Estimates of Value Added

At Top Liberal Arts Colleges*

Roger T. Kaufman
Professor of Economics
Department of Economics
Smith College
Northampton, MA 01063
(413) 585-3612
rkaufman@smith.edu

Geoffrey Woglom
Richard S. Volpert '56
Professor of Economics
Amherst College
Amherst, MA 01002
(413) 542-2433
grwoglom@amherst.edu
Corresponding Author

August, 2008

* We appreciate the financial support of the Teagle Foundation that helped us in conducting this research

Abstract

In ~~second part of~~ this paper we develop and estimate a simple higher education production function for liberal arts colleges. Our measures of output are statistics based on two commonly used indicators of success – the frequency with which a college's graduates subsequently receive doctorates and the college's six-year graduation rate. There is considerable variation in both indicators among the colleges in our sample of the top 110 liberal arts colleges, as ranked by *U.S. News and World Report*. We are able to explain more than half of the variation using independent variables that measure student quality and the intensity of resource utilization. We then construct crude measures of success for each school by comparing their actual performance with the performance that one could expect given the quality of their students and resource intensity.

Introduction

Although much has been written about the goals of undergraduate education, there is a paucity of quantitative comparative data of college performance. Several papers have examined post-baccalaureate incomes, but many observers find this measure insufficient. By developing two other quantitative measures of success, this paper represents a modest attempt to fill that gap. We present and estimate a simple higher education production function. We then construct crude measures of success for each college by comparing its actual performance with the performance that one could expect given the quality of its students and resource intensity.

Education Production Function

Our education production function consists of the following 6 equations

(1) $Y = \alpha_0 + \alpha_1 VA + \alpha_2 SQ + \varepsilon_Y$

where Y is a measure of educational output,

SQ is student quality, and

VA is value added by the college

(2) $VA = \beta_0 + \beta_1 X_1 + \beta_2 LE + \varepsilon_{VA}$

where LE is the logarithm of institutional expenditures per student, and

X_1 is a vector of unobservable school inputs that affect value added

(3) $SQ = \gamma_0 + \gamma_1 Z + \gamma_2 X_2 + \varepsilon_{SQ}$

where Z is a vector of observable variables related to student input quality, and X_2 is a vector of unobservable variables that affect student quality.

The unobservable variables affecting value-added are related to a school's academic reputation REP:

(4) $X_1 = ø_0 + ø_1 REP + \varepsilon_{x1}$

In addition, students may "sort" on the basis of a school's reputation:

(5) $X_2 = \delta_0 + \delta_1 REP + \varepsilon_{x2}$

Academic reputation depends, in part, on observed student quality in the past and on past educational spending, as well as other variables, but we shall treat it as exogenous.

Substituting Equations (2) and (3) into Equation (1) yields

(1') $Y = \alpha'_0 + \alpha_1 (\beta_1 X_1 + \beta_2 LE) + \alpha_2 (\gamma_1 Z + \gamma_2 X_2) + \varepsilon'$

and substituting Equations (4) and (5) in (1') yields

(6) $Y = \alpha'_0 + \alpha_1 \beta_2 LE + \alpha_2 \gamma_1 Z + (\alpha_1 \beta_1 ø_1 + \alpha_2 \gamma_2 \delta_1) REP + \varepsilon''$

where $\varepsilon'' = \varepsilon_Y + \alpha_1 \varepsilon_{VA} + \alpha_2 \varepsilon_{SQ} + \alpha_1 \beta_1 \varepsilon_{x1} + \alpha_2 \gamma_2 \varepsilon_{x2}$

Equation (6) is the basic equation we use in this paper. Note that the coefficients reflect both the effects of the exogenous variables on value-added and student quality and the effects of value-added on our measures of output. Ideally we would like to estimate value added as $\alpha_1 \beta_2 LE + \alpha_1 \beta_1 ø_1 REP$. Since the regression coefficient on REP, however, will equal $\alpha_1 \beta_1 ø_1 + \alpha_2 \gamma_2 \delta_1$, we will only be able to identify the first effect on value added. Under the normal orthogonality assumptions, the estimated coefficients will be unbiased estimates of the true coefficients.

In the next section we describe our two measures of output and compare them among the schools in our sample. In the third section we present our estimated equations for our first measure of output, discuss the results, and then use the estimated residuals from our regressions to "rank" the institutions in our sample. We then discuss the potential relevance of these rankings. In the fourth section we repeat this exercise for our second measure of output.

Measures of Output

Aside from other studies that have used post baccalaureate incomes as a measure of output, there is a paucity of good objective data that measure other dimensions of a college education. The first measure of output we use is the frequency with which a college's graduates subsequently receive doctorates. Our second measure of output is each college's six-year graduation rate, which is the percentage of matriculants who graduate within six years. We readily admit to the limitations of both of these variables. Many educators object to the idea of measuring the success of a college education by examining the future vocation of its graduates. Others object to the inference that the goal of a college should be to prepare more students for graduate school. Surely, many graduates who pursue other activities were also well trained and motivated during their undergraduate experience. Unfortunately, however, there are no data on the frequency with which a substantial number of institutions' graduates pursue professional and other post-baccalaureate degrees.

Although graduation rates are also imperfect measures of success and not all educators would agree that colleges should strive for higher graduation rates, this variable is generally accepted as one measure of success. Nevertheless, neither of these two measures of output capture what many educators would consider to be important ingredients of a successful college experience.

Our sample includes all 110 colleges in *U.S. News and World Report's* top tier of national liberal arts colleges, published in its September 1, 2003 annual edition of *America's Best Colleges*. To calculate the frequency with which graduates of each liberal arts college subsequently earn a doctorate, we first computed the average annual number of doctorates awarded to graduates of each college in the five-year period between 1998 and 2002. We then divided this number by the average annual number of baccalaureate degrees awarded at that school during the eight-year period between 1990 and 1997. We computed the frequency for each sex and for both sexes taken together. Data on the number of doctorates are collected by the National Science Foundation and are available on the WEBCASPAR website (http://webcaspar.nsf.gov/index.jsp). Baccalaureate data were obtained using the Integrated Postsecondary Data System (IPEDS) data on enrollment and degree recipients.

The PhD frequencies by sex for each liberal arts college, listed in the order of their *U.S. News* rankings in 2003, are presented in Table 1.[1] These data depict substantial differences in PhD frequencies among the top 110 liberal arts colleges. Whereas only 1.55 percent of Hillsdale College's graduates eventually earn doctorates, 26.37 percent of

[1] The ties in the *U.S. News* rankings are not indicated in Table 1.

Harvey Mudd's graduates do so. Summary statistics for the entire sample are given at the end of the table. The median (Q2) among all 110 schools is 5.46% and the frequency exceeds 8.09% at only one quarter of the colleges.

Among the 93 coed colleges the male PhD frequency exceeds the female frequency in 62, but the differences are not large. Schools that were single-sex during the entire period 1990-1997 have blanks in the relevant spaces in Table 1.[2] With the notable exception of Bryn Mawr, graduates from the single-sex colleges appear to have PhD frequencies comparable to those of the same sex who graduated from coed schools. The mean PhD frequency among the 14 all-women's colleges is 7.02 percent, compared with a mean frequency of 6.67 percent among women at coeducational colleges, a difference that is not statistically significant.

The six-year graduation rates presented in the last column of Table 1 were obtained from the September 1, 2003 issue of *U.S. News and World Report's* annual edition of *America's Best Colleges.* They refer to the percentage of matriculants who earned a baccalaureate degree in six years or less with 2002 being the most recent year of graduation. The distribution of graduation rates is narrower (relative to the median) than that for PhD frequency. The graduation rates range from 60% at Mills to 96% at Williams and Amherst, and they are generally lower at single-sex colleges. Both sets of

[2] Entries of 0% indicate that the school granted baccalaureate degrees to members of that sex during 1990-1997 but no graduates of that sex and school received doctorates during the period 1998-2002. In a few of the single-sex schools, the PhD frequency for all students is slightly different from the single-sex frequency. This occurs because several single-sex schools occasionally grant degrees to members of the opposite sex. In addition, NSF has a category of doctorate recipients for which the sex is unknown.

data are highly correlated with the *U.S. News* rank, with the simple correlations ranging from 0.51 to 0.73 (in absolute value).

Other Data

For our measure of institutional expenditures per student we used the logarithm of total expenses per full-time student during the fiscal year 1996/97, as reported in the IPEDS financial survey. The IPEDS data do not allow us to separate educational and general costs from other costs in a consistent and/or reliable manner. While the IPEDS questionnaire is very clear about what should be classified as an educational expense, some schools classify room and board costs as educational costs while others classify them as auxiliary enterprises (as they are asked to do). Because of this inconsistency, we used the reported figure for total expenses. Operating and maintenance costs on plant and equipment are included in expenses, while construction expenditures on new buildings and other capital budget items are not. Expenses will be greater for schools that have lower student-faculty ratios, more supporting staff per student, higher wages and salaries per employee, or a greater capital stock (per student) to maintain.

Since only the most capable students are able to pursue doctorates the two measures of student quality in our regressions estimating PhD rates reflect the distribution of high-ability students. The first is HISAT, which represents the 75th percentile combined SAT score for the class that entered each college in the fall of 1996, as published in the September 1, 1997 issue of *U.S. News and World Report*.[3] Thus, HISAT is the combined SAT score that was exceeded by 25% of the entering class. Scores for those colleges that report ACT rather than SAT scores were converted to SAT scores using a conversion table provided by The College Board (http://www.collegeboard.com/sat/cbsenior/html/stat00f.html).[4] ~~Finally, all~~To make the coefficients easier to read, we divided the combined SAT scores by 100.

Our second measure of high-ability students is the percentage of the entering class in September, 1996 that received what is now called the National Merit $2,500 Scholarship. These students receive the most competitive scholarship in the National Merit program, a one-time $2,500 award from the National Merit Scholarship Corporation. We call this group the "non-institutional" National Merit Scholars. Other, less competitive (but often more generous) Merit Scholarships are awarded to many National Merit Finalists by individual academic and corporate institutions. Many of the most selective and prestigious colleges and universities, however, do not participate in this second program.[5] In the fall of 1996 324 non-institutional National Merit Scholars matriculated at the colleges in our sample.[6]

---

[3] Although SAT data for the class entering in 1996 are the appropriate data for the six-year graduation rate, measured in 2002, ideally we should have used SAT data from an earlier year for our doctorate regressions.
[4] Bennington and Juniata reported one SAT score (presumably the median) rather than the 25%-75% range of scores. Since the average difference between the 25th and the 75th percentile scores among those schools that reported a range was 198 points, we estimated the range for these two colleges by subtracting and adding 99 points to their medians. Three of the other top-ranked schools in 2003 (Harvey Mudd, Hillsdale, and Principia) were not included among the top liberal arts colleges in the 1997 report. (Harvey Mudd was not included because it was classified as an engineering school prior to 2003.) For these schools we searched through the *US News* annual reports after 1997 and used the earliest reported scores. Principia and Hillsdale, however, were eventually omitted from our sample for other reasons.
[5] We thank Rod Oto, Director of Student Financial Services at Carleton College, for helping us understand the National Merit Award system of scholarships.
[6] An additional 561 entering first-year students at these colleges received institutional National Merit awards sponsored by their college.

Our reputation variable is based on each school's reputation score as published in the September 1, 1997 issue of *U.S. News and World Report*. These reputation scores, ranging from 1 to 4 are calculated from surveys of college presidents and deans. They comprise a quarter of the total score that *U.S. News* used to determine each college's overall ranking. In one set of regressions we use the absolute reputation score, whereas in others we use a set of dummy variables that identified the decile into which each school fell. The set of regional dummy variables therefore allowed us to test for non-linear reputation effects. Finally, we also included a set of dummy variables for the 9 Census regions.

Predicted PhD Frequencies

In Table 2 we present the results of several regressions of PhD frequencies. In the first two columns of Table 2 we report the results of regressions in which students of both sexes are combined, whereas the last four columns refer to regressions in which students are disaggregated by sex. The odd-numbered regressions include the aforementioned reputation variable and the even-numbered regressions include the set of dummy variables indicating the decile of the reputation variable. Each regression also includes a set of regional dummy variables. The regressions reported in Table 2 were estimated using ordinary least squares to make the coefficients easier to interpret. Since the dependent variable is bounded by 0 and 100 percent, however, we also estimated the equations using the logarithm of the odds ratio log[tphd/(1-tphd)] as our dependent variable and obtained qualitatively similar results.

The adjusted R^2 for every regression in Table 2 exceeds 0.60, indicating relatively good fits, and the coefficients almost always have the expected signs. The regression in column (1) implies that each additional HISAT point (which is equivalent to 100 SAT points) increases the frequency with which graduates subsequently receive doctorates by 1.4 percentage points. The effect appears to be greater for males than for females, where it is not statistically significant.

Although the reputation variable is always positive and has a t-statistic of about 1.0, it is not statistically significant. This is surprising because Equation (6) indicates that its coefficient is the sum of two parameters, each of which is the product of three presumably positive parameters.[7] On the other hand, the coefficient on the percentage of non-institutional merit scholars exceeds 1 in all the regressions. Since some of these students do not receive doctorates, it is likely that this variable is picking up some of the other effects of both student quality and reputation. In none of the regressions, however, is the coefficient on non-institutional merit scholars statistically different from 1.0 at the 5 percent level.

Perhaps most importantly, the coefficient on the logarithm of expenditures is statistically significant in all but the male regressions. When evaluated at the sample

[7] Equation (6) does indicate, however, that the coefficient on reputation will equal zero if the effects of reputation on unobserved determinants of value added and student quality are also equal to zero.

mean level of expenditures per student of $27,962, the coefficients imply that a ten percent increase in expenditures per student will increase the percentage of graduates who eventually receive doctorates by about 0.3 percentage points, even after controlling for student quality. This result implies that the $40,000 spent on each student by the wealthier colleges increases the frequency of doctorates between one and two percentage points (or by about 20 percent) relative to the average college *ceteris paribus*. Recall that this coefficient represents only part of the effect of colleges on value added, i.e., the parameter $\alpha_1 \beta_2$ in Equation (6). It is puzzling, however, that expenditures do not appear to affect male PhD rates, for which the 25th percentile SAT score appears to be more important.

At the bottom of Table 2 we report the F statistics for four hypotheses. Although few of the coefficients on the regional dummy variables are individually significant, the F-statistic testing whether all the regional dummy variables are zero is significantly different from zero at the 10 percent level in both the total PhD and male PhD regressions, where it is significant at the 5 percent level. Because there is no obvious reason why a school's location should be related to success in producing future PhDs the regional dummies are probably capturing unmeasured differences in either student quality or school productivity that happen to be correlated with geographic location.. They could also reflect regional differences in the cost of living, which would impact the amount of real expenditures per student. The coefficients on the regional dummies (not shown) indicate that the schools in the East, Pacific, West North Central, and West South Central regions are more likely to produce PhD candidates *ceteris paribus* than schools in the New England, East South Central and Mountain regions.

In our second F test we test the joint significance of the set of ten reputation dummy variables, and in no case are these dummy variables jointly significant. In the next row of Table 2 we tested the hypothesis that the regressions were the same for the top 26 colleges (approximately one-quarter of our sample) and the remaining 80.[8] This may be viewed as a test for "hierarchies" among the colleges (Winston, 1999). The F statistics of 2.8 and 3.2 indicate that the educational production function appears to be different in these two sets of schools. Differences in expenditures per student and HISAT scores appear to be less important among the top-ranked schools, whereas differences in the enrollment of non-institutional merit scholars appears to be more important.

Finally, in the last row of Table 2 we report the F statistic for the hypothesis that the coefficients in the regression explaining PhD frequencies among all students in women's colleges are the same as those for women in the coeducational institutions. The value of the F statistic indicates that we could not reject the null hypothesis, which was not surprising. When the two sub-samples for a Chow test are very unequal, the power of the tests (the ability of the test to reject the null when it is not in fact true) is relatively low. In our sample of 106 institutions, only 14 are all-women colleges. When we simply added a dummy variable to the female regressions, attending an all-women's college appears to raise the probability of receiving a doctorate by about one percentage point *ceteris paribus*, but the t-statistic on this coefficient was only about 1.0.

---

[8] We originally wanted to examine the top 25 colleges, but there was a tie for the 25th ranking.

In Table 3 we present "scores" for measuring the success of each college in our sample. These "scores" represent the extent to which each college's actual PhD frequency, as reported in Table 1, exceeded its predicted frequency. This methodology is similar to the one used by *U.S. News* in its value-added scores for six-year graduation rates, which we describe below. As noted earlier, it is unclear what the regional dummy variables are capturing.~~of~~ Consequently, our ~~first set of~~ scores are ~~in column (2) is~~ based on a regression of the variables in Column (1) of Table 2 but without the regional dummy variables. In the last column of Table 3, however, we also provide rankings based on the regression that includes the regional dummies.

Colleges are listed in descending order of their unexplained success in producing doctoral candidates. Thus, for example, Reed's PhD frequency of 20.19% (from Table 1) is 10.3 percentage points greater than the value predicted using the regression specification in column (1) of Table 2, but without the regional dummy variables. On the other end of the table, Claremont's frequency of 3.16% (again from Table 1) is 7.6 percentage points below its predicted value.[9] In other words, the "score" for each college is simply the value of its estimated residual in the corresponding PhD frequency regression. In general, those colleges that performed well according to the regression without regional dummies also performed well using the regression with regional dummies, and the correlation between the two rankings was 0.85. Seven colleges were in the top eight in both rankings, and twelve were in the bottom fifteen in both rankings.

[9] Since comparable data on expenses for Virginia Military Institute, St. Mary's, and Principia were not included in the 1996/1997 IPEDS data, they were not included in the regressions. Hillsdale was also excluded because of problems with data availability.

We recognize the difficulty in interpreting these "scores." As we noted earlier, the error term in Equation (6) of our model is a weighted sum of the error terms in the equations determining output, value-added, unobserved variables affecting value-added, student quality, and unobserved variables affecting student quality. Consequently, a college may have a large estimated residual, resulting in a high score, if it is more productive, reflected by large values of the first three error terms. Some colleges, for example, prepare students better for graduate school. Alternatively, the atmosphere at the college may stimulate interest among its undergraduates to pursue graduate education.

On the other hand, a high score would result if we have underestimated the quality of students at a particular college. Students who are attracted to a particular college, for example, might be more likely to pursue graduate education, even after controlling for SAT scores and resource intensity. The three colleges at the top of each column in Table 3, for example, (Reed, Bryn Mawr, and Oberlin) are well known for encouraging their students to pursue graduate training. High school students who are especially interested in graduate school may be characterized in ways that we have not measured, or these colleges may transform students after they arrive on campus. Our data are not sufficient for us to determine the answer conclusively.

Insofar as our measures of student quality are more exhaustive than those measuring the colleges' value-added, we suspect that higher residuals are more likely to reflect greater institutional productivity, but this is speculative. One should view the scores in Table 3 cautiously and be careful about making normative judgments based on them.

Predicted Graduation Rates

We also used six-year graduation rates to compute a second set of success scores similar to that published by *U.S. News*. We begin by estimating regressions for the graduation rates that appear in the last column of Table 1. The results of these regressions are presented in Table 4. Insofar as *U.S. News* does not publish graduation rates for each sex, we could only estimate a combined regression.

As indicated in Table 4 the regressors we used differ from the ones we used in our PhD frequency regressions, with the exception of the logarithm of expenses per student and college reputation. The four additional independent variables are as follows:

(1) a dummy variable for all-women's colleges

(2) the percentage of freshmen entering in the fall of 1996 who graduated in the top 10 percent of their high school class, as reported in *U.S. News.*[10]

(3) the logarithm of the total number of entering first-year students as reported in the IPEDS enrollment data base.

(4) HISAT-LOWSAT, which is the difference in the number of SAT points between the combined 75th percentile SAT score (HISAT) and the 25th percentile score (LOWSAT), divided by 100.

[10] The 1997 issue of *U.S. News* did not include the value of this variable for Harvey Mudd, Reed, Spelman, Hillsdale, and Juniata. As before, we used the earliest reported graduation rate in subsequent issues.

Regression (2) includes a complete set of regional dummy variables whereas regression (1) does not. All of the coefficients in the two regressions enter with the expected sign and most of the coefficients are statistically significant at the 5% level. These results imply that graduation rates increase with both the size of the liberal arts colleges in our sample and total expenses per student. *Ceteris paribus* women's colleges appear to have graduation rates between 2 and 3 percentage points below coed and all-male colleges, but this result may be misleading if women also have lower graduation rates than men at coeducational institutions.[11] The regression in column (1) implies that the six-year graduation rate increases by 0.21 percentage points if the percentage of freshmen that graduated in the top 10 percent of their high school class increases by one percentage point. Colleges with a broader distribution of SAT scores, as reflected in large differences between HISAT and LOWSAT, experience higher attrition rates, presumably because students at the lower end of a college's ability spectrum experience a larger gap between themselves and many of their peers. As in our earlier regressions, the college reputation variable is not statistically significant, perhaps because its effect is already captured by some of the other variables.

The summary statistics at the bottom of Table 4 indicate that the set of regional dummy variables is significant at the 5 percent level. We believe that the regional dummy variables in these regressions are capturing inter-regional effects that may be unrelated to school quality. While these liberal arts colleges are increasingly competing on a national level for students, there is still a strong regional bias in enrollments. Consequently,

[11] The average graduation rate among the fourteen all-women's colleges was 73.4 percent, compared with an overall (male and female) graduation rate of 77.9 percent at the coeducational institutions, a difference that is not statistically significant.

where a school is located tells you something about the residence of many of the school's students, and students from different regions may have different expectations about their college experience. For example, it has been our experience that the transfer rate is very high at West Coast schools, and the regional dummies (not shown) find that these West Coast schools have unusually low graduation rates. We believe these substantially lower graduation rates reflect more about the culture of the West Coast college experience than they do about the productivity of these schools. For example, among our 14 West Coast schools, only 3 had graduation rates in excess of 80%. In con~~s~~trast, 13 of our 15 New England schools ~~, 13~~ had graduation rates above 80%.

We also tested whether the coefficients among the top 26 colleges were different than those among the next 80. Unlike in the PhD regressions, we could not reject the null hypothesis that the coefficients were equal in the two sub-samples. Finally,

although we included a dummy variable for women's colleges in our basic regression, we also tested the more general hypothesis of whether the entire regression explaining graduation rates is different for women's colleges. For this test we re-estimated the regression in column (1) without the dummy variable for women's colleges and for coeducational colleges. The broader hypothesis cannot be rejected at the 10% level, but we suspect this reflects more about the low power of the Chow test with unequal sub-samples than it reflects similarity between women's colleges and the other schools.

Using the same methodology we employed to compute our graduate school "scores," we constructed a similar measures for graduation rates. We used the coefficients from the regressions in Table 4 to compute predicted graduation rates based on the values of the independent variables for each college. We then subtracted the predicted rate for each college from its actual rate to obtain each school's graduation score, which is merely the estimated residual for each college.

In Table 5 we report scores for all schools for both specifications in descending order. ~~*USN*~~Bennington's 83 percent graduation rate, for example (from Table 1) is 16.5 percentage points greater than the rate predicted using the regression in column (1) of Table 4 and 9.9 percentage points greater than the rate predicted using the regression in column (2), which contains regional dummy variables. Looking at the school with the lowest score, the graduation rate at Reed is 12.7 percentage points (or 8.7 percentage points with regional dummies) below what would be predicted for its talented student body.

In Table 5 we also report the "over" or "underperformance" scores for six-year graduation rates published by ~~*USNews*~~*U.S. News*. *U.S. News* estimates a regression to predict graduation rates at these schools and then computes "over" or "underperformance" scores similar to ours. Although they do not report their complete specification, the independent variables they do mention are similar to ours, but not identical. Their expense variable, for example, is calculated differently and they use SAT levels, rather than the gap between the 25th and 75th percentiles. They also don't include regional or all-female dummy variables.

The correlation coefficient between our scores and those reported in ~~*USNews*~~*U.S. News* (2003) are 0.62 for the scores we obtained from the regression in Column (1) of Table 4 (i.e., without regional dummies) and 0.50 for the scores we obtained using regional dummies in Column (2) of Table 4. These correlation coefficients show that our scores are related to those of ~~*USNews*~~*U.S. News*, but they are far from identical. It is not surprising that the correlation is higher for the specification without regional dummies, since we do not believe ~~*USNews*~~*U.S. News* uses such dummies and, as the results in Table 4 indicate, these dummies are important.

Once again, these scores are subject to several interpretations. Reed's disappointing graduation score, for example, may be the result of something that happens at Reed. Alternatively, Reed may simply attract students who enter college less committed to completing their baccalaureate degrees. It is interesting to note, for example, that three of the eight colleges that had the highest PhD scores (relative to their predicted values), Reed, Oberlin, and Harvey Mudd, were among the bottom eight scorers (relative to their predicted values) according to graduation rates.

Conclusion

We have documented that increased expenditures generally have a statistically significant effect on increasing the frequency with which colleges' students eventually received doctoral degrees and on six-year graduation rates, even after controlling for student quality. We have also illustrated the substantial variations in six-year graduation rates among the top liberal arts colleges and the frequency with which their graduates subsequently receive doctorates, even after we control for a variety of factors, including differences in the quality of students and resource intensity. While we do not claim that these two variables measure all dimensions of a college's success, we do believe that they provide useful information for students, faculty, administrators, and other researchers.

Table 1

PhD Frequencies and Six-year Graduation Rates

Among Top Liberal Arts Colleges

Sept, 2003 US News Rank	College	1998-2002 Total PhD Rate	1998-2002 Female PhD Rate	1998-2002 Male PhD Rate	2002 Six-Year Grad. Rate
1	Williams	13.36%	12.58%	14.05%	96
2	Amherst	14.05%	15.07%	13.22%	96
3	Swarthmore	19.98%	18.44%	21.45%	92
4	Carleton	17.48%	16.61%	18.35%	86
5	Pomona	14.72%	18.26%	11.32%	89
6	Wellesley	10.84%	10.84%		91
7	Davidson	7.12%	7.16%	7.09%	89
8	Middlebury	6.05%	5.48%	6.62%	88
9	Haverford	14.53%	11.05%	17.48%	91
10	Bowdoin	8.91%	7.36%	10.31%	90
11	Wesleyan	12.50%	13.28%	11.66%	89
12	Claremont	3.16%	3.13%	3.17%	84
13	Washington & Lee	4.65%	4.56%	4.70%	88
14	Vassar	9.59%	10.71%	7.93%	88
15	Grinnell	13.88%	11.68%	16.26%	84
16	Smith	9.07%	8.99%		80
17	Bryn Mawr	16.39%	16.39%		82
18	Colby	5.95%	5.22%	6.74%	88
19	Colgate	6.40%	6.51%	6.29%	88
20	Harvey Mudd	26.37%	24.51%	26.89%	79
21	Hamilton	6.99%	6.91%	7.05%	84
22	Trinity	5.63%	6.40%	4.84%	83
23	Bates	8.39%	8.40%	8.37%	87
24	Oberlin	16.72%	15.07%	18.76%	76
25	Macalester	9.74%	8.55%	11.20%	82
26	Mt. Holyoke	8.84%	8.80%		81
27	Bucknell	5.31%	4.90%	5.68%	88
28	C. Holy Cross	5.10%	4.38%	5.84%	91
29	Colorado College	5.92%	5.28%	6.58%	83
30	Bard	7.03%	6.79%	7.35%	70
31	Kenyon	6.44%	4.61%	8.49%	83
32	Lafayette	4.32%	5.08%	3.75%	85
33	U. of the South	5.84%	5.48%	6.18%	76
34	Connecticut	5.53%	6.85%	3.72%	83
35	Scripps	5.18%	5.19%		68
36	Whitman	7.56%	6.56%	8.73%	87
37	Union	3.84%	4.38%	3.41%	84
38	Barnard	8.67%	8.63%		86
39	Franklin & Marshall	7.87%	8.08%	7.68%	84
40	DePauw	4.52%	3.99%	5.16%	77

41	Occidental	8.10%	7.70%	8.58%	79
42	Dickinson	4.56%	3.52%	5.93%	79
43	Furman	6.63%	6.12%	7.25%	81
44	Skidmore	3.71%	4.47%	2.60%	75
45	Centre	4.73%	4.56%	4.90%	76
46	Gettysburg	3.96%	3.85%	4.09%	76
47	Rhodes	5.22%	4.08%	6.64%	73
48	Sarah Lawrence	5.24%	4.96%	6.10%	72
49	Denison	4.41%	4.51%	4.29%	76
50	Agnes Scott	7.13%	7.13%		72
51	Wabash	8.07%		8.07%	68
52	Beloit	10.79%	9.79%	12.12%	66
53	Lawrence	9.42%	9.30%	9.54%	67
54	Southwestern	5.99%	6.08%	5.88%	69
55	Wheaton IL	6.79%	4.27%	9.83%	86
56	ILL Wesleyan	4.49%	4.22%	4.84%	81
57	Reed	20.19%	14.99%	24.72%	67
58	Willamette	4.00%	3.52%	4.59%	80
59	College of Wooster	7.83%	8.57%	6.89%	65
60	Drew	6.41%	6.61%	6.12%	75
61	St. Olaf	8.72%	8.14%	9.40%	80
62	St. Lawrence	3.60%	2.82%	4.38%	72
63	Earlham	9.73%	8.13%	11.88%	73
64	Sweet Briar	3.79%	3.79%		61
65	Wheaton MA	3.00%	3.65%	0.00%	70
66	Birmingham Southern	4.25%	4.23%	4.28%	74
67	Kalamazoo	11.49%	12.28%	10.54%	74
68	Mills	3.81%	3.83%		60
69	Wofford	4.15%	4.04%	4.22%	75
70	Bennington	7.81%	8.96%	6.12%	83
71	Gustavus Adolphus	4.39%	3.28%	5.85%	81
72	Hendrix	8.14%	8.56%	7.65%	68
73	Hobart & William Smith	4.02%	3.90%	4.12%	72
74	Muhlenberg	3.32%	3.37%	3.27%	81
75	Pitzer	4.02%	4.33%	3.56%	66
76	Ursinus	4.50%	4.38%	4.65%	78
77	VMI	2.66%	0.00%	2.66%	66
78	Austin	5.59%	5.63%	5.54%	78
79	Knox	8.43%	8.87%	7.95%	72
80	St. Mary's	2.75%	2.56%	3.00%	81
81	Randolph-Macon Women's	3.89%	3.89%		62
82	Spelman	7.82%	7.72%		76

83	St. John's MN	4.79%		4.58%	80
84	U of Puget Sound	3.47%	2.73%	4.49%	74
85	Albion	3.96%	3.93%	3.98%	68
86	Hollins	2.87%	2.87%		68
87	Lewis and Clark	4.73%	3.98%	5.67%	63
88	Allegheny	7.07%	6.79%	7.38%	73
89	Augustana	4.71%	4.46%	4.94%	75
90	Hanover	4.40%	4.99%	3.63%	71
91	Goucher	5.33%	6.20%	0.00%	68
92	Hope	4.86%	3.04%	7.37%	74
93	Millsaps	4.43%	4.20%	4.67%	68
94	Principia	2.80%	2.11%	3.82%	71
95	Thomas Aquinas	5.98%	5.93%	0.00%	75
96	Washington & Jefferson	3.04%	3.10%	3.00%	73
97	C. St. Benedict	2.50%	2.50%		77
98	Hillsdale	1.55%	1.30%	2.11%	69
99	Lake Forest	3.43%	3.21%	3.71%	70
100	Ohio Wesleyan	5.29%	5.26%	5.32%	66
101	Randolph-Macon	2.14%	2.03%	2.26%	73
102	Transylvania	5.58%	4.68%	6.69%	66
103	Washington	2.31%	2.11%	2.55%	70
104	Westmont	3.80%	2.64%	5.49%	68
105	Wittenberg	4.66%	4.09%	5.47%	70
106	Hampden-Sydney	3.59%		3.59%	61
107	Juniata	4.68%	4.77%	4.58%	74
108	Luther	5.39%	4.59%	6.53%	79
109	Presbyterian	2.04%	1.98%	2.09%	72
110	Wells	7.73%	7.73%		63
Correlation with Rank		**-0.55**	**-0.57**	**-0.51**	**-0.73**
	Q1	**8.09%**	**8.13%**	**7.98%**	**83**
	Q2	**5.46%**	**5.19%**	**5.86%**	**76**
	Q3	**4.05%**	**3.95%**	**4.19%**	**70**

Table 2: Predicting PhD Rates[a]

	(1) Total	(2) Total	(3) Fem	(4) Fem	(5) Male	(6) Male
Log Expense per Student	0.032	.0357	.0408	.0433	-.0028	-.000
	(1.98)	(2.02)	(2.56)	(2.48)	(0.13)	(.01)
HISAT	.0139	.0137	.0083	.0103	.0265	.0237
	(1.91)	(1.71)	(1.14)	(1.31)	(2.84)	(2.34)
% Merit Scholars	1.28	1.15	1.29	1.10	1.14	1.08
	(4.59)	(3.73)	(4.67)	(3.65)	(3.41)	(2.89)
REP	.0089		.0089		.0102	
	(0.97)		(0.97)		(0.81)	
Adj. R^2	0.650	0.642	0.629	0.631	0.635	0.615
Std Error of Reg	0.025	0.025	0.025	0.025	0.029	0.030
F tests:						
Regional Dummies	1.87*	1.93*	1.53	1.52	2.4**	2.6**
Decile Dummies		0.80		1.03		0.55
Top 26 vs Bottom 80	2.8**	3.2**				
Women's Colleges			1.42			

a. Each Regression includes a constant term and a set of regional dummy variables. The absolute value of the "*t*' statistic is in parentheses. * and ** indicate *F* values significant at the 10 and 5 percent level respectively.

Table 3

Actual PHD Rates Relative to Regresson Predictions

Rank Without Regional Dummies	College	Score in % points	Rank With Regional Dummies
1	Reed	10.3%	1
2	Bryn Mawr	6.8%	2
3	Oberlin	6.8%	3
4	Carleton	5.2%	8
5	Beloit	5.0%	5
6	Spelman	4.5%	4
7	Harvey Mudd	4.5%	6
8	Kalamazoo	4.1%	10
9	Earlham	3.4%	12
10	Hendrix	3.3%	17
11	Wesleyan	3.0%	7
12	Bennington	2.4%	11
13	Knox	2.4%	22
14	St. Olaf	2.3%	26
15	Occidental	2.1%	19
16	College of Wooster	2.1%	29
17	Haverford	2.0%	18
18	Grinnell	1.8%	58
19	Allegheny	1.8%	13
20	Wells	1.7%	23
21	Luther	1.5%	30
22	Thomas Aquinas	1.5%	28
23	Franklin & Marshall	1.4%	24
24	Wabash	1.4%	39
25	St. John's MN	1.2%	53
26	Juniata	1.2%	20
27	Lawrence	1.1%	45
28	Mt. Holyoke	1.0%	16
29	Swarthmore	1.0%	36
30	Augustana	0.9%	35
31	Bates	0.9%	14
32	Austin	0.8%	67
33	Transylvania	0.8%	9
34	Wofford	0.7%	15
35	Southwestern	0.6%	76
36	Barnard	0.5%	34
37	Vassar	0.5%	57
38	Hope	0.4%	59
39	Westmont	0.4%	32
40	Wittenberg	0.4%	51
41	Goucher	0.4%	27

42	Lake Forest	0.3%	69
43	Smith	0.2%	25
44	Ursinus	0.2%	38
45	Pomona	0.2%	52
46	Agnes Scott	0.2%	33
47	Macalester	-0.1%	84
48	Whitman	-0.1%	41
49	Hobart & William Smith	-0.1%	50
50	Drew	-0.2%	56
51	Bard	-0.3%	70
52	Muhlenberg	-0.3%	49
53	Hanover	-0.3%	74
54	Washington & Jefferson	-0.3%	44
55	Hamilton	-0.4%	65
56	Furman	-0.5%	31
57	Albion	-0.5%	80
58	Sweet Briar	-0.6%	55
59	ILL Wesleyan	-0.6%	88
60	Dickinson	-0.7%	82
61	Hampden-Sydney	-0.7%	43
62	Wellesley	-0.8%	61
63	Hollins	-0.8%	48
64	Millsaps	-0.8%	21
65	Gustavus Adolphus	-0.9%	87
66	St. Lawrence	-0.9%	75
67	Mills	-0.9%	79
68	Denison	-0.9%	90
69	Washington	-0.9%	42
70	C. St. Benedict	-1.0%	93
71	Randolph-Macon	-1.0%	47
72	Ohio Wesleyan	-1.0%	89
73	U. of the South	-1.0%	37
74	C. Holy Cross	-1.1%	40
75	Randolph-Macon Women's	-1.1%	62
76	Lewis and Clark	-1.1%	68
77	Bucknell	-1.1%	78
78	Bowdoin	-1.3%	60
79	Colgate	-1.3%	81
80	DePauw	-1.3%	97
81	Willamette	-1.4%	73
82	Gettysburg	-1.4%	85
83	Scripps	-1.4%	86
84	Presbyterian	-1.4%	54
85	Colorado College	-1.5%	46
86	Sarah Lawrence	-1.5%	83
87	Skidmore	-1.6%	91
88	Connecticut	-1.6%	64
89	Colby	-1.7%	66
90	Amherst	-1.7%	71
91	Lafayette	-1.7%	94

92	Wheaton MA	-1.8%	63
93	Union	-1.8%	92
94	Pitzer	-2.2%	98
95	U of Puget Sound	-2.3%	96
96	Wheaton IL	-2.4%	101
97	Trinity	-2.8%	95
98	Centre	-2.9%	77
99	Kenyon	-2.9%	104
100	Birmingham Southern	-3.1%	72
101	Middlebury	-3.5%	100
102	Davidson	-4.4%	103
103	Williams	-4.5%	102
104	Rhodes	-5.2%	99
105	Washington & Lee	-6.1%	105
106	Claremont	-7.6%	106

Table 4: Predicting Graduation Rates	(1) No Regional Dummies	(2) With Regional Dummies
Women's College	-.019 (1.28)	-.0325 (2.44)
Freshman in top 10% HS Class	0.214 (4.63)	0.269 (6.59)
Log Incoming Freshmen	0.119 (3.87)	.0539 (1.84)
HISAT - LOWSAT	-.049 (3.04)	-.0502 (3.66)
Log Expense Per Student	0.054 (1.53)	0.0127 (0.39)
Rep96	.0229 (1.26)	0.013 (0.81)
Adj. R^2	0.648	0.755
Std Error of Reg	.050	.0419
F Tests:		
Regional Dummies		6.33**
Top 26 vs Bottom 80	0.93	0.74
Women's Colleges	1.44	

See notes to Table 2

Table 5

Actual Graduation Rates Relative to Regression Predictions and U.S. News Scores

Rank without Regional Dummies	College	Score in % Points	*U.S. News % Point* Score	Rank with Regional Dummies	College	Score in % points	*U.S. News* % Point Score
1	Bennington	16.5%	14.0%	1	Whitman	10.6%	5.0%
2	C. St. Benedict	9.5%	5.0%	2	Bennington	9.9%	14.0%
3	Whitman	8.5%	5.0%	3	Earlham	8.4%	5.0%
4	Thomas Aquinas	8.4%	15.0%	4	Lafayette	7.2%	1.0%
5	Gustavus Adolphus	8.4%	13.0%	5	Austin	6.4%	13.0%
6	St. John's MN	8.1%	2.0%	6	Kenyon	6.4%	8.0%
7	Lafayette	7.1%	1.0%	7	Wheaton IL	6.2%	12.0%
8	C. Holy Cross	6.1%	-5.0%	8	Thomas Aquinas	6.2%	15.0%
9	Earlham	6.1%	5.0%	9	Birmingham Southern	5.9%	-4.0%
10	Muhlenberg	5.7%	0.0%	10	Spelman	5.6%	-2.0%
11	Austin	5.3%	13.0%	11	C. St. Benedict	4.8%	5.0%
12	Amherst	5.2%	12.0%	12	Randolph-Macon	4.7%	3.0%
13	Wheaton IL	4.9%	12.0%	13	Willamette	4.3%	-1.0%
14	Williams	4.9%	3.0%	14	Barnard	3.8%	-4.0%
15	Wofford	4.8%	15.0%	15	Gustavus Adolphus	3.8%	13.0%
16	Randolph-Macon	4.7%	3.0%	16	Occidental	3.7%	5.0%
17	Colorado College	4.4%	6.0%	17	Denison	3.7%	-3.0%
18	Ursinus	4.4%	6.0%	18	St. John's MN	3.6%	2.0%
19	Haverford	4.3%	4.0%	19	Muhlenberg	3.4%	0.0%
20	Kenyon	4.2%	8.0%	20	Augustana	3.4%	5.0%
21	Juniata	4.0%	3.0%	21	Hamilton	3.2%	13.0%
22	Washington & Jefferson	3.7%	7.0%	22	Colgate	3.1%	0.0%
23	Bates	3.6%	6.0%	23	Pomona	3.1%	2.0%
24	Hamilton	3.5%	13.0%	24	Claremont	3.1%	8.0%
25	Barnard	3.5%	-4.0%	25	Wofford	2.9%	15.0%
26	Colgate	3.3%	0.0%	26	Dickinson	2.9%	-2.0%
27	Union	3.2%	7.0%	27	U. of the South	2.9%	0.0%
28	Spelman	3.2%	-2.0%	28	Swarthmore	2.5%	6.0%
29	Luther	3.1%	1.0%	29	Hollins	2.3%	10.0%
30	Swarthmore	3.0%	6.0%	30	Lake Forest	2.2%	3.0%
31	Colby	2.9%	0.0%	31	C. Holy Cross	2.1%	-5.0%
32	Connecticut	2.8%	2.0%	32	Union	2.1%	7.0%
33	Mt. Holyoke	2.6%	8.0%	33	Williams	2.0%	3.0%
34	Willamette	2.6%	-1.0%	34	Knox	1.7%	5.0%
35	Hollins	2.1%	10.0%	35	Haverford	1.6%	4.0%
36	Dickinson	2.1%	-2.0%	36	Davidson	1.6%	10.0%
37	St. Olaf	2.1%	0.0%	37	Hope	1.5%	-2.0%
38	Wells	2.0%	7.0%	38	Amherst	1.4%	12.0%
39	Augustana	2.0%	5.0%	39	Washington & Lee	1.3%	-3.0%
40	Franklin & Marshall	1.5%	2.0%	40	Mt. Holyoke	1.3%	8.0%
41	Wesleyan	1.4%	2.0%	41	DePauw	1.2%	-1.0%
42	Presbyterian	1.3%	4.0%	42	Bucknell	1.1%	-3.0%
43	Washington & Lee	1.3%	-3.0%	43	Wellesley	1.1%	6.0%
44	Wellesley	1.2%	6.0%	44	Centre	1.1%	-6.0%
45	Carleton	1.2%	3.0%	45	Furman	1.1%	9.0%
46	Trinity	1.2%	4.0%	46	Ursinus	0.3%	6.0%
47	Davidson	1.1%	10.0%	47	Agnes Scott	0.3%	6.0%
48	Drew	1.1%	0.0%	48	Skidmore	0.3%	7.0%
49	Bucknell	1.1%	-3.0%	49	Wabash	0.2%	-4.0%

50	Bowdoin	0.9%	4.0%	50	Colorado College	0.0%	6.0%
51	Birmingham Southern	0.7%	-4.0%	51	Bates	0.0%	6.0%
52	Grinnell	0.6%	4.0%	52	Vassar	-0.1%	5.0%
53	Knox	0.6%	5.0%	53	Franklin & Marshall	-0.2%	2.0%
54	Lake Forest	0.5%	3.0%	54	Bryn Mawr	-0.2%	6.0%
55	Claremont	0.3%	8.0%	55	Wesleyan	-0.3%	2.0%
56	Furman	0.3%	9.0%	56	Juniata	-0.3%	3.0%
57	Wheaton MA	0.2%	-1.0%	57	U of Puget Sound	-0.5%	-2.0%
58	Denison	0.2%	-3.0%	58	Connecticut	-0.6%	2.0%
59	Sarah Lawrence	0.0%	6.0%	59	Presbyterian	-0.6%	4.0%
60	Occidental	-0.2%	5.0%	60	Washington & Jefferson	-0.7%	7.0%
61	Bryn Mawr	-0.3%	6.0%	61	ILL Wesleyan	-0.8%	4.0%
62	Agnes Scott	-0.5%	6.0%	62	Oberlin	-0.9%	-8.0%
63	Macalester	-0.5%	6.0%	63	Southwestern	-0.9%	-2.0%
64	Skidmore	-0.9%	7.0%	64	St. Olaf	-0.9%	0.0%
65	Vassar	-1.0%	5.0%	65	Colby	-0.9%	0.0%
66	Wabash	-1.0%	-4.0%	66	Trinity	-1.0%	4.0%
67	Washington	-1.0%	-9.0%	67	Rhodes	-1.3%	1.0%
68	Pomona	-1.2%	2.0%	68	Westmont	-1.4%	8.0%
69	Hope	-1.4%	-2.0%	69	Mills	-1.4%	-4.0%
70	Hobart & William Smith	-1.6%	-10.0%	70	Drew	-1.7%	0.0%
71	Goucher	-1.9%	-3.0%	71	Goucher	-1.8%	-3.0%
72	Hanover	-1.9%	-6.0%	72	Kalamazoo	-1.9%	6.0%
73	Gettysburg	-2.3%	-1.0%	73	Washington	-2.0%	-9.0%
74	DePauw	-2.4%	-1.0%	74	Scripps	-2.1%	-4.0%
75	Middlebury	-2.5%	-2.0%	75	Beloit	-2.2%	-4.0%
76	St. Lawrence	-2.8%	-1.0%	76	Carleton	-2.2%	3.0%
77	Southwestern	-2.9%	-2.0%	77	Grinnell	-2.3%	4.0%
78	Westmont	-3.0%	8.0%	78	Wells	-2.5%	7.0%
79	ILL Wesleyan	-3.2%	4.0%	79	Ohio Wesleyan	-2.5%	-8.0%
80	Smith	-3.3%	-4.0%	80	Sarah Lawrence	-2.6%	6.0%
81	Beloit	-3.4%	-4.0%	81	Luther	-2.6%	1.0%
82	Sweet Briar	-3.5%	4.0%	82	Hanover	-2.7%	-6.0%
83	Kalamazoo	-4.0%	6.0%	83	Hobart & William Smith	-2.8%	-10.0%
84	Centre	-4.3%	-6.0%	84	Gettysburg	-2.9%	-1.0%
85	Mills	-4.4%	-4.0%	85	Pitzer	-2.9%	-2.0%
86	U. of the South	-4.5%	0.0%	86	Sweet Briar	-3.0%	4.0%
87	Allegheny	-4.6%	0.0%	87	St. Lawrence	-3.0%	-1.0%
88	U of Puget Sound	-4.7%	-2.0%	88	Wheaton MA	-3.1%	-1.0%
89	Bard	-5.4%	-9.0%	89	Bowdoin	-3.2%	4.0%
90	Scripps	-5.4%	-4.0%	90	Millsaps	-3.6%	-18.0%
91	Pitzer	-6.0%	-2.0%	91	Smith	-3.9%	-4.0%
92	Ohio Wesleyan	-6.1%	-8.0%	92	Macalester	-4.3%	6.0%
93	Hendrix	-6.1%	6.0%	93	Middlebury	-4.8%	-2.0%
94	Wittenberg	-6.2%	0.0%	94	Transylvania	-4.9%	-2.0%
95	Albion	-6.3%	-2.0%	95	Albion	-5.3%	-2.0%
96	Randolph-Macon Women's	-6.5%	-5.0%	96	College of Wooster	-5.5%	-1.0%
97	Rhodes	-6.6%	1.0%	97	Hendrix	-5.5%	6.0%
98	Hampden-Sydney	-6.8%	-5.0%	98	Wittenberg	-5.5%	0.0%
99	Oberlin	-7.0%	-8.0%	99	Randolph-Macon Women's	-6.1%	-5.0%
100	Millsaps	-7.6%	-18.0%	100	Hampden-Sydney	-6.4%	-5.0%
101	Transylvania	-8.4%	-2.0%	101	Harvey Mudd	-6.6%	-12.0%
102	Harvey Mudd	-8.6%	-12.0%	102	Bard	-7.2%	-9.0%
103	College of Wooster	-9.0%	-1.0%	103	Allegheny	-7.4%	0.0%
104	Lawrence	-9.1%	-9.0%	104	Lewis and Clark	-7.4%	-8.0%
105	Lewis and Clark	-10.5%	-8.0%	105	Lawrence	-7.6%	-9.0%
106	Reed	-12.7%	#NA	106	Reed	-8.7%	#NA

References

Kaufman, Roger T. & Woglom G. (2005). "Financial Changes and Optimal Spending Rates Among Top Liberal Arts Colleges 1996-2001. *Review of Higher Education,* 28(3), 339-368.

National Center for Education Statistics (various dates). *Finance survey of the NCES integrated postsecondary education data system.* Retrieved, March 9,2004, from http://www.nces.ed.gov/.

National Science Foundation. *Survey of Earned Doctorates*. Data retrieved from http://webcaspar.nsf.gov/index/jsp.

U.S. News and World Report. America's Best Colleges. Various issues.

Winston, Gordon C. (1999). Subsidies, hierarchy and peers: the awkward economics of higher education, *Journal of Economic Perspectives,* 13 (1), 13-36.