Research in Higher Education, Vol. 47, No. 3, May 2006 (© 2006)
DOI: 10.1007/s11162-005-9388-y

RELATIVE SUCCESS? DETERMINANTS OF COLLEGE GRADUATION RATES IN PUBLIC AND PRIVATE COLLEGES IN THE U.S.

Marc Scott,*,† Thomas Bailey,** and Greg Kienzl**

Amid growing criticism of public universities, there is little discussion of what appropriate institutional evaluation would entail. Six-year graduation rates are commonly used, and public bachelors granting institutions have lower rates than private institutions, but with the growth in non-traditional college attendance, these can be misleading. We develop a regression analysis as a way to evaluate institutions serving vastly different populations. We do this with a dataset constructed from publicly available sources and focus on the evaluation of public colleges. We show that public colleges are able to do more with less: our models suggest that with equivalent resources and student populations, public schools would graduate a slightly larger percentage of students than privates. Since financial resources come from very different sources, we evaluate this finding closely.

KEY WORDS: college performance; educational economics; efficiency.

INTRODUCTION

In the 1960s and 1970s, the federal government and some states expanded their role in higher education, with the goal of increasing access for a larger segment of the population. With education so intricately linked to economic mobility, increased access to college was part of an attempt to provide opportunities to people from less economically advantaged backgrounds. Much of this effort went into increased financial aid, eventually resulting in need-blind admissions policies at many colleges.

*The Steinhardt School of Education, New York University, New York, USA.

**Institute on Education and the Economy Teachers College, Columbia University, 525 West 120th Street, Box 174, New York, NY 10027, USA.

†Address correspondence to: Marc Scott, The Steinhardt School of Education, New York University, 246 Greene Street, Suite 318E, New York, NY 10003, USA. E-mail: marc.scott@nyu.edu

Public colleges played a major role in this effort, with tax revenues offsetting tuition costs, and by the end of the 20th century, about 78% of college students were enrolled in state schools (Mortenson, 2000). By this same time, college graduation rates were in decline and public support for financing higher education had waned. Nationally, five-year graduation rates in the 1980s and 1990s dropped about 6 percentage points from near 58% to near 52% (Mortenson, 1998). Concurrently, a trend toward greater accountability and standards gained strength at all levels of education. Public demand grew for performance-based financing of public two- and four-year institutions.

The goal of this paper is to evaluate whether these trends suggest that public institutions are less effective than the privates. In the process, the paper contributes to the growing literature on the analysis and evaluation of college effectiveness. Assessment of institutional performance can be a difficult and complex task. This is especially true for public colleges, with their clear mandate of access to underrepresented communities. Colleges that enroll many poorly prepared students or students who are working or have family or financial responsibilities that compete with college are likely to have lower graduation rates. Thus, greater selectivity may improve measured performance, but comes into conflict with the public mission of the colleges. In addition, public institutions usually have little control over the state or local appropriations on which their budgets are based, or in many cases, on the tuition that they can charge.

Assessments must also use appropriate outcome measures. Graduation rates measuring the proportion of an entering class that have graduated within a specific number of years are the most common. But scholars have argued that colleges serve many purposes not captured by this measure and they have developed many alternatives (Bolt and Roberts, 1997; Lavin and Hyllagard, 1996; Muffo, 1996). While we acknowledge these concerns, our analysis will be based on six-year graduation rates for three reasons. First, they are at least one important measure of performance. Second, unlike most alternative measures, they are available for a large number of institutions. Third, much of the controversy over public colleges involves the language of graduation rates; using them allows us to compare our findings to the literature.

How should private and public colleges, with their different missions, serving different populations, be compared? We take the following approach to evaluation in this paper. Any cross-institutional comparison of graduation rates must take account of the resources available to the institution and the characteristics of the students who enter. While institution-level analysis of expenditures carries with it some technical limitations (Hanushek, Rivkin, and Taylor, 1996; McPherson and Patterson, 1990), some assessment of their role is clearly appropriate. Further, non-traditional or under-prepared students will take more time to graduate, even when the institution is doing a good job. While there is some debate over what determines the "traditionality" of a student, we follow Bean and Metzer (1985) and Johnson (1997) and define this population as comprised of older, part-time and/or commuter students.

Our study will document the role of institutional and student factors on six-year graduation rates using a regression analysis on data compiled from the following publicly available sources: American Survey of Colleges (ACS; The College Board, 1999); Integrated Post-secondary Education Data System (IPEDS): Institutional Characteristics Survey, Fall Enrollment Survey, Finance Survey (National Center for Education Statistics, 1999). The data source for each variable used in the regression analysis is given in column 1 of Table 1. Details of measure construction and reference to this table will be provided in Section Variables and their Construction. Factors suggested by theory and past research include: institutional resources, student academic characteristics, the traditionality of student attendance patterns and other student demographics, such as race/ethnicity and gender. A meaningful normative judgment about graduation rates should take all of these factors into consideration.

In Section Related Literature and Analysis Framework of this paper, we review the literature on persistence and attainment to motivate and guide the regression analysis. In Section Variables and their Construction, we describe the construction and choice of our measures. In Section Modeling Concerns, we develop the modeling framework and discuss methodological issues, including missing data concerns. In Section Models for Institutional Completion Rates, we fit the models and explore the differences in performance between public and private institutions. In Section Discussion and Implications for Policy, we summarize the findings and discuss their policy implications.

RELATED LITERATURE AND ANALYSIS FRAMEWORK

Much of the literature on college retention analyzes the completion probability for individual students and identifies characteristics of both students and colleges with high degrees of persistence and attainment. Academic and social attachment currently form the foundation of most research on persistence (Pascarella and Terenzini, 1991; Tinto, 1993). Thus institutional or social policy designed to increase retention is generally focused on strengthening student attachment, for example through improving student services or the quality of residential life. Race and

TABLE 1. Characteristics of Sample

			Means and SD		Means only			
			Full sample (across sector)		Full sample		Regression sample	
Variable	Data source[a]	N for full sample	Mean	SD	Private colleges	Public colleges	Private colleges	Public colleges
Six-year graduation rate (%)	ASC97	953	53.29	17.26	57.31	45.30	59.70	46.88
Public (Private)	IC91	1621	.34	.47	.00	1.00	.00	1.00
Urban (non-urban)	IC97	1621	.74	.44	.77	.68	.77	.67
In-state tuition ($1000)	IC91	1578	6.30	4.25	8.53	2.03	10.05	2.22
Instructional expenditures/FTE student ($1000)	F91/EF91	1605	4.57	5.17	4.33	5.02	5.01	3.96
SFR: FTE student to FTE faculty ratio	ASC91	1621	14.57	6.02	13.02	17.60	13.24	17.92
Difference between SAT 75th and 25th percentile	ASC91	1121	244.85	72.81	242.67	249.08	237.34	243.53
SAT 75th percentile	ASC91	1121	1071.29	135.30	1089.85	1035.39	1124.25	1045.58
SAT 75th percentile exceeds 1270 (indicator)	ASC91	1121	.07	.25	.09	.03	.12	.02
Percent full-time undergraduates (×100)	EF91	1620	76.92	18.68	78.21	74.39	82.40	78.82

Percent commuter undergraduates (×100)	ASC91	1581	54.68	29.64	45.71	72.57	36.98	67.86
Average undergraduate age	ASC91	1559	22.80	3.44	22.48	23.42	21.34	22.48
Percent female students (×100)	EF91	1578	56.58	15.02	57.23	55.32	56.97	53.13
Percent minority undergraduates (×100)	EF91	1621	21.06	22.60	20.36	22.45	14.97	17.78
Percent foreign undergraduates (×100)	ASC91	1409	3.80	5.13	4.57	2.20	3.67	2.15
Percent undergraduates w/financial aid (×100)	ASC91	1463	69.72	19.20	75.48	57.18	74.75	57.21
Number of undergraduates (×100)	EF91	1621	4.06	5.58	1.77	8.58	2.43	10.74
Percent graduate student (×100)	EF91	1621	14.34	17.58	13.25	16.47	13.56	16.06
Religious institution (×100%)	IC91	1621	39.54	.49	59.63	.00	60.57	.00
Number in sample		1621	See column 2		1075	546	459	215

[a]*Key to data source*:
ASC97: American Survey of Colleges, 1997–1998. The College Board.
ASC91: American Survey of Colleges, 1991–1992. The College Board.
IC97: Integrated Post-secondary Education Data System, Institutional Characteristics Survey, 1997–1998. National Center for Education Statistics.
EF91: Integrated Post-secondary Education Data System, Fall Enrollment survey, 1991–1992. National Center for Education Statistics.
F91: Integrated Post-secondary Education Data System, Finance survey, 1991–1992. National Center for Education Statistics.
IC91: Integrated Post-secondary Education Data System, Institutional Characteristics survey, 1991–1992. National Center for Education Statistics.

ethnicity may be related to the attachment process if, for example, minorities feel more or less isolated, depending on the ethnic mix of other students on campus (see Martin, 1990; Wells, 1997).

Most of the attainment literature has focused on traditional college-going aged students. Yet in the past 20 years, there has been tremendous growth in the non-traditional student population. Bean and Metzger (1985) and Johnson (1997) define these as older, part-time, and often commuter students. Commuter status is particularly important since non-residential students have less opportunity to develop the type of engagement that underlies the concept of social attachment. Metzger and Bean (1987) find that age and goals have a greater role in persistence and related outcomes for non-traditional than for traditional students. Based on an extensive review of empirical tests of Tinto's model, Braxton, Hirschy, and McClendon (2003) conclude that the "theory lacks explanatory power in commuter institutional setting" (p. 17). For commuter students, attachment to college is a more complex process. Public institutions tend to have relatively larger numbers of commuter and older students; in our data (to be described in detail shortly), about 46% of private college students and 73% of public college students are commuters. Public college students are on average about a year older than their counterparts in private colleges. Given Metzger and Bean's findings, it is imperative that age, commuter status, and related factors be used as controls in any performance assessment.

In addition, institutional characteristics span a broad range of potentially important factors, ranging from financial and other resources to measures of "selectivity." Difficulties in accurately measuring institutional resources are well-documented in Winston (1998), but one would be reluctant to make institutional comparisons without some control for these. For example, college "quality" is often judged by high school GPA or SAT scores of entering students. Pascarella and Terenzini (1991) point out that "'high quality' colleges start with a distinct advantage in terms of the academic ability, educational aspirations, level and clarity of career ambition, and family resources of the students they recruit and enroll" (p. 374). Since selective colleges enroll more accomplished students it is difficult to determine whether higher graduation rates among such colleges result from effective institutional practices or simply from enrolling students who would be more likely to graduate under any circumstances. Similar arguments can be made for allocation of institutional resources; how much money is spent and where it is spent should influence retention and completion. Because many of these factors are under institutional or local government control, their potential effect on outcomes is of direct interest to policymakers.

While we will draw upon the notions of academic and social attachment described in Pascarella and Terenzini (1991) and Tinto (1993), we must depart somewhat from their formulation for several reasons. First, we have noted that the growth in a less traditional college student population necessitates inclusion of additional student variables. Second, measures of student attachment are at best proxied by some of the characteristics we explore (i.e., no direct measures were available). In part, this is because the level of our analysis is institutional; our measures will reflect the aggregated profile of student and school characteristics. We have slightly different aims as well, in that we want to understand how institutions differ, rather than why a particular student-type departs from college. To this end, we also include some features of the institutional profile in the regressors. Recent work by Desjardins, McCall, Ahlburg and Moye (2002) and Smith and Naylor (2001) provide a template for the approach we will employ. Both authors begin with Tinto's model, but then depart from it to suit the differing aims of the analysis. For example, Smith and Naylor (2001) focus on the match between secondary school "majors" and those chosen in college. While the matching process clearly influences attachment, it is a much broader concern, with potentially very different policy implications.

Our evaluation will draw from several institutional analyses from both the United Kingdom and the United States, which we now describe. In the UK, Johnes and Taylor (1989) use a regression framework to evaluate college performance. They model a series of inputs, both student- and institution-based, including gender, ability, program of study, student to faculty ratio, percentage of commuter students, university type (e.g., technical college), and locale. Johnes (1997) looks at age of student and uses a refined set of student-oriented expenditure measures. Smith and Naylor (2001) use individual level data to evaluate the impact of academic and social attachment factors, such as "match" of major field of study, with the ultimate goal of providing an adjusted metric upon which institutional comparisons can be made. Academic attachment is proxied by preparedness measures such as A-level performance, while social attachment is captured in age and residence (on-campus or not) of student. The conclusions from all of these studies are comparable to individual-level analyses in the US in that they find that prior measured ability, age and social factors are important.

In the US, Astin and Oseguera (2002) and Mortenson (1997, 1998) report on research designed to determine which institutional and student characteristics lead to higher retention and graduation. The former use a regression analysis and point out that institution type (public, private, college, university, religious denomination), SAT score, GPA, race and gender all have an impact. Importantly, they find that the gap in six-year graduation rates between public and private colleges diminishes significantly, from 31% to about 7% when all of these characteristics are controlled. These figures are based on Appendix G of Astin and Oseguera (2002). For their weighted sample, private universities had rates of 80%, while public colleges were 49%—a gap of 31%. The expected difference based on the regression coefficients dropped to 7%. Their initial gap is much higher than the gap we find in our data in part because Astin and Oseguera compare private universities to public colleges and do not include four additional categories of four year institutions that they list in their Appendix G. In contrast, we compare all public and private four-year institutions. Because we lack the authors' weights, we cannot replicate their precise comparison using our sample. Mortenson (1997) regresses SAT scores, percent commuter and percent part-time on six-year graduation rates as part of a larger analysis designed to explain why public colleges had lower graduation rates. However, he does not include expenditure data and controls for student socio-economic status. Mortenson (1998) compares public and private school graduation rates using regression models and finds that private schools perform better, even when the "degree of selectivity" is controlled.

We extend these results in the following ways. First, we structure the regression so that four factors are evaluated simultaneously: institutional resources; academic selectivity; student non-traditionality (which may be related to social attachment); and additional student and institutional demographics suggested by theory and prior analyses. Most studies focus on one or two of these factors, rather than the complete set. This may or may not be a limitation, depending on the goals of each analysis. The motivation for our study was to assess institutional performance, particularly across sectors, and the literature suggests that each of the variables that we included had the potential to confound comparisons if omitted.

We provide details of the variables included for each factor in the next section but note here that we include at least three measures for each of these four factors. Our sample includes every not-for-profit four-year degree-granting institution (four-year colleges located in the US territories, such as Puerto Rico, Guam, and the Virgin Islands, were excluded from the study), drawn from publicly available, comprehensive sources, and we address any missing data issues using multiple imputation techniques. An additional contribution we make is methodological. In the past, institutional analyses have included public/private indicator variables. This approach assumes that all of the other variables in the model have the same influence in public and private institutions. We test this assumption, and find strong evidence that it is false. We then document the implications of differential returns in private and public colleges for performance assessment via an Oaxaca (1973) decomposition which, for example, can estimate what graduation rates would be at public colleges if they had the institutional and student characteristics of private colleges. This gives us a better understanding than has been available so far of what differentiates public from private institutions.

VARIABLES AND THEIR CONSTRUCTION

Dependent Variable

To understand how our institutional measures are constructed, one must first understand the way in which students are tracked over time. Our dependent variable is an institution's six-year graduation rate in 1997, which is the percentage of the full-time freshman class entering in 1991 that has graduated by 1997. As a result, some students will have been in the institution for six years, yet variables for institutional characteristics can change from year to year. We choose to use data from this cohort's first year of enrollment, 1991, to characterize the institutions. There is a reasonable consensus in attainment theory that the first few years of enrollment are crucial to persistence. Moreover, many students who drop out or graduate never experience the college in the later years of this time period. Our graduation rates come from the 1997 to 1998 ACS (The College Board, 1999); sources for this and all other covariates are summarized in the first column of Table 1. We note that nearly every remaining data source is based on the 1991–1992 academic year, so the origin date is omitted from the data discussions to follow.

The graduation rates do not account for transfers out of school, which in many cases result in an eventual graduation at some other institution. From this point of view, transfer may be a "successful" outcome and neglecting it makes colleges look worse than some notions of accountability might suggest. Nevertheless, many state accountability systems are based on institutional rather than system wide graduation rates (which would account for transfers). And enrollment managers are also concerned with retention at their institutions, rather than system-wide student experience. Moreover, this will be a problem for our comparative analysis of public and private institutions only if transfer rates were significantly different in the two sectors. We assessed the potential distortion from transfer using a publicly available sample from the Beginning Post-Secondary Students (BPS; National Center for Education Statistics, 1996) of first-time college freshmen. We found that transfer rates for public and private institutions did not differ significantly (we restricted the BPS sample to students initially enrolling in four-year colleges, and evaluated the transfer rates over a five-year period). We found that transfer activity occurs for about 30% of students in both public and private colleges, making our six-year graduation rate comparable across sectors.

Explanatory Covariates

Institutional

We use several available and reliable revenue and expenditure measures: tuition, instructional expenditures per student, and student to faculty ratio (SFR). While the source of revenues varies widely between and within college sectors, aggregated per student expenditure reflects both available resources and institutional distributive policy. The SFR indicates how a particular set of resources, with clear potential to influence academic achievement and persistence, is allocated. We include tuition primarily for empirical reasons; the variation in tuition among private colleges is substantial—the interquartile range is about $4000 in 1991, with a median value of $8000. While the variation reflects selectivity, in part, it also represents available resources (tuition increases are often justified by linking them to increased costs). So our three institutional components reflect one measure of expenditures per student, one aspect of faculty deployment, and one measure of financial resources. While many more could be considered, issues of multicollinearity and interpretability impelled us to limit our analysis to these three.

For a measure of tuition, we use the in-state tuition at public institutions, and use the common tuition at private ones, given in the IPEDS institutional database. For expenditures per student, we take the 1991 annual instructional expenditures divided by the number of full-time equivalent (FTE) students in the fall of that year. These measures come from IPEDS finance and enrollment databases. We construct the student-to-faculty ratio from fall FTE student enrollments and FTE faculty appointments, available in the ACS. We note that other related datasets, such as the US *News and World Report* college rankings data (*US News and World Report*, 2004) used in Mortenson (1997) construct many ratio measures from raw counts of student enrollments. FTE-adjusted measures are much more appropriate (see Winston, 1998); for example, some universities have shifted faculty appointments from majority full-time to part-time, so any per student or per faculty measures should be adjusted to reflect this. Because expenditures for graduate and undergraduate students are aggregated in the IPEDS, we had to include graduate students in the denominator of our two "per student" measures. To correct this potential bias, controls for percent graduate student were included in the regression.

Academic Selectivity

Typical measures of college selectivity are SAT and high school ranking; incoming freshman SAT scores provide very reliable comparisons, so we use these. Rather than take average SAT score, we obtain more information by including the 25th and 75th percentiles of math and verbal scores computed within each institution. For example, at SUNY at Buffalo, the SAT 25th and 75th percentiles for verbal are 450 and 550, while they are 550 and 650 for math. We average the corresponding verbal and math quartiles, in this case $500 = (450 + 550)/2$ for the 25th percentile (combined score) and $600 = (550 + 650)/2$ for the 75th percentile. This is not strictly equivalent to taking the quartiles of student-level combined SAT scores at each college, but it is a close proxy. In our analyses, we use the combined 75th percentile to track the academic ability/preparedness of the upper tier of students within a school. We use the difference between the 75th and 25th percentiles to capture the variability within an institution—or the extent of its homogeneity. When the 75th percentile exceeds 1270, the school ranks in about the top 10% of schools in our regression sample, and we set a "highly selective" indicator variable to one. The source for these covariates is the ACS database.

Non-Traditionality

All three aspects of non-traditionality were available, so we include measures of percent part-time, percent commuter and average student age (all undergraduates) in our analysis. These measures are only available in aggregated form in the IPEDS enrollment and ACS data, but given the aggregated level of analysis, this is appropriate. Since this variable is based on the entire student population, it captures potential indirect (contextual) effects of student characteristics. All three measures are expected to indirectly gauge the degree of student involvement in campus life, or social attachment. The percent part-time variable warrants additional mention in that no one in our cohort begins enrollment part-time, therefore, these measures reflect contextual effects, in this case of the impact of large part-time populations on full-time enrollees.

Demographics

Our demographic measures were chosen to reflect features of the school or its students that could confound comparisons, were they omitted. For example, individual graduation rates differ by gender and race/ethnicity, so colleges with disproportionate numbers of women or minorities would be expected to have different graduation rates, all else equal. We include several other demographic factors that have similar potential to confound: foreign student populations (who may or may not be as well prepared for English instruction), urbanicity, size of school, and religiosity. The latter was included because the sample contains a large number of religiously affiliated schools, and social and academic integration could be very different on such campuses; the same argument can and has been made for school size (Pascarella and Terenzini, 1991). We did not include the full array of Carnegie classification indicators, as these reflect a very different attempt to construct comparable schools (see Carbone and Winston, 2004, for a related comparison made at that level). Lastly, a crude measure of socio-economic status was included using information on financial aid received.

A college is designated as urban if it is geographically located in a Metropolitan Statistical Area (MSA) as defined by the Census Bureau (For the basic standards for defining MSAs see http://www.census.gov/population/www/estimates/mastand.html). The IPEDS 1997–1998 institutional database identifies the school location, and we link that with Census data to derive an urban indicator. We calculate school-level percent female and percent minority from the IPEDS undergraduate enrollment figures (full- and part-timers are included). Percent foreign-born and percent receiving financial aid variables are based on first-year undergraduate enrollments given in the ACS database. Percent receiving financial aid is an imperfect measure of socio-economic status (SES); moreover, at the institution-level, (presumed) negative impacts of lower SES are confounded with the (presumed) positive impact of financial aid. While our measure is far from perfect, some control for SES is clearly warranted. We define minority as anyone self-identified as non-White, or Hispanic. We do not have a direct measure of the prevalence of English as a Second Language, nor do we have a direct measure of remediation needs, so percent of foreign-born students will serve as a very rough proxy for the former. School size and graduate to undergraduate ratio are derived from the IPEDS enrollment database. Whether the school is religiously affiliated comes from IPEDS institutional database.

Sample Description

Table 1 presents summary statistics that portray a sense of the scale of the explanatory measures and of compositional differences between public and private institutions. In what follows, we discuss the full sample, but as a result of missing data, we are not able to use the full sample in our regressions. We will provide some indication of how the full sample differs from the regression sample in the next section. Table 1, column 1 summarizes the data sources used in the variable construction just described, and column 2 the number of non-missing records for each variable. Columns 3 and 4 provide means and standard deviations for the full sample, pooling across sector. We see that the overall graduation rate is a bit over 50%, but that the variation is substantial, with a standard deviation of 17%. The average undergraduate age, which is nearly 23, with standard deviation 3.4, is bit surprising. This average is based on all undergraduates, so 20 is to be expected at a more ''traditional'' college. This clearly shows the importance of non-traditional students in this sample. About 40% of the colleges in the sample are religiously affiliated.

Private and public college means for these same variables in the full sample are compared in columns 5 and 6. We confirm that the gap between these sectors is substantial, with average graduations rates of 57% at private college and 45% at publics. Tuition costs are, not surprisingly, much higher at a private college—over four times as high in 1991 (using the Consumer Price Index, $10,000 in 1991 is about $15,000 in 2006). Some of the larger differences between sectors include SFR, which is 13 at privates and over 17 at publics; percent commuter is vastly different, at 46% and 73%, respectively; undergraduates a solid year older in public schools; public schools are about five times as large. Financial aid is received by 75% of private school students, as compared to 57% of public school students. We see that this measure reflects the dramatic difference in cost of attending private colleges more than pure SES, and given the variation in aid policy particularly in private schools (e.g., non-need based allocations), we expect that this measures does a better job assessing SES in the public sector. In terms of selectivity, the upper quartile of private school students score about 55 points higher on their combined SAT (total scale of 400–1600). Nine percent of private and three percent of public schools are highly selective.

MODELING CONCERNS

Missing Data

A total of 1676 four-year institutions meeting our criteria were identified. However, not all of the desired variables were available for each college. Among the most heavily missing variables were SAT scores and graduation rate itself. The six-year graduation rate came from the American Survey of Colleges (College Board, 1999), and most of the missing rates reflect survey non-response. In a few cases, particularly for large public universities, the differences in how institutions recognize branch campuses and/or affiliated colleges resulted in some data loss. For example, all the branch campuses of a multi-campus university (such as Rutgers University) might report the same graduation rate as the main campus. Minor data cleaning was performed in these cases to avoid multiple instances of the same information, yielding 1621 cases. To recover some SAT scores, we imputed values from available ACT scores using a conversion chart provided by the College Board (2001). We know that schools that only report ACT scores are commonly located in the central portion of the US. In order to be sure that our findings were not driven by a "location" effect, we included an ACT indicator in several of our models (not reported here). We found no such differential effect and hereafter ignore this issue. We were able to recover 293 cases with this effort, increasing our complete sample by about a third. For the remaining explanatory variables, missing data was much less of an issue, but when a listwise deletion approach is taken (cases missing any variable are dropped), we are left with 947 complete cases, ignoring the dependent variable, and 674 cases if we include graduation rate in the deletion procedure.

Biases can result when a regression sample is much smaller than the full representative sample. How does our regression sample relate to the full sample? The means for the full sample were already discussed; in Table 1, columns 6 and 7, we provide the means for the restricted regression sample. A comparison of this smaller sample versus its complement (the excluded cases) revealed that the differences are statistically significant for more than half of the covariates. We note, for example, that our restricted sample consists of larger schools (in aggregate, about 1000 more students), with higher tuition, and lower expenditures per student on average. Sometimes the differences between samples are stronger for private colleges, but certainly not always.

Thus the missing data that would result in our sample reductions are not "missing completely at random" (MCAR, Little and Rubin, 1987). When data are not missing completely at random, then deleting observations that have missing data can result in biased estimates.

On the other hand, data could be missing in a systematic way that can be predicted using remaining explanatory variables. This is referred to as "missing at random" (MAR) and resulting biases can be corrected using imputation techniques. We test whether the data are MAR by trying to predict whether each record contains any missing data (that is, was it deleted from the regression sample?) using a subset of the explanatory covariates (those for which few data are missing). We performed a logisitic regression, by sector, of these explanatory covariates on an indicator variable set to one for the regression sample and zero otherwise. The findings, not presented here, suggest that the independent variables are reasonable predictors of whether a data point is missing. More importantly, the outcome variable, graduation rate, was not significant in the models that included it, suggesting that whether a data point is missing does not depend on the magnitude of the dependent variable—this is required for the MAR assumption to be plausible.

When the MAR assumptions hold, multiple imputation (Alison, 2001; Schafer, 1997) makes use of all of the information in the full sample and yields unbiased parameter estimates. Thus the findings from the restricted sample are representative of the results we would obtain from an analysis of the complete sample. Multiple imputation involves using the covariance matrix estimated from all observed data to impute every missing covariate—including any missing responses—several times, to create several complete datasets. (Imputing response variables is standard in the statistical literature. See Horton and Lipsitz (2001) for some discussion and related references.) In our case, we created eight complete datasets. The imputation procedure adds a random component to each observation to ensure that the same level of variability exists in each imputed sample as in the observed data. The regression analysis is performed on each imputed dataset, and the results are combined in a manner that accounts for the uncertainty due to missing data. We will gauge the robustness of our findings by the similarity of our multiply imputed and restricted sample findings.

Level of Analysis

In our models, we do not examine the individual effect of any student-level characteristics. For example, we do not directly know how being a commuter affects one's likelihood of graduating, but we do know the aggregate effect of having a large proportion of commuting students on the graduation rate of an institution. Thus the coefficients for a given variable, say percent commuter, will combine the direct effect of commuting on graduation (commuters themselves may graduate less often) and the indirect effect (the presence of many commuters may influence graduation rates of other students perhaps by diverting resources to the special needs of commuters or through peer effects). Hanushek et al. (1996), in the context of state-wide secondary school evaluation, note that aggregated analyses tend to overstate the effects of resources. We will keep this in mind when interpreting our regression findings.

MODELS FOR INSTITUTIONAL COMPLETION RATES

Six-year graduation rates take a value between zero and one. We model the response using grouped logistic regression, which is appropriate for an aggregated dependent variable that is a rate or percentage. The frequently used linear probability model in which the ratio between zero and one is entered directly into the ordinary least squares regression is problematic because expected values based on the regression coefficients can fall outside of the possible range of the dependent variable (between zero and one). The grouped logit specification constrains expected values of the dependent variable to be between zero and one, providing for a non-linear response curve, dampening the absolute impact of each factor on graduation rate when the rate is very low or high (see Dey and Astin, 1993 for a comparison of logit, probit, and OLS models for studying retention). Formally, our model can be written down as follows:

$$\text{logit}(p_i) = X_i\beta,$$

where $\text{logit}(p_i) = \log(p_i/(1-p_i))$, and X_i is the row vector $[1, X_{1i}, X_{2i}, X_{3i}, X_{4i}]$, consisting of a constant, institutional characteristics X_{1i}, academic selectivity factors X_{2i}, non-traditionality factors X_{3i}, and demographic controls X_{4i}. The behavioral aspects of this model operate in the aggregate, so that all n_i individuals attending institution i are given the same overall treatment: a combination of institutional resources and effectively a "profile" of characteristics of the student population, yielding a proportion p_i of graduations in six-years. The size of the treated population, n_i, is given by total FTE enrollment.

In practice, grouped logistic regression models can be estimated using two-stage weighted least squares, using the logit of the observed graduation rate as the outcome. The function `glogit` in STATA™; uses this approach and was employed in all of the findings herein (see Greene, 1993 for details). The R^2 reported by STATA™ reflects the proportion of the variance in the transformed responses ($\text{logit}(p_i)$) explained by the model, and thus is not a "pseudo-R^2" as is commonly reported in logistic models.

We begin by fitting a model using all of the available cases for which the graduation rate and the main institutional, student, and control variables are available. The variables are entered in their raw form (untransformed), with some attention to the "unit," so dollars are given in thousands; percentages are represented by 100× proportion (e.g., 1%, rather than .01). SAT level and range are rescaled to units of 100 points, while age is centered near the full-sample private college mean of

22.5. Some covariate distributions are skewed; we considered non-linear transforms of the variables but decided against it for the following reason. The transform to symmetry for this data is the natural log, and logistic modeling captures much of the same effect; that the impact of covariates enters in a way that is proportional to the level at which it is being evaluated (the effects are relative). We initially fit a pooled public and private college model, including indicator variables for sector. Specification tests revealed that the covariates function differently in the public and private colleges ($p < .0001$, overall). Specification tests on each of the four variable groupings indicate that institutional, non-traditionality measures, and demographic control variables all function differently ($p < .0001$, each group), while the school selectivity effects were not statistically different ($p > .10$). Based on these tests, we decided to model each sector separately rather than have a common model with a large number of interaction terms. Tests for multicollinearity suggest it is not a problem; the variance inflation factors (VIFs) for the covariates had a maximum of 3.4, well below the commonly accepted threshold of 10.

In Table 2, we report the raw regression coefficients, their significance, and the marginal effect. For the latter, we evaluate the change in probability of the outcome (in this case proportion graduating) for a one unit change in the covariate, evaluated at the mean for all covariates, following the guidelines given in Petersen (1985). We made three exceptions: for the urban, highly selective, and religious affiliation indicators, we evaluate the probability first at level 0, and then at level 1. The marginal effect, often referred to as ΔP, has a regression-like interpretation: it is the expected change in the absolute level of the outcome (proportion graduating) for a one-unit change in the covariate *ceteris paribus*. In what follows, we emphasize differences between sectors or samples and defer a discussion of the substantive implications of the findings until the final section.

Looking first at private colleges, in the first two columns, the only institutional factor that is significant is instructional expenditures per student. Raising these by $1000 is associated with a tiny .44% increase in graduation rates, holding all else constant. There is a notable absence of student-to-faculty ratio effects, but the effect may be masked by a number of additional controls. In particular, expenditures per student clearly include allocations to faculty salaries. Tuition effects are potentially masked here as well. While these factors are less influential than one might have expected, there is very little empirical evidence in the literature that contradict our results. Allison (2001) and Johnes and Taylor (1989) find insignificant student-faculty-ratio effects, while

TABLE 2. Graduation Models for Schools in Private, Public Sectors

	Restricted sample				Full sample (multiply imputed)			
	Private colleges		Public colleges		Private colleges		Public colleges	
Independent variables	Coefficient	Mg. effect	Coefficient	Mg. effect	Coefficient	Mg. effect	Coefficient	Mg. effect
Institutional resources								
In-state tuition ($1000)	.0098	.0024	.0366	.0091	.0195**	.0048	.0423†	.0103
	(.0082)		(.0262)		(.0073)		(.0214)	
Instructional expenditures/ student ($1000)	.0183**	.0044	.0702**	.0174	.0163*	.0040	.0412**	.0101
	(.0071)		(.0166)		(.0075)		(.0138)	
Student to faculty ratio	.0047	.0011	−.0019	−.0005	.0029	.0007	−.0049	−.0012
	(.0048)		(.0056)		(.0061)		(.0036)	
Academic attachment								
Difference SAT 75th and 25th percentile (×100)	−.1761**	−.0432	−.1925**	−.0470	−.1836**	−.0458	−.1716**	−.0411
	(.0365)		(.0314)		(.0395)		(.0386)	
SAT 75th percentile (×100)	.2980**	.0698	.3186	.0794	.2709**	.0659	.3042**	.0753
	(.0264)		(.0270)		(.0279)		(.0263)	
SAT 75th percentile exceeds 1270	.3119**	.0729	.3061†	.0763	.2869**	.0697	.1885**	.0464
	(.0744)		(.1578)		(.0877)		(.1218)	
Social attachment								
Percent full-time undergraduates	.0042**	.0010	.0133**	.0033	.0048**	.0012	.0107**	.0026
	(.0015)		(.0023)		(.0016)		(.0017)	
Percent commuter undergraduates	−.0017	−.0004	−.0059**	−.0015**	−.0017	−.0004	−.0050**	−.0012
			(.0016)		(.0012)		(.0014)	
Average undergraduate age (centered)	−.0294**	−.0071	−.0372*	−.0092	−.0143	−.0035	−.0439**	−.0106
	(.0108)		(.0148)		(.0111)7		(.0125)	

Demographics								
Urban	.0485	.0117	.1054*	.0262	−.0040	−.0010	.1215*	.0298
	(.0526)		(.0501)		(.0499)		(.0553)	
Percent female students	.0050**	.0012	.0131**	.0032	.0049**	.0012	.0116**	.0028
	(.0014)		(.0031)		(.0013)		(.0025)	
Percent minority undergraduates	−.0041*	−.0010	−.0030*	−.0007	−.0020	−.0005	−.0026*	−.0006
	(.0017)		(.0014)		(.0013)		(.0010)	
Percent foreign undergraduates	−.0260**	−.0063	.0011	.0003	−.0148**	−.0037	.0045	.0011
	(.0055)		(.0074)		(.0040)		(.0066)	
Percent receiving financial aid	−.0060**	−.0014	−.0004	−.0001	−.0044**	−.0011	−.0001	.0000
	(.0013)		(.0011)		(.0015)		(.0012)	
Number of undergraduates (×1000)	.0069*	.0017	.0071*	.0018	.0100**	.0025	.0077*	.0019
	(.0034)		(.0030)		(.0034)		(.0028)	
Percent graduate student	.0036**	.0009	−.0044†	−.0011	.0027†	.0007	−.0011	−.0003
	(.0013)		(.0026)		(.0014)		(.0023)	
Religious institution	.1173**	.0281	n/a	n/a	.0931†	.0230	n/a	n/a
	(.0395)				(.0495)			
Constant	−2.9249**	n/a	−4.6828**	n/a	−2.8066**	n/a	−4.2953**	n/a
	(.3459)		(.4570)		(.3822)		(.4295)	
R^2	.7437		.8476		.6843		.8192	

Source: Authors' estimates based on data compiled from IPEDS and College Board ACS.
Standard errors in parenthesis. **Significant at the .01 level; *significant at the .05 level; †significant at the .10 level.

Johnes (1997) finds significantly negative effects, but only in 2 of the 15 major fields of study evaluated. The latter two studies are based in the UK; we know of no recent US study that includes both tuition and expenditure data in a college graduation model. Astin and Oseguera (2002) may have included institutional financial data in their step-wise regression procedure, but none were reported in the final models chosen (implying non-significance).

Selectivity effects are quite strong in private colleges. Whether this is prior academic ability or subsequent academic attachment is not known. Whatever the mechanism, an increase of 100 points in the upper quartile of student SATs yields an expected gain of about 7% in graduation rates. The negative effect for the gap between upper and lower SAT quartiles suggests that a school with a broader *range* in achievement can expect 4.3% lower graduation rates for every additional 100 points. The indicator for high selectivity has substantial positive impact, at 7.3%. The strength of the prior achievement effect reported here is comparable to Astin (1997) and Mortenson (1997); the sign of the effects are consistent with models in Astin and Oseguera (2002), which contain a much larger set of individual-level variables.

Measures of non-traditionality matter as well—only commuting was found to be insignificant for private colleges. Percentage full-time matters significantly in both statistical and substantive terms. A private college that is 90% full-time will have a 4% higher graduation rate than one that is 50% full-time. A student population that is on average one year older, being that much less traditional, lowers the school's expectations by about 1%, all else equal. These findings match the expected direction of similar effects reported in the literature, under a broader set of simultaneous controls. Mortenson (1997) reports a large negative effect for colleges with large part-time attendance (this is the indirect effect on full-timer completion), DesJardins et al. (2002) report a similarly signed age effect, while Smith and Naylor (2001) in the UK and Astin and Oseguera (2002) in the US find significantly negative effects for commuting students.

Our additional demographic controls function about as expected and all are significant. We find that schools that are more female, larger, and religiously affiliated are expected to have higher graduation rates. Female and religious affiliation effects are well known. Of course, our female effects are likely to be direct rather than contextual effects: women graduate more frequently individually, so this influences the institution as a whole. School size effects are less consistent in the literature (Pascarella and Terenzini, 1991); Allison (2001) and Astin and Oseguera (2002) find evidence of positive effects for larger schools, but these effects tend to be small, like ours. The negative effects for percent foreign-born, minority, and financial aid recipients (our proxy for SES) are reasonably consistent with the literature (see also Alfonso, Bailey, and Scott, 2005). Urban effects are not significant in the private sector.

Turning our attention to public colleges, in columns 3 and 4, there are some surprising differences. These have broad implications that we will subsequently explore using the Oaxaca decomposition. We note that in this sector, an institutional increase of $1000 in instructional expenditures per student is associated with near 2% gains—nearly four times the effect for private colleges. The academic attachment effects are comparable, but highly selective public schools do not show a significant differential graduation rate. The lack of significance may result from a small sample size (there are only five such colleges in this sample). Measures of non-traditionality matter much more. If we compare a public college that is 90% full-time to one that is 50%, then the former should have about a 13% higher graduation rate; this is about three times the comparable effect at private colleges. It is instructive to note that public colleges apparently do better at the margin with *more* traditional students, yet they are more likely to enroll non-traditional students. The biggest shift in the role of controls is that our SES proxy (financial aid received) matters less, which is consistent with the mission of public colleges. Urbanicity has a positive and significant effect in this sector, and more demographic controls are found to be insignificant. The female effect is 2.5 times as strong in public colleges; a larger gap between male and female rates emerges as a crucial feature of the public sector.

We turn to the findings from the multiply imputed (full) dataset, presented in Table 2, columns 5 through 8. For private colleges, a pairwise comparison of the restricted and full sample findings reveals some change in the significance of coefficients, but no statistically significant differences between the two samples. Tuition effects became significant, so perhaps we increased the variation in that variable using the full sample, obtaining a more precise coefficient estimate. Age and many of the demographic controls are no longer significant in private colleges. For public colleges, there are no changes in significance, but the role of instructional expenditures drops by nearly half. These comparisons suggest that our restricted sample findings are robust, up to our level of uncertainty, since there is little change in the magnitude and no change in the direction of the effects.

Oaxaca Decomposition

The goal of this research is to document the factors that help to explain the substantial gap in private vs. public college graduation rates. A challenge when there is an interaction between sector and covariate effects is that assessment of any pure sector effect is problematic. Viewing such an analysis as an ANCOVA, we would require that the slopes for all covariates have the same influence on graduation rates across sectors, but they do not. Taking account of these differential effects greatly weakens prior arguments that private colleges outperform publics. One way to compare the two sectors is to evaluate the differences between them at various levels of the covariates: we might find, for example, that for lower levels of some inputs, public colleges outperform privates, and vice versa. The Oaxaca decomposition is a variant of this approach and provides a useful framework for comparisons.

The Oaxaca decomposition partitions an observed gap into a portion attributable to compositional differences between groups and a portion attributable to differential returns to the inputs. Formally, we estimate two models for graduation rates:

$$\text{logit}(p_{iv}) = X_{iv}\beta_v,$$

$$\text{logit}(p_{ib}) = X_{ib}\beta_b,$$

where 'v' and 'b' identify private and public colleges, respectively. The overall gap is $\text{E}(\text{logit}(p_v)) - \text{E}(\text{logit}(p_b))$. The gap due to different characteristics is $\bar{\text{X}}'\hat{\beta}_b$, while the gap due to differing response is $\bar{X}_v'\hat{\beta}$, where $\bar{\text{X}}$ is the mean vector of characteristics for private colleges minus the means for publics and$\hat{\beta} = \hat{\beta}_v - \hat{\beta}_b$. The gap associated with characteristics reflects the gap predicted by the models when both sectors function identically. The gap associated with response of the dependent variable to the covariates predicts the gap under a scenario in which public colleges have the same inputs as private colleges. The sum of these two gaps is the overall gap.

Table 3, Panel A presents this decomposition for both restricted and full model estimates. In addition to reporting the aggregated, or total, compositional and response effects, we list the contribution of institutional, academic, social, and demographic components to these sums in the table rows. The first row provides an "unexplained" returns component that reflects the difference in the constants for each model. This estimates the gap when all covariates are zero, which is a meaningless direct comparison, but nevertheless can be interpreted as a public college handicap, that diminishes at more meaningful levels of the

TABLE 3. Oaxaca Decomposition

	Restricted sample GAP (*.606*)		Full (imputed) sample GAP (*.430*)	
	Composition	Response	Composition	Response
(A) Summary by category				
Unexplained		1.758**		1.489**
Institutional resources	.369†	−.443	.269†	−.201
Academic selectivity	.291**	−.192	.188**	−.383
Non-traditionality	.271**	−.606*	.216**	−.314‡
Demographic	.016**	−.859**	−.001	−.833**
Total	.948**	−.342	.672**	−.242
(B) Category components				
Unexplained		1.758**		1.489**
Institutional resources				
In-state tuition	.287	−.270	.275*	−.194
Expenditures/FTE student	.074**	−.260**	−.028	−.108
SFR	.009	.087	.023	.102
Academic selectivity				
SAT 75th–25th percentiles	.012	.039	.011	−.029
SAT 75th percentile	.251**	−.232	.166**	−.363
Highly selective school	.029	.001	.011	.009
Non-traditionality				
Percent full-time	.047**	−.751**	.041**	−.462*
Percent commuter	.181**	.154*	.134**	.149
Average undergraduate age	.042*	−.009	.041**	.000

TABLE 3. *(Continued)*

	Restricted sample GAP (*.606*)		Full (imputed) sample GAP (*.430*)	
	Composition	Response	Composition	Response
Demographic				
Urban	.011	−.044	.011	−.097
Percent female students	.050**	−.458*	.022*	−.385*
Percent minority	.008	−.018	.005	.012
Percent foreign	.002	−.100	.011	−.088*
Percent w/financial aid	−.006	−.419**	−.001	−.329*
Number of undergraduates	−.059**	−.001	−.053**	.004
Percent graduate student	.011	.109**	.003	.050
Religious institution	.000	.071**	.000	.056
Total	.948**	−.342	.672**	−.242

**Significant at the .01 level; *significant at the .05 level; †significant at the .10 level; ‡ $p = .11$

Note: Due to the complexity of significance assessment for the full (imputed) sample, upper bounds on p-values were computed—actual p-values may be smaller covariates. The total for the response column evaluates each model at a single level of the covariates, providing a more meaningful comparison. With private colleges as the reference category, positive component values reflect better performance for that sector.

Focusing on the restricted sample, we first note that the overall gap between the graduation rates for the two sectors is .606 on a logit scale. Since the regression is conducted on this scale, it is appropriate to perform the decomposition on it as well. Nearly all of the compositional effects are large and positive. The response effects associated with the variables are all negative suggesting that at the margin public college rates are more responsive than private rates to changes in variables. On the other hand, the unexplained gap is large and positive. A significant total compositional effect of .948 suggests that the gap would be even wider, given input differences, if all resources were utilized equivalently across sectors. The total returns effect of −.342 suggests that were inputs equalized, public colleges would perform a bit better than privates, but we note that the difference between this total and zero was not statistically significant. Evaluation of this hypothetical is complex; we will discuss competing explanations in the next section.

Turning to the components of these effects, we highlight the significant factors that drive the aggregated findings. For compositional effects, the non-demographic categories all contribute significantly and about a third each to widening the gap. (Institutional resources are significant at the .10 level ($p = .07$). All other effects are significant at the .05 level.) These categories are in turn driven by a set of sub-effects displayed in Panel B of Table 3. Namely, institutional resource gaps are driven by a significant and large difference in instructional expenditures per student; the SAT 75th percentile drives the academic effect; and while all measures of social attachment are significant and positive, differences in percent commuter accounts for about two-thirds of this component. One way to interpret these findings is that private college success is driven by larger numbers of more traditionally enrolled and academically prepared students. Student selectivity and unmeasured aspects of SES must be considered in any evaluation, and the Oaxaca breakdown for compositional effects highlight this.

With respect to the response of graduation rates to the different variables, it is surprising that the resources category is not significant overall. However, the category aggregates both positive and negative effects (private returns to SFR are a bit better), and the tuition effect is estimated somewhat imprecisely, contributing to aggregate imprecision. The instructional expenditures effect is large, significant and positive, suggesting that at the margin, public colleges have greater response in this domain. This should be considered when policies for public college expenditures are being debated. We find that selectivity effects are not significant overall, but that most of the non-traditionality effects are, as are most of the demographic components. The three largest contributions are for percent full-time, female and receiving financial aid, suggesting that publics do much better with populations that are higher on these inputs.

Putting the pieces together, the overall gap of .606 is predominantly driven by differences in inputs. Private colleges have student characteristics and college inputs that tend to be associated with higher graduation rates. But the differential responses to inputs in the public sector offsets the extant gap somewhat. Thus, a very different scenario is ruled out by our Oaxaca decomposition; namely, that the gap is largely due to lower response to resources in the public sector—the opposite conclusion is more consistent with our findings. We have pointed out that Hanushek et al. (1996) argue that aggregate institutional data may yield upwardly biased estimates of expenditure effects. However, were the institutional factor effects reduced to zero, the aggregate gap associated with returns would still be statistically zero, so our overall finding is robust.

The fact that the gap widens (total for column 1) when responses are equalized (when the same coefficients are used for both public and private institutions) is important. First, it highlights the disparities in public/private inputs, since this is the only remaining sectoral difference. The widening also reflects differences in input utilization across sectors, because the public sector coefficients employed yield higher response rates in our models. Given the relative magnitudes of compositional and response factors, it is clear that compositional differences drive the overall gap.

The theoretical closing of the gap (total for column 2) when compositional differences in inputs are eliminated warrants more discussion. Our models suggest that were public colleges to have the same inputs as privates, they would render the original gap of .606 insignificant. Breaking this down into four components, the first two of which are not statistically significant, yields: .443 units for higher responses to expenditures; .192 for selectivity; .606 for traditionality of student body; and .859 for higher responses to demographics (in particular, percent female and receiving financial aid). All of these reductions counterbalance an unexplained portion of the gap, which is quite large at 1.758, yielding a negligible net expected gap. A key finding is that at the margin, public schools show higher returns to most of their resources. This finding reflects well on public institutions: were public schools to have increased resources, more academically prepared students, and a more traditional student body, our models suggest that the graduation rates at public institutions would be slightly higher than at the privates, all else equal.

The decomposition for the full (multiply imputed) sample is given in the remaining two columns of Table 3. We see that an overall smaller gap of .430 is partitioned into a compositional effect totaling .672 and a response effect totaling −.242. While the magnitude of these effects is reduced, the overall pattern is about the same. Turning to the components within each effect-type, they are all about two-thirds as large as for the restricted sample, and all but one (demographic composition effects, which are trivial) maintain their signs. Significance patterns mirror the restricted sample reasonably well; tuition and expenditure compositional effects notably swap significance. The evidence still strongly suggests that public colleges do better with their resources and inputs—at the margin—that the gap would widen if the public and private response rates were equalized.

DISCUSSION AND IMPLICATIONS FOR POLICY

We have established that the gap in public vs. private college graduation rates is best understood through a regression model that captures the differing responses to the inputs simultaneously. We have shown that much of the gap between the average graduation rates of public and private institutions can be explained by the different characteristics of their students. But we have presented evidence that at the margin, the public institutions do at least as good a job with their population of students than the privates would were they to enroll a similar student body. Stated another way, the gap would be larger if public and private colleges used their inputs in precisely the same manner. Public colleges apparently use their inputs more effectively than the private institutions; that is, our models suggest that were their inputs to resemble private college levels, the gap would effectively be zero. Institutional resources, selectivity and student traditionality and demographics matter significantly in most partitions of the gap, but the aggregate impact of institutional resources seems to be measured too imprecisely to obtain overall significance. We find, for example, that instructional expenditures per student is significant in the Oaxaca decomposition for compositional and response effects only in the more restricted regression sample.

It is important to avoid attributing the superior responses in the public sector suggested by the Oaxaca decomposition to pure institutional efficiencies. The college is only one actor; the student is the other. Public colleges are making "better use" primarily of student inputs, such as their academic and social attachment, which tend to be lower. This is consistent with a scenario in which these students are more motivated to complete or public colleges are more effective at moving them through their schooling. If the latter were true, it could be due to institutional efficiencies or lower absolute standards. Reconciling these competing hypotheses is beyond the scope of this analysis.

While we have indicated the need for caution when interpreting these findings, particularly with respect to institutional resources, we now expand this discussion. The potential for omitted variable bias is serious: we do not use faculty quality or salary information, nor do we use information on additional funding available to public schools from state budgets. One could argue that we are missing a "total assets" measure for both public and private colleges, since the latter tend to have larger endowments as well. This limitation should be offset somewhat by our instructional expenditures per student measure, which captures two things at once: it proxies for overall college resources—the numerator should be highly correlated with this; and it captures institutional policy, since it is an amount allocated per student. Non-linear returns to scale effects may partially drive the potential efficiency of public colleges, since these schools tend to be larger.

Limited resource and allocation measures raise an additional concern over the practicality of some of the hypotheticals we are evaluating. Equalizing tuition across public and private sectors, which implies dramatic increases in public school tuition is unrealistic. So the tuition portion of the returns effect is severely constrained, practically, and this argument extends to other factors, such as SAT scores. Public schools may institute policies that make them more selective, but then where will less academically prepared students enroll?

In terms of our other covariates, we would like to have included more direct measures of academic attachment, such as first year GPA, but these are more appropriate to a multi-level and/or multi-year analysis, such as DesJardins et al. (2002). Social attachment, as proxied by measures of traditionality is easier to measure institutionally, as percent commuter provides substantial indirect information, but individual-level questionnaires could certainly yield other useful indicators. We would like to have examined social attachment variables such as working while in school, family obligations, transfer prevalence, and even stop out incidence at the institutional level. Stop-outs are negatively correlated with graduation rate, so their inclusion as covariates is problematic unless an event history modeling framework with enrollment-based time lines is employed (see Scott and Kennedy, 2005, for some discussion). Again, interest in the effects of certain covariates necessitates taking an individual-level approach to the analysis.

While we express some important concerns here, we emphasize that this analysis is one of a very small number that attempt to make sense of institutional college graduation rates. While some research questions may be best approached using individual-level data, nationally representative public data, such as the Beginning Post-Secondary Students (BPS) or the National Educational Longitudinal Study (NELS), cannot be used to assess specific institutional performance, because only a subset of colleges are available in the sample. And institutional performance, while not highlighted here, is the focus of increased national assessment. Our regression models yield residuals for each college that measure the extent of over- or under-performance, relative to schools with similar institutional and student characteristics; these may be of interest to a variety of audiences.

Our findings strongly suggest that evaluation of public colleges based on raw graduation rates is inappropriate. An adjusted rate such as that proposed by Astin (1997), Astin and Osaguera (2002), or one based on our modeling, takes into account the diverse financial and student inputs to colleges. Contrary to what has been discussed in the literature and what is commonly supposed, we have provided evidence via the Oaxaca decomposition that public colleges are doing a relatively good job when one considers all of the constraints they face.

Our study, therefore, adds to the small but growing literature on college institutional performance. In particular, we take a promising step towards analyzing the relative effectiveness of public and private institutions, presenting a more optimistic portrayal of public sector performance than is generally assumed. In doing this we have focused attention on how graduation rates in the two sectors are related in different ways to various institutional characteristics. This raises interesting new questions about what explains these differences in the functioning of public and private institutions of higher education.

ACKNOWLEDGMENTS

The authors wish to thank the editor and two anonymous reviewers for their comments and suggestions. We additionally thank Linda Edwards, Amy Schwartz, and David Weiman for their suggestions and encouragement on prior drafts of this manuscript.

REFERENCES

Alfonso, M., Bailey, T., and Scott, M. (2005). The educational outcomes of occupational sub-baccalaureate students: Evidence from the 1990s. *Economics of Education Review* 24(2): 197–212.

Allison, P. D. (2001). *Missing Data*, Sage, Thousand Oaks, CA.

Astin, A. W., and Oseguera, L. (2002). *Degree Attainment Rates at American Colleges and Universities*, UCLA Graduate School of Education Higher Education Research Inst, Los Angeles, CA.

Astin, A. W. (1997). How "good" is your institution's retention rate? *Research in Higher Education* 38(6): 647–658.

Bean, J. P., and Metzger, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research* 55(4): 485–540.

Bolt, J. R., and Roberts, K. J. (1997). What are performance indicators and can they be used to manage colleges? A comparative study of similar colleges in the US and the UK. Orlando, FL: Association for Institutional Research Annual Forum Paper.

Braxton, J. M., Hirschy, A. S., and McClendon, S. A. (2003). *Understanding and Reducing College Student Departure. ASHE-ERIC Higher Education Report: Volume 30, Number 3*, Jossey-Bass, San Francisco.

Carbone, J. C., and Winston, G. C. (2004). Saving, wealth, performance, and revenues in U.S. colleges and universities. *The Review of Higher Education* 28(1): 97–128.

The College Board. (2001). SAT I-ACT Score Comparison Table (electronic datafile); SAT I and SAT II Equivalence Tables (obtained directly from the College Board).

The College Board. (1999). American Survey of Colleges (electronic datafile).

DesJardins, S. L., McCall, B. P., Ahlburg, D. A., and Moye, M. J. (2002). Adding a timing light to the "tool box". *Research in Higher Education* 43(1): 83–114.

Dey, E. L., and Astin, A. W. (1993). Statistical alternatives for studying college student retention: a comparative analysis of logit, probit, and linear regression. *Research in Higher Education* 34(5): 569–581.

Greene, W. H. (1993). *Econometric Analysis*, MacMillan, New York.

Hanushek, E. A., Rivkin, S. G., and Taylor, L. L. (1996). Aggregation and the estimated effects of school resources. *Review of Economics and Statistics* 78(4): 611–627.

Horton, N. J., and Lipsitz, S. R. (2001). Multiple imputation in practice: comparison of software packages for regression models with missing variables. *The American Statistician* 55(3): 244–254.

Johnes, J. (1997). Inter-university variations in undergraduate non-completion rates: a statistical analysis by subject of study. *Journal of Applied Statistics* 24(3): 343–361.

Johnes, J., and Taylor, J. (1989). Undergraduate non-completion rates: difference between UK universities. *Higher Education* 18(2): 209–225.

Johnson, J. L. (1997). Commuter college students: what factors determine who will persist and who will drop out? *College Student Journal* 31(3): 323–332.

Lavin, D. E., and Hyllegard, D. (1996). *Changing the Odds: Open Admissions and the Life Chances of the Disadvantaged*, Yale University Press, New Haven.

Little, R. J. A., and Rubin, D. B. (1987). *Statistical Analysis with Missing Data*, John Wiley and Sons, New York.

Martin, O. L. (1990). The college milieu for the 1990s: increasing Black student retention on White campuses. Boston, MA: Paper presented at the Annual Meeting of the American Educational Research Association [ED333091].

McPherson, A., and Patterson, L. (1990). Undergraduate non-completion rates: a comment. *Higher Education* 19(3): 377–383.

Metzger, B. S., and Bean, J. P. (1987). The estimation of a conceptual model of nontraditional undergraduate student attrition. *Research in Higher Education* 27(1): 15–38.

Mortenson, T. G. (2000). Poverty, race, and the failure of public policy: the crisis of access in higher education. *Academe* 86(6): 38–43.

Mortenson, T. G. (1998). Institutional graduation rates by control, academic selectivity and degree level 1983 to 1998. *Postsecondary Education Opportunity* 73: 1–10.

Mortenson, T. G. (1997). Actual versus predicted institutional graduation rates for 1100 Colleges and Universities. *Postsecondary Education Opportunity* 58: 1–24.

Muffo, J. A. (1996). Lessons learned from a decade of assessment. Albuquerque, NM: Association for Institutional Research Annual Forum Paper.

National Center for Education Statistics (1999). *Integrated Postsecondary Education Data System (IPEDS) Fall Enrollment Survey*, US Department of Education, Washington DC.

National Center for Education Statistics (1996). *Beginning Postsecondary Students (BPS): second follow-up Computation by DAS-T Version 3.0*, US Department of Education, Washington, DC.

Pascarella, E. T., and Terenzini, P. T. (1991). *How College Affects Students: Findings and Insights from Twenty Years of Research*, Jossey-Bass, San Francisco.

Petersen, T. (1985). A comment on presenting results from logit and probit models. *American Sociological Review* 50(1): 130–131.

Schafer, J. (1997). *Analysis of Incomplete Multivariate Data*, Chapman and Hall, New York.

Scott, M., and Kennedy, B. (2005). Pitfalls in pathways: some perspectives on competing risks event history analysis in education research. *Journal of Educational and Behavioral Statistics* 30(4).

Smith, J. P., and Naylor, R. A. (2001). Dropping out of university: a statistical analysis of the probability of withdrawal for UK university students. *Journal of the Royal Statistical Society Series A—Statistics in Society* 164(3): 389–405.

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition*, University of Chicago Press, Chicago, IL.

US News and World Report. (2004). America's Best Colleges 2004: Ranking criteria and weights. tttp://www.usnews.com/usnews/edu/college/rankings/about/04rank_brief.php.

Wells Jr., R. N. (1997). The Native American experience in higher education: turning around the cycle of failure II. Research report [ED414108].

Winston, G. C. (1998). A guide to measuring college costs. Williams Project on the Economics of Higher Education Discussion Paper No. 46, January 1998.

Received August 30, 2004.