EDUCATION ECONOMICS
https://doi.org/10.1080/09645292.2023.2206985

College quality as revealed by willingness-to-pay for college graduates

John J. Green[a], Peter F. Orazem [b] and Nicole S. Swepston[c]

[a]University of California at Davis, Davis, CA, USA; [b]Iowa State University, Ames, IA, USA; [c]Ohio Department of Job and Family Services, Columbus, OH, USA

ABSTRACT

This study measures college quality by the amount by which the college adds to the salary of its students above what the median market value would be for the same majors and student quality. Commonly used national rankings of colleges such as U.S. News and World Report or Forbes are heavily biased by a college's average salaries and the quality of the students it enrolls, and not by the actual value-added by the colleges. Once student quality and mix of majors are controlled, salary differences between elite and nonelite schools largely disappear.

ARTICLE HISTORY
Received 4 August 2021
Accepted 18 April 2023

KEYWORDS
College quality; major; salary; revealed preference; willingness-to-pay; value-added

JEL
I23; I26

Relative wages for bachelor's recipients compared to high school graduates have risen from 49% to 74% between 1980 and 2016.[1] Over that time, the share of high school graduates enrolled in 2- and 4- year tertiary educational institutions increased from 49% to 68%.[2] The annual investment in colleges and universities in 2015 totaled $536 billion or nearly $27 thousand per student. Annual average tuition was almost $19 thousand in public universities and $38 thousand in private schools.[3] Meanwhile, in addition to the services traditionally provided by public and nonprofit private colleges and universities, the share of college training provided by for-profit colleges has risen. A natural question is whether investments of this magnitude are warranted and whether the returns are higher in some schools than others.

To ensure quality, all higher education institutions are subject to external accreditation that aims to ensure minimum standards are met. However, accreditors do not provide relative rankings of the institutions they monitor. To help inform college choice decisions, numerous sources have provided rankings of colleges and universities. Traditionally, these have been based on school selectivity such as *Barron's,* a combination of students' surveys and academic reputation such as the *Fiske Guide*, or exclusively on student surveys such as the *Princeton Review*. More recently, *US News and World Report, Forbes*, and the *Wall Street Journal* have produced rankings based on importance-weighted factors believed to affect college quality. All of these rankings, whether based on reputations, reactions to social or college amenities, or importance weights reflect the tastes of the respondents answering the questions and the researchers conducting the studies. These aggregations of tastes may not accurately reflect the totality of information in the market for academic services.

A related problem clouded commonly used rankings of best places to live. That led to alternative rankings based on local wages and housing prices. If a place is nicer to live in, residents would bid up the price of housing and workers would be willing to live in the area at lower real wages than they

CONTACT Peter F. Orazem pfo@iastate.edu Program for the Study of Midwest Markets and Entrepreneurship and Department of Economics, 267 Heady Hall, Iowa State University, Iowa State University, Ames, IA 50011-1070, USA
A copy of the stata dataset is available at http://www2.econ.iastate.edu/faculty/orazem/college%20rankings_final%20data.dta Supplemental data for this article can be accessed online at https://doi.org/10.1080/09645292.2023.2206985.

would require elsewhere (Rosen 1979; Roback 1982; Gyourko and Tracy 1991). This analysis has been extended to show how, by providing local amenities, cities have been able to increase their productivity by supporting greater density of educated workers (Shapiro 2006) or by attracting educated couples (Costa and Kahn 2000). This study examines whether one could rank the quality of colleges and universities using a strategy similar to the market-based rankings of local quality of life. Colleges that offer higher quality should produce graduates who are atypically productive. Hiring firms will target those atypically high-quality colleges and universities. Graduates of the best colleges and universities should receive higher starting salaries than their contemporaries from lower quality institutions, and so college quality should be apparent through the employer's revealed willingness to pay for their graduates.

Complicating this effort is that universities offer different academic disciplines that command different market wages. According to data provided by the National Association of Colleges and Employers, the gap in starting salaries between the highest and lowest paying college major is 157%. Engineering colleges will generate higher priced graduates than will design colleges, regardless of the quality of instruction. Consequently, any wage-based measure of university quality must control for the specific labor markets that the university's graduates enter. A further complication is that universities do not attract the same students. A university may have high quality graduates, not because the university is adding value to their students, but because the university brings in students who were high quality to begin with. Hence, rankings may differ if they are based on the quality of graduates or on the value-added by the university.

While other rankings have incorporated the salaries of graduates including those issued by the Brookings Institution, the *Wall Street Journal, Forbes,* and *PayScale*, ours has some added features. First, we generate an expected salary for each college based on its mix of majors and the median salary of each major. Value-added college quality is then implied by how much firms are willing to pay for each college or university's graduating class compared to what their graduates would have been paid had they earned the median starting salary for their major. Our preferred value-added index also holds fixed the local cost of living, the demographic composition of the student body, and the performance of its students on standardized exams before attending the university. The resulting hierarchy of schools based on value-added net of student attributes differs greatly from the most commonly used university rankings. Elite private schools look much more like flagship public schools or even modestly priced liberal arts colleges once the quality of the student body is held fixed.

We are also able to generate confidence intervals for college rankings based on the sampling variation of the model parameters. We find that there is substantial variation in the ranking of colleges, especially in the middle of the quality distribution. Only at the extreme upper and lower tails of college quality is there a tight error distribution around the mean college ranking. This corroborates our conclusion that willingness to pay measures suggest that college quality is spread broadly across schools of many types, and it explains why college rankings can vary so much across rating systems or across time.

1. Literature review

College rankings such as the ones produced by *US World and News Report, Forbes*, and the *Wall Street Journal,* or traditional compilations by *Barron's*, the *Fiske Guides* or the *Princeton Review* are meant to add information for consumers of academic services. The coexistence of so many different rankings reflects the lack of agreement on how school quality should be measured. Rankings differ on whether they focus on college inputs such as average faculty salary, library resources, expenditures/student, and student to faculty ratio; college outputs such as graduation rates or salaries of graduates; or college ability to attract students such as average SAT scores, admission rates, or enrollments relative to admissions (Ehrenberg 2005; Monks and Ehrenberg 1999). These measures of quality have then been used by studies examining how college quality affects student learning or earnings (Dale & Krueger, 2002; Long 2010; Thomas & Zhang, 2005), alumni satisfaction with the school (Rothwell 2019), or college graduation (Kelchen and Harris 2012).

It is apparent that these rankings matter. Analyses of the *US World and News Report* rankings of national universities, liberal arts colleges, and law schools (Ehrenberg 2005; Sauder and Lancaster 2006) show that lower rankings are associated with worse outcomes. Monks and Ehrenberg (1999) found that schools receiving poorer rankings became less selective, enrolled students who were less likely to matriculate, and attracted freshman classes with lower average SAT scores. A one-position improvement in the law school ranking was associated with receiving almost 19 additional applications and a 0.18% improvement in the yield rate (Sauder and Lancaster 2006).

Schools take the rankings sufficiently seriously that they alter their behavior in order to influence their scores. Colleges send promotional information to fellow school administrators in an effort to influence their reputational scores. Colleges have altered their admissions policies so that their enrolled students have higher scores on the attributes valued by the rankings (Hossler 2000; Monks & Ehrenberg, 1999; Sauder and Lancaster 2006; Bastedo and Bowman 2010). Some universities have even admitted to falsifying data in order to raise their rankings.[4] In this way, the rankings can become circular, where ranking based on student attributes are used to attract students who share those same attributes. The Black and Smith (2004) finding that better students tend to attend higher ranked schools while weaker students attend lower ranked schools is due in part to schools being ranked at least in part by the SAT or ACT scores of the students they attract and on the selectivity of the college which is also based on those admissions scores.

There is mixed evidence on the link between college quality and student learning or earnings after graduation. It seems clear that higher ranked schools generate more able and higher earning graduates, but at least part of that outcome reflects the abilities the graduates brought with them from high school rather than the value-added by the colleges. The disagreements in findings also reflect differences across studies in how they control for observed and unobserved differences in abilities between graduates. Black and Smith (2004) and Long (2010) found large earnings gains from attending a higher quality college. Thomas and Zhang (2005) found very small gains from attending a higher quality college. Dale and Krueger (2002) argued that the wage premium was due entirely to pre-enrollment student abilities although there was a premium attached to school quality for students from lower-income families.

However, the mixed evidence of whether school quality matters ignores the more fundamental question of whether the school quality measures are actually measuring the quality of the students. Past studies that examined how school quality affects probability of graduation or post-graduation earnings such as Light and Strayer (2000), Dale and Krueger (2002), Thomas and Zhang (2005), Eide, Hilmer, and Showalter (2016), Hoxby (2019), and Dillon and Smith (2020) all rank colleges in whole or in part by the entering SAT or ACT scores or on college selectivity scores that are based on standardized exams and high school rank. While they may use an independent measure of school quality such as *Barron's*, those ratings are themselves driven by the attributes of the students at the college. If school quality is measured by the pre-enrollment qualifications of the students attending the school, it is not clear if the effects attributed to school quality should really be attributed to the quality of the students.

Whether salaries are attributable to school or student quality carries over to the role of majors on salaries. In England, Walker and Zhu (2018) found that salaries differ by major and by the selectivity of the school. Eide, Hilmer, and Showalter (2016) found similar results for the United States. Average salaries for the schools differ by the types of majors offered by the schools, and then by the salaries the schools' majors receive. There are two reasons why graduates within the same major receive different starting salaries depending on the school they attend: the schools differ in their ability to attract talented students, or the schools differ in their ability to train the students they enroll.

Our proposed measure of a college's quality is based on the presumption that the salaries paid by employers reflects the productivity of the schools' graduates. These starting salaries will reflect the markets the schools' students compete in, as reflected by the mix of their majors. High quality schools will be the ones whose graduates earn more than others in the same mix of majors, controlling for the students' abilities before they went to the school.

Cunha and Miller (2014) proposed such a measure of school value-added based on the extent to which graduate earnings exceeded what would be predicted on the basis of individual demographics, standardized test scores, and the socioeconomic composition of the individual's family and high school. Their analysis was confined to public universities in Texas and it did not control for the mix of majors by school. As a result, schools with heavy undergraduate emphasis on engineering or business were rated as high quality while those more oriented toward the arts and humanities were graded less favorably.

Rothwell and Kulkarni's (2015) strategy is similar to ours in that they use the salary a school's graduates would expect to earn based on the average salaries of the mix of majors offered at each school. However, we compare the starting salaries of college graduates with the starting salaries of the college's majors, while Rothwell and Kulkarni make the comparison at midcareer. The logic of using starting salaries is that there is a temporal link between college and student attributes and the salaries of the current graduates, while midcareer salaries lag the college and student data by an average of 15 years. Our results also differ in that the Rothwell and Kulkarni (2015) paper uses the average SAT math score as a control for student ability before entering college, while we also include the SAT verbal and ACT composite scores along with the SAT math score. The latter two measures are much more important in explaining starting salaries, and so their inclusion provides a more complete control for student ability before attending college.

Our study combines the Cunha and Miller (2014) and Rothwell and Kulkarni (2015) strategies by controlling for the value-added by the university after controlling for both the college's mix of majors and the average ability of its students. As we show, using a national sample of colleges and universities including public, private nonprofit, and private for-profit institutions, conclusions regarding college value-added are very sensitive to whether controls for mix of majors and student aptitudes are included in the analysis. Once both of these controls are incorporated, the differences in value-added across colleges and universities become much less pronounced. As one might expect, high quality students in valued majors earn more, regardless of what school they attend. Weak students in low-demand markets earn less.

2. Methodology

We measure the quality of a college's graduates by the average starting salary of the university's graduates compared to the average the graduates would have earned if paid the median starting salary for their majors. Specifically, designate the average wage for graduation cohort t in college k as W_{tk}. The market norm for the college's graduates is

$$W_{tk}^{M} = \sum_{i=1}^{I} \omega_{itk} W_{it}^{M}$$

where ω_{itk} is the share of college k, year t graduates in major i, $\sum_{i=1}^{I} \omega_{itk} = 1$; and W_{it}^{M} is the national median starting salary for graduates in major i in year t. Our assumption is that a college offering average quality will graduate students earning the national median starting salary for its majors. Walker and Zhu (2018) and Eide, Hilmer, and Showalter (2016) have shown that there are significant differences between salaries earned by gradates in the same major across colleges. Higher quality colleges will earn more than the median, while lower quality colleges will earn less. A natural indicator of college quality is the college's average graduate salary relative to the market norm: $RW_{tk} = \frac{W_{tk}}{W_{tk}^{M}}$.

There are alternative measures for college quality, but none share the universality of employer revealed preference through their willingness to pay. Because of differences in grading and admissions policies, outcomes such as GPA or graduation rate are not comparable across colleges. Outcomes that may make sense for one major (passage of licensing exams, accreditation reviews) will not generalize to others (admissions to law or graduate school, placement rates). Average graduate starting salaries are also not comparable across colleges that differ in the mix of majors offered. Colleges whose graduates will enter public or nonprofit sectors such as social work or education will have lower salaries than those focused on engineering. By comparing relative starting salaries paid by employers for the same major, the variation in willingness to pay due to differences in the importance of public versus private versus nonprofit shares of employment demand is controlled.

There are several reasons why a college may have a high value of RW_{tk} that are not related to the college quality. Wages for comparably skilled workers are higher on the east and west coasts and in metropolitan areas than they are in the Midwest and South and in rural areas. These differences reflect market variation in housing and other living costs that would affect nominal wages, independent of any real value-added attributable to skills acquired in college. Because a disproportionate share of college graduates find employment near their college Winters (2020), potentially because large employers focus their recruiting on nearby colleges (Weinstein 2022), starting salaries will reflect local prevailing wages. Consequently, we would want to control for local market variation in wage levels using a vector of regional dummy variables, D_k. In a regression of the form:

$$RW_{tk} = \alpha_0 + D_k'\alpha_D + \varepsilon_{tk}^R \tag{1}$$

the values of the error term, ε_{tk}^R, would reflect the real relative earnings of the school's graduates net of local cost of living variation.

Schools may have higher quality graduates because they attract higher quality students as indexed by a vector of student attributes at the time they enter the university, Q_{kt}. In that case, the university may provide a service to the labor market by signaling higher quality graduates due to their admissions policies, but the university is not adding any additional value. To accommodate that possibility, we add in controls for student quality measured at the time of enrollment in a regression of the form:

$$RW_{tk} = \gamma_0 + D_k'\gamma_D + Q_{kt}'\gamma_Q + \varepsilon_{tk}^V \tag{2}$$

The error term from estimating regression (2), ε_{tk}^V, is the salary of a college's graduates net of the wage attributable to college location and the quality of the students it attracts. Therefore, the error term in (2) represents the value-added by the college to its student's salary at completion.

Colleges may add value to their students using observable inputs into the educational production process or unobservable and perhaps idiosyncratic efficiencies in generating productive graduates. Let Z_{kt} be a vector of the observable inputs for college k. The vector could include measures of faculty or staff quality, resources devoted to students, school infrastructure, and other factors commonly used to identify high quality universities such as the school's selectivity, acceptance rate, and yield rate. That suggests an alternate regression,

$$RW_{tk} = \beta_0 + D_k'\beta_D + Q_{kt}'\beta_Q + Z_{kt}'\beta_Z + \varepsilon_{tk}^{UV} \tag{3}$$

The unexplained college quality due to the idiosyncratic college value-added that is independent of the observable college attributes is measured by the error term in equation (3), ε_{tk}^{UV}. While this measure identifies colleges that produce salaries above and beyond their observed inputs, it overcontrols as a measure of value-added because observed college inputs should also be incorporated. Consequently, our preferred value-added measure is ε_{tk}^V in equation (2).

We will rank all colleges using these four measures of college graduate starting salary relative to the market, the measure controlling only for cost of living, ε_{tk}^R; the student quality controlled measure, ε_{tk}^V; and the idiosyncratic college value-added unrelated to its observed attributes, ε_{tk}^{UV}.

2.1. Generating standard deviation of school ranks

Sampling variation in the parameters from estimation in equations (1–3) can have a large potential effect on the magnitude of the error terms. As one example, consider the role of college major in the willingness to pay for graduates. As Eide, Hilmer, and Showalter (2016) show, returns to college major can vary substantially between more and less selective schools. However, students pick the major based on comparative advantage, and so the returns to the major reflect the match between the individual abilities and the field. Similarly, students pick the school based on their individual observable talents and unobservable preferences. Using Norwegian data, Kirkeboen, Leuven, and Mogstad (2016) found that when they correct for these complex selection problems, returns to a major were substantial, even after controlling for individual ability and school quality. What's more, the major is more important than the institution in defining graduate success. That suggests that rankings that do not control for major and for student ability will generate misleading information on school quality. In addition, random variation in the estimated returns to major and to student backgrounds could have a substantial impact on school rankings.

That means that we need to develop a distribution of possible rankings that reflects the sampling distribution of the parameters in the regressions explaining starting salaries by college. Mogstad et al. (2022) discuss strategies for defining the statistical uncertainty in ranking journals and economics departments based on the existence of a model that relates true quality to a vector of observable factors. While their applications do not have the data necessary to conduct a bootstrap on the parameters of the model, we are able to follow the nonparametric bootstrap method that they discuss. First, suppose that the true rank is given by:

$$R_{tk} = X'_{tk}\theta_X + \varepsilon^X_{tk} \tag{4}$$

where R_{tk} is the true rank for college k, X_{tk} is a vector of college characteristics that define quality, θ_X is a vector of estimated parameters, and ε^X_{tk} is a random error. The predicted rank is given by:

$$\widehat{R_{tk}} = E(R_{tk}) = X'_{tk}\widehat{\theta_X} \tag{5}$$

Sampling variation in the parameter vector, θ_X, will create variation of the predicted rank around the true rank. For each of equations (1-3), we generated 100 random samples of the college rankings with replacement to derive 100 draws of the parameter vector which were then used to produce 100 estimates of each college's rank. These bootstrapped rankings will provide an estimate of the variation around each college's mean ranking.

3. Data

The empirical analysis requires information on average starting salaries for graduates by college, the distribution of graduates by major by college, and the starting salaries by major across colleges. We will also require information on student quality and academic inputs by college. We compile this information from three main data sources.

The National Association of Colleges and Employers (NACE 2022) is a nonprofit organization with over 50 years of experience in collecting starting salary data. Beginning in 1960, NACE began publishing Salary Surveys, which contain information on starting salary offers among institutions that were NACE members. Over time, NACE has revised its methods in order to mitigate self-selection bias and attain a more robust and representative sample. In 2012, NACE began using data from employers rather than colleges and universities which lowered the selection on school response. In addition, their focus shifted from starting salary offers to actual starting salaries, reducing the measurement error in their reported salaries by major. NACE was able to accomplish this shift in methodology by using salary data gathered by Job Search Intelligence (JSI) (Koncz 2016). JSI uses business establishment surveys and data from governmental agencies and college institutions to formulate its salary dataset (JSI 2023).

Their estimates represent the salary offers for 1.4 million recent college graduates and 400,000 employers (*NACE* 2013).[5]

We use the 2013 edition to generate our measures of W_{it}^{M}. Note that an advantage of using starting salaries is that it holds work experience constant at 0 for all colleges so that our results are not clouded by differences in work experience across schools or majors. It also ensures that the data on school and student attributes correspond to the college graduate salaries from the same year. The choice of year allows us to compare our results with several other college rankings that were released that year.

Information on average starting salaries of the 2013 bachelor's degree graduating class by college, W_{tk}, was provided by *PayScale* who survey employed graduates about their starting pay.[6] Rothwell and Kulkarni (2015) showed that the *PayScale* earnings by major corresponded closely to comparable data reported by the US Census (correlation 0.85). Starting salaries by college could be compared to reported starting salaries for Texas Colleges, and again they matched reasonably well (correlation 0.73). They concluded that the *PayScale* data were the most promising source of consistently collected data on college graduates across all states.[7]

Information on each college's 2013 graduating class by major is available from the Integrated Postsecondary Education Data System (*IPEDS*). As a condition of their receipt of federal aid, academic institutions are required to report a broad range of information on their students, staff, costs, and infrastructure. We use *IPEDS* to generate measures of our region dummies, D_k, including the college's state and city size (rural, town, urban or metro). The same source provides measures of student quality and demographics, Q_{kt}, including ACT scores and SAT scores, race, ethnicity, and national origin of the student body. Institutional quality and inputs, Z_{kt}, includes an index of selectivity, the share of applicants accepted, the share of accepted applicants who enroll, and measures of instructional and non-instructional expenditures per student. Instructional expenditures represent an average of 37% of total expenditures. Appendix Table 1 contains the descriptive statistics for these various variables. We were able to match *PayScale* data with contemporaneous *IPEDS* and *NACE* information for 1047 universities.

Table 1 illustrates the range of the average salaries across colleges and the importance of controlling for the mix of majors in assessing market willingness to pay for graduates. The highest average salary is at the US Military Academy, 184% higher than the lowest paid graduates at Coker College. Controlling for the mix of majors reduces the variance only modestly with a top-to-bottom range of 142%. The US Military Academy and US Naval Academy also rank 1 and 2 after controlling for market salaries by major, with the Air Force Academy ranked 12th. For that reason, it may be appropriate to exclude the military academies from computing the relevant range, as the salary is fixed by rank upon commission and not by major. As a result, the military academies have the highest paid history, psychology and humanities graduates in the nation. Excluding the military, the range in starting salary is 155% and the range in salary relative to market norm for the mix of majors is 83%. Therefore, mix of majors explains 46% of the variation in average starting salary across colleges and universities.

In Table 1, it is clear that the mix of majors can greatly alter the pay of graduates. Engineering colleges such as Rose-Hulman and the Colorado School of Mines have top ten average starting salaries, but they are barely paying above the market norm for the majors they produce. However, there is considerable variation left to explain after controlling for mix of majors, the task we turn to next.

After calculating our four value-added measures, we further analyze our results by controlling for college type. The data used comes from the 2000 Edition of The Carnegie Classification of Institutions of Higher Education, which categorizes institutions based upon degrees offered and whether they are for-profit or non-profit.

4. Results

In Table 2, we report the results of our estimation of equations (1-3). The first equation controls for local wages and prices, presuming that most graduates find work in the same region as their college.

Table 1. Ten best and worst ranked universities using salary of graduates as a measure of value-added.

		Average starting salary		Relative wage ratio	
Rank[a]	Institution	Rank	Starting salary	Rank	RW_{tk}
1	US Military Academy	1	$76,000	1	1.576
2	US Naval Academy	2	$72,200	2	1.410
3	Mass. Institute of Tech.	3	$68,400	16	1.214
4	California Institute of Tech.	4	$67,400	10	1.245
5	Harvey Mudd College	5	$66,800	14	1.220
6	US Air Force Academy	6	$65,400	12	1.242
7	Colorado Schools of Mines	7	$64,200	197	1.018
8	Loma Linda University	8	$63,400	4	1.318
9	Thomas Jefferson University	9	$63,200	6	1.270
10	Rose-Hulman Institute of Technology	10	$62,300	160	1.035
1038	St. Bonaventure University	1038	$31,100	1039	0.692
1039	Univ. of Arkansas-Pine Bluff	1039	$30,900	1008	0.734
1040	Berea College	1040	$30,800	1018	0.725
1041	Daemen College	1041	$30,700	1044	0.663
1042	Benedict College	1042	$30,200	1038	0.698
1043	Everest College – Phoenix	1043	$30,000	1043	0.669
1044	Howard Payne University	1044	$29,700	1014	0.729
1045	Wayne State College	1045	$29,300	1041	0.678
1046	College of the Ozarks	1046	$28,400	1042	0.671
1047	Coker College	1047	$26,800	1045	0.662

[a]Out of 1047 institutions.

Table 2. Regressions explaining variation in RW_{tk}: The college average salary relative to the market norm for the 2013 graduating class.

Variable	Equation (1)	Equation (2)	Equation (3)
Rural	−0.023	0.009	0.006
	(1.20)	(0.61)	(0.43)
Town	−0.018**	0.004	0.003
	(2.08)	(0.56)	(0.47)
Urban	−0.012	0.003	0.001
	(1.43)	(0.40)	(0.19)
% Male		0.121**	0.104**
		(4.71)	(4.11)
% White		0.029	−0.031
		(0.73)	(0.80)
% Black		0.073*	0.020
		(1.71)	(0.47)
% Hispanic		0.148**	0.113**
		(2.74)	(2.15)
% Asian		0.525**	0.403**
		(6.49)	(5.04)
% Foreign		0.117*	−0.038
		(1.66)	(0.54)
SAT Math /100		0.002	0.002
		(0.52)	(0.49)
SAT Critical Reading/100		0.031**	0.013
		(2.51)	(1.07)
ACT Composite/100		0.714**	0.653**
		(2.86)	(2.64)
Acceptance Rate			0.032*
			(1.68)
Yield			0.052**
			(2.45)
Ln (Instruction Exp/Student)			0.056**
			(5.99)
Ln (Non-Instruction Exp/Student)			−0.003
			(0.42)
N	1047	1047	1047
R squared	.23	.50	.54

All regressions include fixed effects for the state of residence.
**t*-statistic in parentheses.

The first column shows that state fixed effects and size of college community explain 23% of the variation in starting salary. In practice, correcting for local cost of living has only a modest effect on college ranking, but colleges on the coasts tend to fall in rank while colleges in more rural areas rise.

The second equation controls for the quality of the students in the college. In addition to the controls for local and regional cost of living, we add measures of student quality and the demographic composition. Interestingly, colleges with greater fractions of Black, Hispanic and Foreign graduates have higher average relative salaries, as do schools with larger shares of male graduates. Schools whose students arrive with higher verbal SAT scores and higher average ACT scores also have higher relative pay. Adding these measures raises the share of variance in pay explained to 50%. The error terms from this equation, ε_{tk}^{V}, represent the pay received by graduates of a university controlling for their major and controlling for student attributes and local cost of living. As such, we view ε_{tk}^{V} as the valued added by the school, where value is the additional salary received by their students above and beyond what they would get from the skills they brought before coming to college.

Equation (3) adds in the college's investments in the quality of its academic programs through expenditures and its enrollment management policies. Interestingly, higher per student expenditures on instruction raise graduate salaries relative to market norms, but per student expenditures on noninstructional inputs do not translate to higher salaries. Graduates of schools that get a greater yield from their admitted students are paid more, but so are graduates of schools that admit a larger share of their applicants. Holding student quality constant through information on test scores, there does not appear to be a further value to being more selective on admissions policies. Also interesting is that these measures of school's infrastructure and enrollment policies only explain 4 percentage points more of the starting pay. Much more of the variation is explained by the attributes the students bring into the school than what the school provides in inputs.

The error term to equation (3), ε_{tk}^{UV}, represents the starting salary relative to market norms students receive that is not explained by either the students' or the school's attributes. This is the idiosyncratic value the school adds, whether by being more efficient with its resources or by fostering uniquely productive matches between its faculty and staff and its students.

In Tables 3 and 4, we summarize the top and bottom 10 colleges using the results of Table 2 to define quality. Our preferred measure is the value-added by the college, ε_{tk}^{V}, holding constant student quality and demographics and controls for local cost of living. The top ten colleges based on college value-added to its graduates' earnings relative to their major's market salary proved quite interesting. Two military academies; two Catholic colleges (Molloy and Holy Family); an

Table 3. Ten best ranked universities using alternative willingness-to-pay measures of value-added.

		Relative wage ratio		Real relative wage			Value-added			Idiosyncratic value-added		
Rank[a]	Institution	Rank	RW_{tk}	Rank	ε_{tk}^{R}	Rank ± σ	Rank	ε_{tk}^{V}	Rank ± σ	Rank	ε_{tk}^{UV}	Rank ± σ
1	US Military Academy	1	1.576	1	0.618	0.000	1	0.505	0.000	1	0.444	0.000
2	Molloy College	5	1.300	5	0.335	1.258	2	0.367	0.578	2	0.350	0.595
3	US Naval Academy	2	1.410	2	0.407	0.744	3	0.340	1.842	7	0.265	4.818
4	Thomas Jefferson University	6	1.270	6	0.331	1.314	4	0.302	1.282	12	0.237	5.347
5	Art Center College of Design	3	1.385	3	0.359	1.149	5	0.301	2.188	6	0.266	3.320
6	Kettering College	34	1.141	20	0.234	5.360	6	0.287	1.889	11	0.238	4.161
7	Loma Linda University	4	1.318	8	0.304	1.519	7	0.287	1.370	10	0.241	4.195
8	Holy Family University	55	1.114	44	0.175	9.728	8	0.264	1.713	9	0.248	2.343
9	LIU Brooklyn	23	1.177	33	0.201	5.379	9	0.248	1.604	13	0.222	2.901
10	Charter Oak State College	20	1.185	26	0.218	9.301	10	0.246	2.850	4	0.297	1.930

[a]Out of 1047 institutions.

Table 4. Ten worst ranked universities using alternative willingness-to-pay measures of value-added.

		Relative wage ratio		Real relative wage			Value-added			Idiosyncratic value-added		
Rank[a]	Institution	Rank	RW_{tk}	Rank	ε^{R}_{tk}	Rank ± σ	Rank	ε^{V}_{tk}	Rank ± σ	Rank	ε^{UV}_{tk}	Rank ± σ
1038	Haverford College	616	0.898	645	−0.030	60.024	1038	−0.183	5.559	1040	−0.188	3.982
1039	King's College	1012	0.731	1037	−0.207	5.520	1039	−0.183	5.683	1043	−0.194	4.174
1040	Marietta College	999	0.746	990	−0.155	22.662	1040	−0.185	18.205	1044	−0.198	9.648
1041	Nichols College	1019	0.725	1044	−0.259	2.163	1041	−0.195	4.179	1034	−0.166	9.635
1042	Bluffton University	1040	0.679	1041	−0.221	5.898	1042	−0.197	4.167	1041	−0.191	5.018
1043	CUNY Bernard M Baruch College	597	0.901	828	−0.075	36.936	1043	−0.199	12.223	1020	−0.142	26.766
1044	Sullivan University	1047	0.644	1036	−0.207	12.935	1044	−0.205	9.001	1036	−0.177	16.371
1045	Daemen College	1044	0.663	1046	−0.301	0.474	1045	−0.231	1.845	1045	−0.227	2.398
1046	St. Bonaventure University	1039	0.692	1045	−0.266	1.077	1046	−0.243	1.120	1046	−0.242	0.986
1047	Culinary Inst. of America	1046	0.652	1047	−0.313	0.359	1047	−0.302	0.100	1047	−0.371	0.100

[a]Out of 1047 institutions

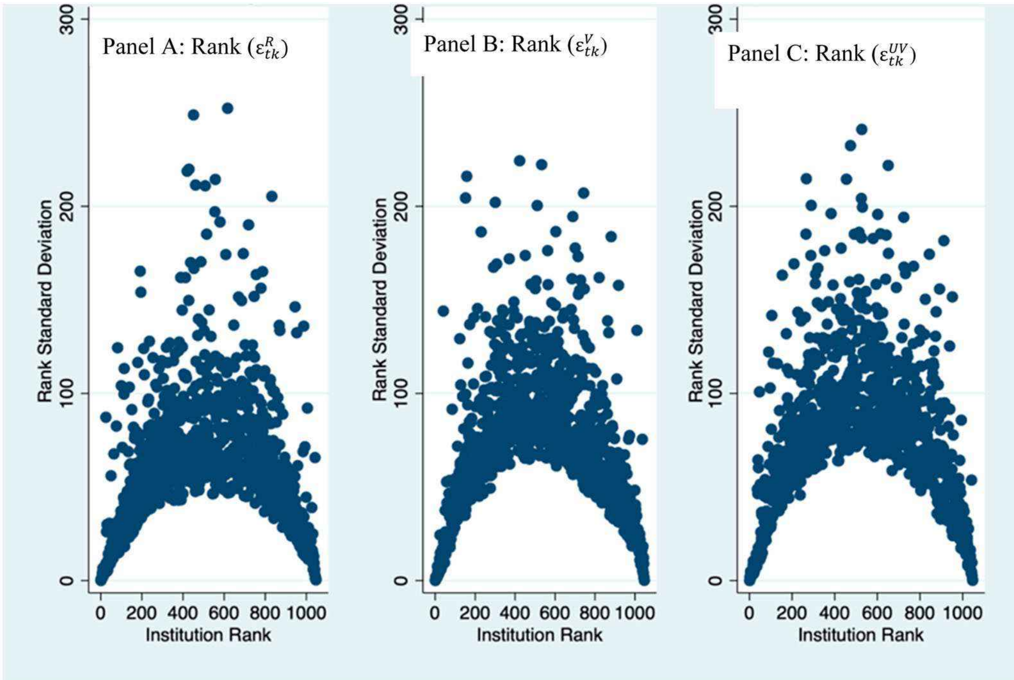

Figure 1. Bootstrapped standard deviation of college ranking by expected rank, $\widehat{R_{tk}}$. Notes: Standard deviations of rankings of each college derived from the coefficients from 100 draws with replication of the rankings based on estimating equations (1-3), using the PayScale and National Association of Colleges and Employers data to measure college salary relative to the market wage for their majors. Standard deviations in Panel A use the Real Relative Wage, ε^{R}_{tk}, in equation (1). Panel; B uses the Value-added, ε^{V}_{tk}, in equation (2). Panel C uses the Idiosyncratic Value-Added, ε^{UV}_{tk}, in equation (3).

online school (Charter Oak State College); two Seventh Day Adventist colleges (Loma Linda and Kettering College); a design school; and two private nonprofit colleges. Other than the military academies, none of these schools are even listed among the Forbes top 660 schools. It is clear that many relatively unheralded schools actually have among the highest starting salaries after controlling for mix of major, geographic differences and student quality. Meanwhile, 4 of the bottom 10 institutions in Table 4 are on the Forbes list of the best colleges. The bottom 10 includes 2 Catholic schools, a Mennonite college, one online school, a culinary school, a public university and 3 private colleges. Clearly, quality as measured by willingness to pay for the graduates of colleges and universities can be found in unusual places.

Some of the lowest ranked schools in Table 4 actually have reasonable average starting salaries (Haverford, Baruch). However, once pre-enrollment student quality is controlled, it is apparent that these schools added little additional value in the form of better starting pay for their graduates. Students came to those colleges with the attributes necessary to earn high salaries already in hand as opposed to acquiring those skills after enrolling.[8]

Also included in Tables 3 and 4 are the standard deviations of the ranks for the ten worst and best schools when we generate 100 bootstrapped random samples of the colleges with replacement. We do this for each of the three equations (1–3). Figure 1 illustrates how the variance in ranking varies by college rank. The inverse-U shape of the relationship between the standard deviation of the college's ranking to its mean ranking implies that there is the greatest uncertainty about the college's true rank if it is in the middle of the quality distribution. For example, using our value-added measure, schools ranked in the top 20 have an average standard rank deviation of 3.2 ranks. Schools ranked in the bottom 20 have an average standard rank deviation of 12.4 ranks. Schools ranked in the middle 20 have an average standard rank deviation of 102, or a 95% confidence interval of ±204 ranks.

Therefore, for colleges whose ranking lie between 50 and 950, the 95% confidence interval of their ranking with variation due solely to uncertainty about the parameter values spans an average of 320 ranks. This is true for rankings that only control for major and local labor market, for rankings that control for student quality, and for rankings that arguably over-control for college expenditures and admissions policies. Colleges that locate in the middle of the distribution are equivalent on observable quality based on value-added.

4.1. How does the willingness to pay ranking correspond to other rankings?

There are many rankings of colleges and these are very important to students' and parents' decisions regarding where to apply. To the extent that these rankings are transparent in what they measure, they will not induce poor choices. However, they may not provide accurate information on the value-added by the college to the student's future earnings potential.

We compare our rankings based on observed salaries, salaries relative to mix of major, and the three error terms from equations (1–3). Our comparison rankings include three chosen rankings systems because they include a large number of colleges and because they present clear numerical hierarchies of academic institutions. *Forbes* 2018 listing of *America's Top Colleges* includes rankings of 660 institutions, 537 of which matched our group. The 2018 *US News and World Report* ranks 301 *Best National Universities* with 199 ranked progressively and the last in an unranked group. The magazine also rates other colleges and universities within other categories defined by region, specialization, and graduate and professional programs, but those schools are excluded because the rankings are not comparable to the national university ranking. To make comparisons easier, we invert these two ranking so that higher values mean a higher quality rating.

The *Forbes* and *US News and World Report* rankings are not meant to provide guidance on the added value of the school net of the student's attributes. In fact, higher ability students raise the ranking. This is justified in the sense that having higher ability peers may enhance the learning environment of the college, but it does confuse inputs and outputs if one is interested in the value-added by the school. The *Brookings Institution's* 'Beyond College Rankings' uses value-added measures very similar to ours. The most similar one to ours uses the *PayScale* mid-career salary while we use the *PayScale* starting salary. The two measures are correlated at 0.75. The *Brookings* measure use a measure of major salary based on the America Communities Survey which is correlated with our *NACE* measure at 0.82. These differences, plus differences in model specification result in outcomes that are consistent with ours, but as we show, even minor differences in measurement and model result in low correlation among school quality measures.

Cunha and Miller (2014) also use salaries of graduates as their outcome measure, but they confine their analysis to public institutions in Texas. The strength of their analysis is that they have data on individual salaries, but they lack data on major which has proven to be a key factor explaining salary differences in ours and others analysis. Similar to our approach, they use the error terms from an equation similar to (2) but with the individual graduate as the unit of observation and with actual rather than relative salary as the dependent variable.

In Table 5, we show the interrelationships between the 5 measures that we employ in our analysis. Any of the rankings might be a plausible metric for college quality, whether we are interested in the average graduate starting salary; the average salary relative to others in the same major (RW_{tk}); the average salary relative to others in the same major and local market (ε_{tk}^{R}); the average salary relative

Table 5. Relationship between alternative willingness-to-pay rankings.

	Relative wage ratio RW_{tk}	Real relative wage ε_{tk}^{R}	Value-added: ε_{tk}^{V}	Idiosyncratic value-added ε_{tk}^{UV}
Starting Salary	0.836	0.749	0.522	0.521
Relative Wage: RW_{tk}		0.877	0.706	0.680
Real Relative Wage: ε_{tk}^{R}			0.805	0.776
Value-Added: ε_{tk}^{V}				0.963

Table 6. Correlations between willingness-to-pay rankings and other college rankings.

	Brookings: Value-added – Midcareer Earnings[a] (848)	Forbes: America's Top Colleges[b] (537)	USNWR: National Universities[b] (199)	Cunha-Miller[c] (26)
Starting Salary	0.670	0.532	0.580	0.146
Relative Wage: RW_{tk}	0.541	0.649	0.699	0.184
Real Relative Wage: ε^{R}_{tk}	0.458	0.555	0.582	0.202
Value-Added: ε^{V}_{tk}	0.189	0.097	−0.085	0.068
Idiosyncratic: ε^{UV}_{tk}	0.196	0.071	−0.139	0.160

[a]We use the VA_sal measure reported by Brookings (Rothwell, 2015).
[b]These are listed in Forbes (2017) and in U.S. News & World Report (2018) as rankings rather than ratings, and so the rankings were reversed so that higher numbers meant higher quality.
[c]Value-added measure taken from Table 3, column 5 of Cunha and Miller (2014). Correlation reflects only public universities in Texas.
Number of matches listed in parentheses.

to others in the same major, local market, and college entry qualifications (ε^{V}_{tk}); or the average salary controlling for all the above plus college resources expended (ε^{UV}_{tk}). The last measure is an overcontrol Using the fact that the bivariate R^2 is the square of the simple correlation coefficient, we can assess how much of the college's average graduate salary is attributable to the various components. About 56% of the variation in starting salary across colleges can be attributed to its real wage relative to the market norm for its majors, ε^{R}_{tk}. College valued added, ε^{V}_{tk} explains about 27% of the variation in college starting salary, and almost all of the college's value-added is independent of its expenditures and its enrollment management policies.

Turning to Table 6, the *Brookings, Forbes,* and *US News and World Report* rankings are highly correlated with the average salary of the graduates. The correlation rises for the *Forbes,* and *US News and World Report* when we adjust the starting salaries for the mix of majors in the college. The correlation between the *Brookings* rating and RW_{tk} falls. All three rankings fall further as we adjust for local prices, but the correlation drops dramatically after we control for the quality of the students in the school. The closest correspondence is with the *Brookings* measure which is correlated at 0.2. Our university value-added measure explains less than 1% of the variation in the *Forbes,* and *US News and World Report measures,* and the latter ranking is negatively correlated with our measure.

The Cunha and Miller (2014) rankings are restricted to public universities in Texas, but their methodology and ours are similar except that they cannot adjust university salaries for the mix of majors offered. Whether due to that difference or to differences in sample of schools, use of individual versus school as unit of observation, or choice of student and school controls, our measures are never correlated above 0.2.[9]

What Tables 5 and 6 show is that reasonable attempts to measure school quality systematically result in very different measures of school quality. Because colleges and universities are complex institutions with heterogeneous objectives, the variation in school quality measures is not surprising. However it does suggest that any ranking must be clouded by considerable measurement error, as we will demonstrate later.

There are several studies that have examined how improvements in school quality alter the probability that a student graduates and the amount that they earn after graduation. Some, such as Light and Strayer (2000), Dale and Krueger (2002), and Thomas and Zhang (2005), found very small gains from attending a higher quality college. Hoxby (2019) finds positive earnings gains, although about two-thirds are due to selection of higher ability students into higher ranked colleges. Dillon and Smith (2020) find that quality matters for both college completion and, in most cases, for earnings later in life. All of these studies rely on measures of college quality that are largely or entirely based on information about the quality of the students the college attracts. Consequently, at least some of the reported college quality effect should be credited to the quality of the students apparent in test scores and high school performance that predate any value-added by the colleges.

4.2. How different are college salaries after controlling for majors, cost of living and student quality?

Perhaps the best way to illustrate how important student quality is to a school's ranking is to examine self-selected groupings of schools. Schools that compete with one another in athletic conferences are made homogenous as they level the playing field, setting rules related to student academic ability, academic standards, and graduation rates. In Table 7, we show the average ranking of these schools under our various ranking methods. Based on starting salary alone, the Ivy League is clearly the highest rated while the Southeastern Conference is a cut below the other Power 5 conferences. The Ivy League rises when we control only for mix of majors, but this gain reverses when we control for the higher cost of living in urban, east coast locations. Meanwhile, the disadvantaged position of the Southeastern Conference disappears when we control for the lower cost of living in their more rural, southern locations. Once we control for the pre-enrollment student quality measures, the differences in starting salaries relative to major largely disappear between all the conferences. The advantage in ranking attributed to the Ivy League schools reflects the students they attract.

A similar story holds when we look at other groupings of schools. Land Grant Universities have higher starting salaries, but that reflects their larger business and engineering programs. Controlling for majors, the gaps between Liberal Arts Colleges and Land Grant schools disappears. Controlling for student quality at time of entry, the differences between Historically Black Colleges, and Private For-Profit schools and the Liberal Arts and Land Grant Programs are largely eliminated. All the school groups have comparable value-added, holding the quality of their students constant.

4.3. Alternative data sets

There are alternative measures of starting salary by college and market value of college majors. In this section, we reexamine our results using alternative data sources. The essential result that there is substantial variation in college rankings due to dispersion in the parameters relating school and student attributes to the earnings of the school's graduates.

College Scorecard (U.S. Department of Education 2022) reports the median salary of individuals 10 years after first receiving federal Title IV grants or loans. The data are taken from federal tax returns of individuals with unique W-2 forms or with self-employment earnings from Schedule SE. About 37% of college students receive federal loans and 42% receive federal grants, and so these data cover

Table 7. Average willingness-to-pay college rank[a] by college athletic conference and by type of college.

Conference	Number of Institutions	Average Rank by willingness-to-pay measure				
		Starting Salary	Relative Wage Ratio RW_{tk}	Real Relative Wage ε_{tk}^{R}	Value-Added: ε_{tk}^{V}	Idiosyncratic Value-added ε_{tk}^{UV}
Atlantic Coast Conference (ACC)	15	218	229	168	441	441
Big East Conference	10	217	269	248	329	289
Big Ten Conference	14	183	220	161	394	389
Big 12 Conference	10	282	372	246	364	400
Pac-12 Conference	12	209	187	263	496	563
Southeastern Conference (SEC)[b]	13	416	491	221	374	395
Ivy League College Type	8	50	33	46	479	539
Liberal Arts	109	486	355	354	452	484
Land Grant Universities	51	264	310	272	401	402
Private, For-Profit Colleges	18	528	628	621	550	447
Historically Black Colleges and Universities (HBCU)	42	682	676	620	460	524

[a]Out of 1047 Institutions.
[b]Louisiana State University (LSU) not included due to incomplete data.

about two-thirds of college students (Hanson 2021). More importantly, only 64% of college students graduate within 6 years (NCES 2022), and so nearly a third of the *Scorecard* data represents college dropouts. The sample selection matters. Average *Scorecard* salary is $35 thousand, 16% below the average *PayScale* salary of $42 thousand that only includes college graduates. Nevertheless, the series track one another reasonably closely. The correlation between the *PayScale* and *Scorecard* measures of a college's starting salary is 0.75.

The *College Scorecard* only releases earnings information when the data include at least 30 students. For that reason, most of the data on earnings by school and major are listed as 'Privacy Suppressed'. However, Carnevale, Cheah, and Hanson (2015) reported data on earnings by bachelor's degree recipients by major for 2013 that they compiled from pooled 2009 to 2013 editions of the *American Community Survey*. Median earnings are reported for year-round, full-time workers between the ages 21 - 59. These data are not starting salaries, and so they overstate relative market wages for newly minted graduates. Average value of the *ACS* major salary is $58.5 thousand compared to $45 thousand for the *NACE* measure. However, the two series are reasonably highly correlated at 0.82. The relative wage, $RW_{tk} = \frac{W_{tk}}{W_{tk}^M}$ computed using the *Scorecard* data for the college graduate salary W_{tk}, and the *ACS* for the weighted major salary W_{tk}^M, had a mean value of 0.60 compared to 0.92 for the earlier measure. The ratios demonstrate that the *Scorecard* data likely underestimates the salary at graduation and the *ACS* data overstates the market salary for majors at graduation, biasing downward the average salary relative to market by about one-third. Measurement error in this version of the relative wage compared to the earlier measure leads to a smaller correlation in relative wages at 0.53.

We replicated the regressions in Table 2 using this version of the relative wage based on *Scorecard* and *ACS* data. The results are reported in the appendix. The R^2 with the new relative wage measure was always lower than the corresponding specification in Table 2, presumably because of the greater measurement error from the noisier measures of each college's starting salary and the corresponding starting market wage for the college's majors. In the Appendix Table, we found no significant differences in coefficients between the two versions of equation (1). When we apply equation (2) to the new data, the only significant difference was a reversal from a positive to a negative coefficient on the %Black. For estimates of equation (3), there is only one significant sign reversal from positive to negative on the Acceptance Rate.

The rankings based on the new data are reported in the Appendix. While the *Scorecard* data do not include the military academies, of the remaining 8 original top ranked schools using the value-added specification in Table 3, 3 were in the top 10 using the new data and 2 more were in the top 25 (upper 2.5%). Of the bottom 10 in Table 4, 7 were in the lower 25% of schools ranked using the new data.

The key takeaway from the earlier analysis, as summarized in Figure 1, is that random variation in the coefficients that relate schooling outcomes to student and school attributes can have a huge effect on rankings. Figure 2 shows the results. Other than the very best and very worst schools, the rankings are essentially meaningless. In fact, the less precise coefficients when we used the *Scorecard* and *ACS* data added even more noise to the rankings. At the median rank, the 95% confidence interval is ±222 ranks.

4.4. What types of colleges have the highest value-added?

Even within our college groupings, there is considerable heterogeneity in the attributes of the best and worst schools. While the online appendix has the rankings for all 1047 schools, we can make some generalizations regarding which schools seem to have better value-added. One set of indicators reflects the type of programs the school offers, using the Carnegie Classification of Institutions of Higher Education. We distinguish between Doctoral, Master's, Bachelor's degree and liberal arts

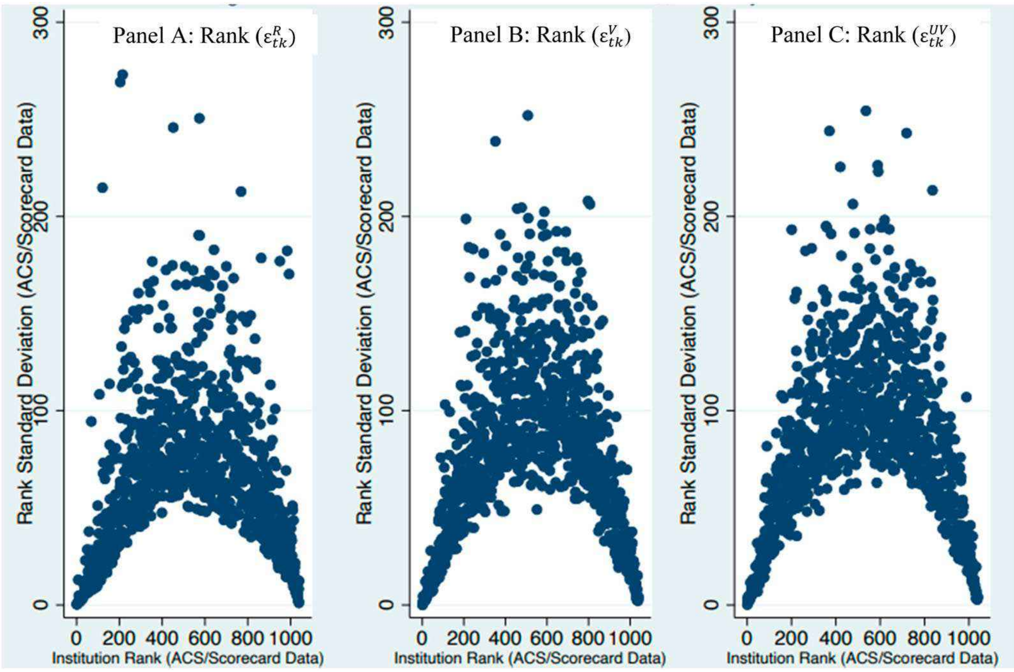

Figure 2. Bootstrapped standard deviation of college ranking by expected rank, $\widehat{R_{tk}}$. Notes: Standard deviations of rankings of each college derived from the coefficients from 100 draws with replication of the rankings based on estimating equations (1–3), using the Scorecard and American Community Survey data to measure college salary relative to the market wage for their majors. Standard deviations in Panel A use the Real Relative Wage, ε^{R}_{tk}, in equation 1. Panel; B uses the Value-added, ε^{V}_{tk}, in equation (2). Panel C uses the Idiosyncratic Value-Added, ε^{UV}_{tk}, in equation (3).

colleges with more specialized colleges being the base. An open question is whether undergraduates benefit from the presence of graduate programs or if the focus on undergraduates and liberal arts colleges improves learning. The second set of factors represent the type of institution: private for-profit, private non-profit, and public, with the latter being the base. The third group includes controls for whether the institution has a medical school, whether it is a land grant school, or whether it is an historical Black college.[10]

We also add two measures made available by Third Way. The first measures the graduation rate for each institution, using newly available data that includes full time, part-time and transfer students. The second is an estimate of the fraction of graduates earning more than the average for someone with a high school degree. The former measure can be viewed as a check on selection so that high average salaries were not masking a weeding out of weaker students. The second is a check on our *PayScale* data as we should see that high average salaries are sufficiently distributed across graduates so that a larger fraction earn more than would someone who did not go to college.

The results are in Table 8. While we include results for all five measures, our interest is in column 4 where we identify the schools having the highest value-added relative to the market salary for the majors it offers and the quality of the students it serves. Graduate programs do not help undergraduates. Offering PhD or Master's degree programs lowers relative starting pay by 3-4%. In contrast, Grunig (1997) found that expenditures on research and more robust graduate program were strongly tied to the *US News and World Report* undergraduate reputation. Liberal arts college graduates earn 3% more than graduates of more specialized undergraduate programs. There is no difference in value-added between public, private nonprofit and private for-profit schools.[11] Land grant schools and historically black colleges add 3% to their graduates relative starting pay. The value-added is reasonably broadly shared, so that a 10 percentage point increase in the share earning above a high school salary results in a 0.013 increase in the salary premium relative to the

Table 8. Regressions identifying factors that affect alternative measures of willingness-to-pay for college graduates.

Dependent Variable	1 Starting Salary	2 Relative Wage Ratio RW_{tk}	3 Real Relative Wage ε_{tk}^{R}	4 Value-Added: ε_{tk}^{V}	5 Idiosyncratic Value-added ε_{tk}^{UV}
Constant	21506.6**	0.67**	−0.16**	−0.04*	−0.05**
	(18.24)	(27.29)	(7.10)	(1.70)	(2.30)
Doctoral	−2036.9**	−0.04**	−0.03**	−0.05**	−0.04**
	(2.85)	(2.70)	(2.18)	(3.64)	(3.41)
Masters	−4773.5**	−0.07**	−0.07**	−0.04**	−0.04**
	(7.34)	(5.48)	(5.30)	(3.74)	(3.77)
Liberal Arts	1521.2**	0.06**	0.05**	0.03**	0.03**
	(2.80)	(5.68)	(4.94)	(3.54)	(3.04)
Bachelors	−5687.2**	−0.10**	−0.07**	−0.06**	−0.05**
	(7.94)	(6.83)	(5.08)	(4.28)	(4.20)
Private Non-Profit	−565.7*	−0.01**	−0.01*	−0.01	−0.01
	(1.77)	(1.97)	(1.91)	(1.08)	(1.30)
Private For-Profit	614.9	<0.01	−0.01	−0.02	0.01
	(0.56)	(0.20)	(0.71)	(1.26)	(0.46)
Medical School	−401.5	<0.01	0.02**	<0.01	−0.02*
	(0.83)	(0.21)	(2.18)	(0.09)	(1.94)
Historically Black	4976.9**	0.06**	0.04**	0.03**	0.02
	(6.70)	(3.60)	(2.99)	(2.42)	(1.44)
Land Grant	552.0	0.01	0.02	0.03**	0.03**
	(0.92)	(0.63)	(1.42)	(2.71)	(3.00)
% Earning More than	33160.4**	0.24**	0.16**	0.13**	0.17**
H.S. Grad. Avg.	(18.27)	(6.31)	(4.64)	(3.98)	(5.40)
Graduation Rate	3595.0**	0.26**	0.17**	−0.01	−0.04*
	(2.89)	(10.17)	(6.95)	(0.64)	(1.72)
N	1016	1016	1016	1016	1016
R-Squared	.496	.396	.304	.057	.065

market. There is only modest evidence of selection so that a higher completion rate resulted in a small decrease in average salary. The graduation rate is uncorrelated with our preferred value-added measure in column 4.

While there are modest differences in value-added between colleges, the overall conclusion is that differences are small and that able students will earn about the same amount regardless of what college they attend. Toutkoushian and Smart (2001) found that the students' perceptions of the amount learned did respond to college expenditures. That suggests that the true value-added by colleges may be more subjective than objective.

5. Conclusions

Given the rising concerns about rising college costs and student debt, the public should have information on the earnings graduates can expect from a college. However, our analysis shows that half of the variation in college graduate starting salaries is attributable to the mix of majors offered by the school, local cost of living, and the pre-enrollment quality of the students the school admits. In other words, the colleges themselves add less than half of the value of their graduates. Because the most popular national college rankings are so heavily weighted toward the pre-enrollment quality of their undergraduates (SAT and ACT score, high school rank, measures of school selectivity) the rankings reflect the quality of the students at the school rather than the actual quality of instruction provided by the school.

Controlling for major and student quality plus local variation in cost of living generates a hierarchy of schools based on the added value provided by the college, independent of the quality of the students it attracts, the majors it offers, or its location. The commonly used national rankings are virtually uncorrelated with the resulting value-added index. If parents or prospective students want to know which colleges will provide the learning environment that will raise their value on the labor market, the national rankings do not do that.

The rankings on the basis of value-added show great heterogeneity in the types of schools that score well or that score poorly. There are some small advantages to some types of schools. Holding student attributes and majors fixed, undergraduates earn more on average in liberal arts colleges and schools without graduate programs. There are advantages to getting degrees from land grant universities and from traditionally Black colleges. However, there are no systematic differences in starting salaries between public, private nonprofit, or private for-profit institutions.

This analysis was conducted at the school level, and so it cannot control for individual variation in school preparation, parental socio-economic status, high school quality, personality attributes, or other important factors that contribute to individual success. These factors have affected graduate earnings in other contexts, and they might create missing variables bias in college quality rankings. For example, parental education did affect alumni satisfaction with their undergraduate school (Rothwell 2019). To our main point, these factors would add yet another reason to suspect that the colleges are more similar in value-added than is implied by the most commonly used rankings of college quality.

Notes

1. Estimated using the U.S. Bureau of the Census Series P-32: Educational Attainment-Full-Time, Year-Round Workers 18 Years and Older.
2. Estimated using information compiled by the U.S. Bureau of Labor Statistics.
3. Information obtained from the U.S. Department of Education, National Center for Education Statistics.
4. Anderson (2013) reported that 91% of college admissions officials believe other universities doctor their statistics to inflate their rankings.
5. Our use of a weighted average of starting salaries by major, $W_{tk}^{M} = \sum_{i=1}^{I} \omega_{itk} W_{it}^{M}$, also reduces the potential measurement error bias from using possibly noisy reported salaries by major. Suppose that $W_{it}^{M} = W_{it}^{T} + e_{it}$, where W_{it}^{M} is the reported median starting salary for major *i* in year *t*, W_{it}^{T} is the true median value, and e_{it} is a random error with mean zero and variance σ_{it}. The weighted average of major salaries will have variance $\sum_{i=1}^{I} \omega_{itk}^{2} \sigma_{it} < \sum_{i=1}^{I} \sigma_{it}$ because each weight $\omega_{itk} \leq 1$. As a result, the use of a weighted average of many observed median values will have a lower error variance than had we used the individual measures of major salaries.
6. Details on the *PayScale* methodology is provided at https://www.payscale.com/docs/default-source/pdf/data_one_pager.pdf.
7. *PayScale* generates a ranking of colleges by salary, but they do not provide the reference by major that we undertake in this study or the corrections for local cost of living and student quality.
8. A complete list containing the ε_{tk}^{R}, ε_{tk}^{V}, ε_{tk}^{UV}, and RW_{tk} rankings for all 1,047 institutions and comparisons with the other rankings can be found online at: https://www2.econ.iastate.edu/faculty/orazem/rankings_explanation/.
9. One difference may be in the measure of graduate salaries. The three highest ranked Texas public schools using our value-added measure were Prairie View A&M, Texas A&M-Kingsville and Lamar, all of which were ranked toward the bottom by Cunha and Miller (2014), but whose starting salaries reported by *PayScale* were 6-15% above the Texas public university average.
10. Details on the variable definitions are in Appendix Tables 2A and 2B.
11. It may seem surprising that there is no systematic disadvantage to for-profit degrees. Deming et al. (2016) found that in experimental settings, graduates of for-profit business and health institutions were significantly less likely to receive callbacks than were identical graduates of nonprofit schools. However, that does not mean that their pay is lower once they have landed a job.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Partial Support from the Charles Koch Foundation is gratefully acknowledged.

ORCID

Peter F. Orazem http://orcid.org/0000-0003-2092-5089

References

Anderson, N. 2013. "Five Colleges Misreported Data to U.S. News, Raising Concerns About Rankings, Reputation." *Washington Post*.

Bastedo, M. N., and N. A. Bowman. 2010. "US News & World Report College Rankings: Modeling Institutional Effects on Organizational Reputation." *American Journal of Education* 116 (2): 163–183. doi:10.1086/649437

Black, D. A., and J. A. Smith. 2004. "How Robust is the Evidence on the Effects of College Quality? Evidence from Matching." *Journal of Econometrics* 121 (1-2): 99–124. doi:10.1016/j.jeconom.2003.10.006

Carnevale, Anthony P., Ban Cheah, and Andrew R. Hanson. 2015. *The Economic Value of College Majors*. Washington, D.C.: Georgetown University Center on Education and the Workforce.

Costa, D. L., and M. E. Kahn. 2000. "Power Couples: Changes in the Locational Choice of the College Educated." *The Quarterly Journal of Economics* 115 (4): 1287–1315. doi:10.1162/003355300555079

Cunha, J. M., and T. Miller. 2014. "Measuring Value-Added in Higher Education: Possibilities and Limitations in the Use of Administrative Data." *Economics of Education Review* 42: 64–77. doi:10.1016/j.econedurev.2014.06.001

Dale, S. B., and A. B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *The Quarterly Journal of Economics* 117 (4): 1491–1527. doi:10.1162/003355302320935089

Deming, D. J., N. Yuchtman, A. Abulafi, C. Goldin, and L. F. Katz. 2016. "The Value of Postsecondary Credentials in the Labor Market: An Experimental Study." *American Economic Review* 106 (3): 778–806. doi:10.1257/aer.20141757

Dillon, E. W., and J. A. Smith. 2020. "The Consequences of Academic Match Between Students and Colleges." *Journal of Human Resources* 55(3): 767–808.

Ehrenberg, R. G. 2005. "Method or Madness? Inside the U.S. News & World Report College Rankings." *Journal of College Admission* 189: 29–35. https://hdl.handle.net/1813/75853

Eide, E. R., M. J. Hilmer, and M. H. Showalter. 2016. "Is it Where You Go or What You Study? The Relative Influence of College Selectivity and College Major on Earnings." *Contemporary Economic Policy* 34 (1): 37–46. doi:10.1111/coep.12115

Forbes. 2017. *America's Top Colleges 2017*. https://www.forbes.com/top-colleges/#6d2e3dc61987

Grunig, S. D. 1997. "Research, Reputation and Resources." *The Journal of Higher Education* 68 (1): 17–52. https://doi.org/10.1080/00221546.1997.11778976.

Gyourko, J., and J. Tracy. 1991. "The Structure of Local Public Finance and the Quality of Life." *Journal of Political Economy* 99 (4): 774–806. doi:10.1086/261778

Hanson, Melanie. February 15, 2021. *Financial Aid Statistics* EducationData.org. https://educationdata.org/financial-aid-statistics.

Hossler, D. 2000. "The Problem with College Rankings." *About Campus* 5 (1): 20–24. doi:10.1177/108648220000500105

Hoxby, C. M. 2019. "The Productivity of US Postsecondary Institutions." In *Productivity in Higher Education*, edited by C. M. Hoxby, K. Strange, and K. Stange, 31–66. Chicago, IL: University of Chicago Press..

Job Search Intelligence (JSI). 2023. *JSI – The People and the Company*. https://www.jobsearchintelligence.com/about/.

Kelchen, R., and D. N. Harris. 2012. "Can 'Value-Added' Methods Improve the Measurement of College Performance? Empirical Analyses and Policy Implications." *Context for Success Working Paper*.

Kirkeboen, L. J., E. Leuven, and M. Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *The Quarterly Journal of Economics* 131 (3): 1057–1111. doi:10.1093/qje/qjw019

Koncz, A. 2016. "Salary Trends Through Salary Survey: A Historical Perspective on Starting Salaries for New College Graduates." https://www.naceweb.org/job-market/compensation/salary-trends-through-salary-survey-a-historical-perspective-on-starting-salaries-for-new-college-graduates/.

Light, A., and W. Strayer. 2000. "Determinants of College Completion: School Quality or Student Ability?" *Journal of Human Resources* 35 (2): 299–332. doi:10.2307/146327

Long, M. C. 2010. "Changes in the Returns to Education and College Quality." *Economics of Education Review* 29 (3): 338–347. doi:10.1016/j.econedurev.2009.10.005

Mogstad, M., J. P. Romano, A. Shaikh, and D. Wilhelm. 2022. "Statistical Uncertainty in the Ranking of Journals and Universities." *AEA Papers and Proceedings*.

Monks, J., and R. G. Ehrenberg. 1999. "The Impact of US News and World Report College Rankings on Admission Outcomes and Pricing Decisions at Selective Private Institutions". No. w7227. *National Bureau of Economic Research*.

National Association of Colleges and Employers (NACE). 2013. *Salary Survey: September 2013 Executive Summary*.

National Association of Colleges and Employers (NACE). 2022. *About the Association*. https://www.naceweb.org/aboutus.aspx?msclkid=914ee874d09311ec96e13fc0127229e6.

National Center for Education Statistics. (NCES). . 2022. "Undergraduate Retention and Graduation Rates." *Condition of Education*. U.S. Department of Education, Institute of Education Sciences.

Roback, J. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–1278. doi:10.1086/261120

Rosen, S. 1979. "Wage-Based Indexes of Urban Quality of Life." In *Current Issues in Urban Economics*, edited by Peter Mieszkowski, and Mahlon Straszheim, 74–104. Baltimore: Johns Hopkins University Press..

Rothwell, J. 2015. *Using Earnings Data to Rank Colleges: A Value-Added Approach Updated with College Scorecard Data*. Washington, DC: Brookings Institution.
Rothwell, J. 2019. "Assessing the Validity of Consumer Ratings for Higher Education: Evidence from a New Survey." *The Journal of Consumer Affairs* 53 (1): 167–200. doi:10.1111/joca.12201
Rothwell, J., and S. Kulkarni. 2015. "Beyond College Rankings: A Value-Added Approach to Assessing Two- and Four-Year Schools." *Metropolitan Policy Program at Brookings*.
Sauder, M., and R. Lancaster. 2006. "Do Rankings Matter? The Effects of US News & World Report Rankings on the Admissions Process of Law Schools." *Law and Society Review* 40 (1): 105–134. doi:10.1111/j.1540-5893.2006.00261.x
Shapiro, J. M. 2006. "Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital." *The Review of Economics and Statistics* 88 (2): 324–335. doi:10.1162/rest.88.2.324
Thomas, S. L., and L. Zhang. 2005. "Post-Baccalaureate Wage Growth Within Four Years of Graduation: The Effects of College Quality and College Major." *Research in Higher Education* 46 (4): 437–459. doi:10.1007/s11162-005-2969-y
Toutkoushian, R. K., and J. C. Smart. 2001. "Do Institutional Characteristics Affect Student Gains from College?" *The Review of Higher Education* 25 (1): 39–61. doi:10.1353/rhe.2001.0017
U.S. Department of Education. 2022. *College Scorecard*. https://collegescorecard.ed.gov/data/.
U.S. News & World Report. 2018. *The 10 Best Universities in America*. https://www.usnews.com/best-colleges/rankings/national-universities
Walker, I., and Y. Zhu. 2018. "University Selectivity and the Relative Returns to Higher Education: Evidence from the UK." *Labour Economics* 53 (August): 230–249. doi:10.1016/j.labeco.2018.05.005
Weinstein, R. 2022. "Firm Decisions and Variation Across Universities in Access to High-Wage Jobs: Evidence from Employer Recruiting." *Journal of Labor Economics* 40 (1): 1–46. doi:10.1086/714367
Winters, J. V. 2020. "In-state College Enrollment and Later Life Location Decisions." *Journal of Human Resources* 55 (4): 1400–1426. doi:10.3368/jhr.55.4.0916-8255R2