ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Economics of Education Review ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Economics of Education Review

www.elsevier.com/locate/econedurev

Institutional research productivity and the connection to average student quality and overall reputation

Stephen R. Porter[a,*], Robert K. Toutkoushian[b]

[a]Department of Educational Leadership and Policy Studies, N243 Lagomarcino Hall, Iowa State University, Ames, IA 50011, USA
[b]Department of Educational Leadership and Policy Studies, Education 4220, Indiana University, Bloomington, IN 47405, USA

Received 6 October 2003; accepted 3 June 2005

Abstract

We posit that institutions of higher education attempt to maximize their reputation, and that an institution's reputation, research output, and average student quality are determined simultaneously. Because these outputs are produced jointly, three-stage least squares is used to estimate the parameters of the model. We find that faculty research productivity is positively related to reputation but negatively related to student quality at research universities, but that reputation and student quality have little impact on research productivity at liberal arts colleges.

JEL classification: I23; C31

Keywords: Human capital; Productivity; Input output analysis

1. Introduction

Now more than ever, institutions of higher education (IHE) face pressure from higher education stakeholders to provide evidence of the productivity and efficiency of their operations. As tuition increases continue to outpace the growth in inflation and family incomes, legislators, parents, and students are increasingly asking whether the benefits of a particular institution outweigh the costs of attending, and how the net benefits compare across IHE. The trend toward comparing institutions based on metrics such as the average SAT scores and graduation rates has risen dramatically, fueled in part by the media attention given to institutional ranking schemes devised by commercial entities such as *US News and World Report* (*USNWR*).

Most of the institutional rankings and performance indicator systems in higher education, however, focus on the average quality of students and in the process overlook the research activities and accomplishments of an institution. It is widely acknowledged that the role of faculty and their institutions is to produce teaching, research, and service. Despite this, publications such as *USNWR* tend to concentrate on the entering student characteristics of IHE. This is especially curious given that the vast majority of research on the productivity of faculty members and academic departments

*Corresponding author. Tel.: +1 515 294 7635.
E-mail addresses: srporter@iastate.edu (S.R. Porter), rtoutkou@indiana.edu (R.K. Toutkoushian).

0272-7757/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.econedurev.2005.06.003

has centered around their scholarly accomplishments. The lack of attention given to research at the institutional level is troubling because the resulting assessments and rankings of institutions will overlook an important facet of their mission. Because research activities are omitted from institutional rankings and evaluations, stakeholders will not observe and appreciate the resources and effort required to fulfill this part of IHE missions, nor the benefits accrued to society through university research. Perhaps more troubling is the possibility that if there is a negative relationship between average student quality and an institution's level of research productivity, then performance measures that reward schools for having a high-ability student body may inadvertently work against the research mission of IHE. Evidence on the relationship to date is mixed (Bray, Braxton, & Smart, 1996; Fox, 1992; Grunig, 1997; Porter & Umbach, 2001; Volkwein & Carbone, 1994). While the National Research Council, the Carnegie Foundation, and The Center at the University of Florida have devised rankings that rely on research-related measures of faculty and institutions, they often ignore the multi-product nature of IHE and the effects of other factors on the production of research.

Complicating matters is the fact that the outputs in higher education are produced jointly, and thus interrelated. For example, the quality of students at an institution can affect the production of research, and likewise research activities may influence the quality of instructional services and in turn the types of students who are attracted to the institution. Grunig (1997) found that there was a strong interrelationship between the reputations of an institution's academic programs, and that institutional size and selectivity were common determinants of an institution's undergraduate reputation and the perceived scholarly quality of faculty. Furthermore, researchers have not agreed on a model of IHE behavior, theorizing that institutions attempt to maximize budgets, minimize costs, or maximize prestige/reputation. An institution's reputation is also likely to be interrelated with its research and teaching outcomes, although the extent is not clear. Researchers need to model and examine these interrelationships when trying to gain a better understanding of the research productivity of an institution and derive consistent estimates of how various factors influence research productivity.

In this study, we use a simple model of institutional behavior that describes the interrelationships among selected aspects of its mission, and use this to examine the determinants of institutional research productivity. We posit that institutions attempt to maximize their reputation, and that an institution's reputation, research output, and average student quality are determined simultaneously. Given the multi-product nature of IHE and the fact that their student quality, research output, and reputation are interrelated, we use three-stage least squares to estimate the parameters in our model using data from 1996 for all "national universities" and "national liberal arts colleges" as categorized by *USNWR*.

2. Explaining variations in institutional research productivity

To understand differences in the research productivity of IHE, it is necessary to introduce a model of university behavior that takes into account the multi-product nature of institutions and how they are interrelated. The natural starting place for an investigation of institutional productivity is to ask the question: what are institutions trying to maximize? Most analysts would agree that, unlike firms in the private sector, IHE do not attempt to maximize profits. By definition, public institutions are non-profit in nature, and the fact that many highly-selective private colleges charge tuition rates that are below market-clearing levels suggests that they, too, do not act in ways so as to maximize profits. It has been argued that IHE attempt to maximize revenues, maximize expenditures, or maximize their discretionary budget (see Bowen, 1980; Paulsen, 2001). In this paper, we offer a very simple model of institutional behavior patterned after the work of Dolan and Schmidt (1994), who posited that IHE attempt to maximize their output, and that output, average student quality, and average faculty quality were jointly determined. With regard to specific variables, they measured output by the number of alumni who subsequently earned a doctorate/professional degree, average student quality by the 3rd quartile of SAT scores of freshmen, and average faculty quality by the average salary of associate professors. While their choice of particular measures may be questioned by some, Dolan and Schmidt's (1994) study makes a valuable contribution to the literature in that it demonstrates how to model the interrelationships among outcomes.

As an alternative to Dolan and Schmidt's model, we begin by assuming that IHE seek to maximize their reputation rather than their output, and that an institution derives its reputation through the average research accomplishments of its faculty, the average quality of its students, and other exogenous factors such as the institution's age, public/private status, geographic location, and size. Notationally, this may be expressed as

$$\text{Rep} = f(R_{\text{F}}, Q_{\text{S}}, \text{IC}_{\text{Rep}}), \tag{1}$$

where Rep is the institutional reputation, R_{F} the research output per faculty member at the institution, Q_{S} the average quality of students at the institution, and IC_{Rep} the set of additional institutional characteristics thought to influence reputation. Both R_{F} and Q_{S} can be thought of as "per-capita" measures of research output and student quality, respectively.

The literature on the determinants of departmental reputation has consistently found that a positive relationship exists between the number of publications produced by faculty in a program and the program's reputation (e.g. Ehrenberg & Hurst (1998); Toutkoushian, Dundar, & Becker (1998)). Toutkoushian, Dundar, and Becker (1998) also found that the factors explaining the scholarly quality of faculty were similar to those contributing to the perceived teaching quality of faculty, and questioned whether the reputation measures of scholarly quality and teaching quality were capturing the same concept. An institution's research output per faculty member is in turn posited to be a function of the average quality of faculty (Q_{F}), the average quality of students, and other institutional characteristics (IC_{F}) that would affect per-capita research productivity, such as the size of the institution and the emphasis of research in its mission:

$$R_{\text{F}} = f(Q_{\text{F}}, Q_{\text{S}}, \text{IC}_{\text{F}}). \tag{2}$$

Naturally, there will be some overlap between the exogenous factors in IC_{Rep} and IC_{F}. The average quality of faculty at an institution is not exogenous, however, since the type of faculty employed at an institution will be greatly affected by both supply and demand considerations. The average quality of faculty at an institution is likely to be affected by the institution's reputation, other characteristics of the college or university (IC_{F}) that could affect research productivity and the types of faculty that would be demanded by (supplied to) institutions, and the average human capital of faculty (HC_{F}) as measured by the age, educational attainment, and gender composition of the faculty:

$$Q_{\text{F}} = f(\text{Rep}, \text{IC}_{\text{F}}, \text{HC}_{\text{F}}). \tag{3}$$

Substituting (3) into (2) yields an alternative expression for per-capita research productivity that depends on institutional reputation, average student quality, average faculty human capital, and institutional characteristics that influence faculty supply/demand/productivity:

$$R_{\text{F}} = f(\text{Rep}, Q_{\text{S}}, \text{HC}_{\text{F}}, \text{IC}_{\text{F}}). \tag{3'}$$

Finally, we posit that the average student quality of an institution will be influenced by the institution's reputation, and other characteristics that affect student supply and demand (IC_{S}):

$$Q_{\text{S}} = f(\text{Rep}, R_{\text{F}}, \text{IC}_{\text{S}}). \tag{4}$$

The institutional characteristics would include factors that would possibly affect supply and demand of students. It is not clear a priori whether per-capita research output would have a positive or negative effect on the average quality of students at an institution. On the one hand, higher per-capita research output may send a signal to students that the quality of instruction is likely to be enhanced by the research activities of the faculty. Alternatively, higher per-capita research output could lead to declines in higher-ability students if they feel that research activities will lead faculty to focus less attention on their teaching obligations.

It can readily be seen that Eqs. (1), (3′) and (4) are interrelated in that the dependent variables also appear as explanatory variables in the other equations. Provided that each equation can be identified, three-stage least squares can be used to obtain consistent estimates of the parameters in the model and show the connections between the exogenous variables and the dependent variables, as well as the interrelationships among dependent variables.

3. Data and methodology

We collected a wide range of institution-level data for 4-year IHE identified as being either "national universities" ($n = 203$) or "national liberal arts colleges" ($n = 143$) by *USNWR* in 1996. The national universities offer baccalaureate, master's and doctoral degrees, and many are heavily engaged in research, while national liberal arts colleges stress undergraduate education and award at least 40 percent of their degrees in liberal arts disciplines. Information on institutional characteristics, expenditures by source, revenues by source, enrollments, and faculty are taken from the National Center for Education Statistics' Integrated Postsecondary Education Data System (IPEDS) surveys. Additional institution-level data were obtained from the June 1998 issue of *USNWR*. Finally, we obtained data on the number of faculty publications by institution in 1996 from the Institute for Scientific Inquiry (ISI) databases.

3.1. Dependent variables

Our first dependent variable, institutional reputation (Rep), is measured by the mean academic reputation score from the *USNWR* annual survey of campus presidents, provosts, and deans of admission. Each respondent was asked to place schools in their category into one of four quartiles, with $1 =$ lowest quartile and $4 =$ highest quartile, based on their perception of the quality of the academic programs at each institution. Because respondents were asked to rate schools only within their institutional category, such as national universities or national liberal arts colleges, different processes may have driven how respondents rated institutions across categories, making it necessary to estimate separate three-stage least-squares models for national universities and national liberal arts colleges.

The second dependent variable is the per-capita research output of IHE (R_F), measured as the log of the ratio of institutional publications to the number of full-time faculty. Most studies have used either individual- or departmental-level data to understand faculty research productivity, with only a few studies modeling research productivity at the institutional-level. However, many of the studies at the individual- or departmental-level have also included institutional variables in an attempt to control for organizational differences that may affect the production of research. The most commonly-used measure of research productivity is the number of faculty publications (Baird, 1991; Bell & Seater, 1978; Creamer, 1999; Dundar & Lewis, 1998; Gerrity & McKenzie, 1978; Lawrence & Blackburn, 1988; Porter & Umbach, 2001; Tien & Blackburn, 1996), while other approaches have been used (e.g. Buchmueller, Dominitz, & Hansen, 1999; Gander, 1999; Porter & Umbach, 2001).

Our publication data are based on counts of articles published in academic journals monitored by the ISI for the 1996 calendar year. The Institute is responsible for producing the Science Citation Index, Social Science Citation Index, and the Arts and Humanities Index. While the three indexes do not include all academic journals in each field, together they include over 6600 scholarly journals in over 200 academic disciplines. The process for developing counts for each institution is similar to the process used by Toutkoushian, Porter, Danielson, and Hollis (2003). This measure should not be confused with the average publications per faculty because not all publications that are attributed to an institution were written by faculty members, papers co-authored by faculty at the same institution are only counted once, and the publication counts shown here include only articles published in academic journals monitored by ISI.

The final dependent variable in our study is the average student quality variable (Q_S). This is represented by the midpoint of the 25th and 75th percentiles of SAT scores for incoming freshmen. This approach is similar to that used by Dolan and Schmidt (1994), who used the third-quartile SAT scores of freshmen. The scores for schools using the ACT rather than the SAT were converted to their SAT equivalents using the College Board conversion chart.

3.2. Independent variables

Research on the factors that influence an institution's reputation is limited. Studies conducted on the *USNWR* reputation scores, for example, are usually limited to correlations between the variables used to derive the overall rankings (e.g. Schmitz, 1993; Webster, 2001). However, this research is illustrative in that it has shown that acceptance rates, the average SAT scores of incoming freshmen, and the percentage of freshmen in the top 10 percent of their high school class are highly correlated with reputation (Schmitz, 1993). To the extent that these variables serve as proxies for average student quality, they indicate that the quality of undergraduates is correlated with an institution's reputation.

To explain variations in research productivity across individuals, researchers usually rely on human capital theory (Becker, 1993). Proxy measures of a faculty member's human capital, including educational attainment (Porter & Umbach, 2001; Wanner, Lewis, & Gregorio, 1981), prestige of Ph.D.-granting department (Buchmueller et al., 1999; Long & McGinnis, 1981; Singell & Lillydahl, 1996), and age/experience (Lawrence & Blackburn, 1988) have been found to affect faculty research productivity. Differences in research output by gender is also a common finding, even after controlling for field and other differences between men and women (see, for example, Creamer, 1999). Other analysts also include organizational characteristics as explanatory variables in studies of productivity at the individual level, such as the average productivity of the individual's academic discipline to recognize differences in publication practices across fields (Porter & Umbach, 2001), the departmental reward structure (Fox, 1992; Wanner et al., 1981), and institutional type or mission (Singell & Lillydahl, 1996). Studies of departmental productivity also use the human capital and organizational frameworks to explain variations in output across departments, since an academic program's productivity and reputation are derived by summing the productivity and reputation of its faculty. Generally these studies use averages of individual-level data, such as the percentage of faculty who are tenured (Dolan & Schmidt, 1994), the percentage of female faculty (Gander, 1999), the percentage of faculty with a Ph.D., or the average salaries of faculty by field (Graves, Marchand, & Thompson, 1982).

Much of the literature at both the individual-level and departmental-level has tried to measure an institution's emphasis on research and teaching to help explain variations in research productivity. One set of variables measures the presence and/or size of an institution's graduate programs. Fox (1992), for example, separated her regression models by the degree-granting level of department, while Graves et al. (1982) included a variable indicating the presence of a Ph.D. program. Others have used the ratio of graduate students to faculty (Dundar & Lewis, 1998; Graves et al., 1982) and Carnegie classifications, which in turn are based in part on presence and scope of graduate programs (Gander, 1999; Milem, Berger, & Dey, 2000), to help explain variations in departmental-level productivity. Organizational size also appears to play a role in research productivity, due to possible economies of scale (Dundar & Lewis, 1998; Jordan, Meador, & Walters, 1988). Little attention has been paid to the role of finances and other resources in organizational productivity. Graves et al. (1982) included average faculty salary as a predictor of departmental productivity, and some scholars have distinguished between public and private institutions to control for institutional aspects of finances (e.g. Dundar & Lewis, 1998; Jordan et al., 1988).

While there is a substantial literature at the individual-level, few studies have compared IHE on the basis of either objective or subjective institutional-level data. Hughes (1925) first surveyed scholars in specific fields and used the results to rate 38 institutions on the basis of their aggregate ratings. Similarly, Webster (1986) and Keith (1999) also used subjective (survey) data to compare institutions. With regard to objective data, scholars have used average faculty salaries (Dolan & Schmidt, 1994) or research expenditures (Gander, 1999) to represent average faculty quality. One approach used to measure institution-level research productivity has been to sum responses from individual faculty surveys (Bentley & Blackburn, 1990) or departmental-level surveys (deGroot, McMahon, & Volkwein, 1991). More recently, others have used publication counts derived from the Institute for Scientific Information bibliographic databases (Toutkoushian et al., 2003; Zheng & Stewart, 2002).

In the same way that departmental-level productivity could be viewed as the aggregate of the productivity of individuals within the department, an institution's productivity is the sum of the productivity of its departments and hence individuals. Accordingly, studies attempting to explain variations in institutional-level productivity would rely on the same human capital, departmental, and institutional factors used in studies of individual and department productivity, albeit at possibly different levels of aggregation. With regard to human capital, for example, the variables used have included the percentage of faculty with a Ph.D, the percentage of faculty tenured or with senior rank, and the percent of faculty who are female (Dolan & Schmidt, 1994; Gander, 1999; Milem et al., 2000). Turning to attributes of the institution, the most commonly-used variables have been measures of institutional mission, often defined by the Carnegie classification of the institution (Bentley & Blackburn, 1990; Gander, 1999; Milem et al., 2000). Other organizational attributes used as proxies for the institution's emphasis on research and teaching include the presence of a medical school (Gander, 1999), expenditures per student, and the student-to-faculty ratio (Dolan & Schmidt, 1994).

In their review of the literature on student college choice, Hossler, Braxton, and Coopersmith (1989) review the institutional factors that affect the decision to attend a particular college and their relative impacts. Not surprisingly, the perceived academic quality of the institution has been found to have a large effect on student demand, in that students tend to select institutions where the quality of students is similar to their own (Venti & Wise, 1982). Cook and Frank (1993) provided evidence that there is a direct connection between the student quality and the reputation of IHE. While the evidence also suggests that there is a negative relationship between cost and demand, this relationship is generally weak (see McPherson & Shapiro, 1998; Wetzel, O'Toole, & Peterson, 1998). There is also evidence that successful sports teams can positively affect both the number of applicants and the quality of the incoming freshman class (Murphy & Trandel, 1994; Tucker & Amato, 1993).

In this study, we collected information on a number of variables that relate to the characteristics of an institution. The age of an institution was obtained from *USNWR*, and represents the years since the institution was founded as of 1996. Using data from the IPEDS Institutional Characteristics survey, we created dummy variables for the public/private status of each institution, whether the institution was affiliated with a hospital and/or medical school, whether the institution was a historically black college or university (HBCU), whether the institution was located in New England, and whether the institution was classified as a Research I or II institution according to the 1994 Carnegie classification system. The IPEDS Fall Enrollment survey for 1996 was used to estimate the percentage of students who were graduate students. From the 1996 IPEDS Completions survey, we obtained data on the numbers of bachelor degrees awarded by field for each institution, and then calculated the percentage of these degrees awarded in the sciences. The 1996 IPEDS Finance survey was used to derive measures of expenditures per student and the percentage of expenditures devoted to instruction.[1] From the 1996 IPEDS Faculty Salary survey, we calculated the number of full-time faculty and the percentage of faculty who were female. We relied on the 1994 IPEDS Academic Library survey to compute the number of volumes in institutional libraries. Finally, we obtained data from the NCAA on whether the school had a football or basketball team ranked in the top 10 in the AP poll at the conclusion of the 1993 or 1994 seasons as a measure of athletic team success.

We posit that the institutional characteristics affecting reputation (IC_{rep}) would include the age of the institution, public/private status, whether the institution was a HBCU or was affiliated with a hospital or medical school, whether the institution was located in New England, and the size of the institution, as measured by the number of faculty. We include age of the institution as a predictor of reputation, hypothesizing that older schools are viewed as more prestigious than younger schools. Similarly, because many prestigious institutions are located in New England, there may also be a regional effect of reputation. We include the sports team measure in this equation, as having successful teams may enhance or perhaps even detract from institutional prestige. Because different emphases on research and instruction may affect reputation, we include the size of the graduate program (as measured by the percentage of graduate students) and the percentage of degrees awarded in the sciences. The latter variable distinguishes between schools such as Johns Hopkins and Caltech, which focus on the sciences, and schools such as Harvard and Princeton, which offer a full range of academic programs. Finally, as a control for size of institution, the number of faculty and the number of faculty squared are also used in the reputation equation.

Turning to the institutional characteristics that might influence the per-capita research output of IHE, we control for whether the institution was

[1] Instructional spending is defined by IPEDS as follows: "Enter all instruction expenses of the colleges, schools, departments, and other instructional divisions of the institution and expenses for departmental research and public service that are not separately budgeted. The instruction category includes general academic instruction, occupational and vocational instruction, special session instruction, community education, preparatory and adult basic education, and remedial and tutorial instruction conducted by the teaching faculty for the institution's students. (FARM para. 452.11). Include expenses for both credit and non-credit activities. Exclude expenses for academic administration if the primary function is administration (e.g. academic deans). Such expenses should be entered on line 04." See http://nces.ed.gov/ipeds/web2000/SpringDataItems.asp.

classified by the Carnegie Foundation as being a Research I or II institution, the public/private status, whether the institution was affiliated with a hospital or medical school, whether the school was a HBCU, the percentage of enrollments at the graduate level, and the number of faculty and the number of faculty squared. Whether a school is public is used as a control variable to test whether public and private institutions differ in productivity, perhaps due to the service demands on faculty in the public sector. We also include the dummy variable measuring presence of a hospital or medical school, as institutions with these facilities are likely to have higher ratios of publications to faculty. Whether the institution was a Carnegie Research I or II institution, and the percentage of the student body who were graduate students both measure an institution's emphasis on research, and may contribute positively to an institution's per-capita research output. The relative mix of academic programs at an institution may also affect per-capita research output. Publication practices vary considerably by field, with faculty in the physical sciences on average generating more publications in scholarly journals than many of their counterparts in the humanities and liberal arts. To control for this possibility, we constructed a variable based on the percentage of bachelor degrees awarded in the sciences at each institution (the definition of sciences that we used is the same as that used by ISI to classify publications). We included two variables—the total number of books and journals in the library and the expenditures per student—to capture the effects of resources on the production of research. The percentage of female faculty is an additional variable often used to control for differences in the human capital of faculty, and is expected to be negatively correlated with per-capita research productivity.

Finally, the institutional characteristics entered in the student quality equation (IC_S) include public/private status, whether the institution was affiliated with a hospital or medical school, whether the school was a HBCU, the Carnegie classification, the average fellowships per student, the percentage of expenditures allocated to instruction, and the faculty size variables. As the student demand literature demonstrates that cost is a factor in the decision to attend college, the average fellowships offered by the institution capture some differences in cost between schools. The percentage of the annual budget spent on instruction is a proxy for the institution's emphasis towards teaching, and the percentage of degrees awarded in the sciences measures differences in academic focus between institutions. We also include the presence of a top 10 sports team, as some students may be more attracted to schools with nationally recognized sports teams. Finally, we control for the size of the institution by including the number of faculty at the institution, as well as a squared term to capture any nonlinearities. Table 1 provides the descriptive statistics for both datasets.

Clearly, all three equations are overidentified due to the exclusion restrictions that we impose on the model. Some of the faculty human capital measures HC_F can be safely excluded from the student quality and institutional reputation models. Likewise, factors such as institutional age, geographic location, and per-student financial aid expenditures in and of themselves are unlikely to impact the research productivity of an institution. As a further test of the appropriateness of our restrictions, we applied Basmann's (1960) test for overidentifying restrictions to each of the models presented below, and failed to reject (at $p < .05$) the null hypothesis that predetermined variables excluded from the models have zero coefficients.

4. Results

Our first goal is to obtain estimates of the structural parameters in our system of equations. We use 3SLS to do this because OLS will be biased and inefficient given the simultaneity of equations (1), (3′) and (4). The estimated structural parameters are shown in Tables 2 and 3. These can be thought of as the "direct effect" of each X on each Y of interest. For the purpose of predicting the total impact of a change in X on a given Y, we have to take into account the fact that each X may impact multiple Y's. This is important for evaluating the impacts of changes in policy-related variables on outcomes, because changing the number of faculty, for example, will have a direct effect on the ratio of publications to faculty as well as an indirect effect through the resulting change in institutional reputation and average student quality. In other words, we can think of the reduced form parameters as "the long-run multipliers associated with the model" (Kennedy, 2003, p. 181). Accordingly, the total impacts are estimated using the reduced-form parameters from the model.

Table 1
Variable descriptive statistics

Variable	National universities				National liberal arts colleges			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Reputation score	2.4	0.7	1.3	4.0	2.4	0.7	1.3	3.9
Average publications per faculty (logged)	0.67	0.44	0.00	2.16	0.14	0.10	0.00	0.48
SAT score	1134.2	125.9	885.0	1505.0	1189.9	90.5	910.0	1415.0
Age of institution	119.1	52.9	27.0	360.0	139.6	41.5	27.0	254.0
Public institution	0.66	0.48	0.00	1.00	0.04	0.20	0.00	1.00
Hospital or medical school	0.42	0.49	0.00	1.00	0.00	0.00	0.00	0.00
HBCU	0.01	0.10	0.00	1.00	0.01	0.08	0.00	1.00
New England region	0.09	0.29	0.00	1.00	0.10	0.31	0.00	1.00
Top 10 football or basketball team	0.13	0.34	0.00	1.00	0.00	0.00	0.00	0.00
Carnegie Research I or II	0.56	0.50	0.00	1.00	0.00	0.00	0.00	0.00
Number of faculty (1000's)	0.71	0.42	0.06	2.30	0.11	0.05	0.04	0.26
Number of faculty squared (1000's)	685.86	852.12	3.25	5271.62	15.04	13.16	1.52	69.70
% degrees awarded in sciences	0.33	0.18	0.00	0.98	0.22	0.07	0.05	0.47
% graduate students	0.31	0.14	0.09	0.80	0.05	0.11	0.00	0.60
Books and journals (100,000's)	18.83	17.85	1.02	131.43	3.04	2.15	0.00	11.75
Average fellowships ($1000's)	2.08	1.77	0.33	8.86	5.27	1.87	0.41	10.01
% spending on instruction	0.43	0.09	0.22	0.65	0.42	0.05	0.29	0.55
Expenditures per student ($1000's)	21.19	16.18	6.36	132.05	16.29	5.12	5.47	31.91
% female faculty	0.26	0.07	0.08	0.75	0.34	0.10	0.09	0.62

Table 2 presents the estimated coefficients for Eqs. (1), (3′) and (4) for the national universities, and Table 3 presents the estimated coefficients for the same equations when applied to the national liberal arts colleges. The left half of each table shows the 3SLS estimates, and the right half contains the 2SLS results for comparison.[2] Beginning with the national universities (see Table 2), we can see that the 3SLS models explain 88% of the variations in institutional reputation, and over 70% of the variations in average student quality and the ratio of publications to faculty. The estimates for the reputation equation show that both research productivity and SAT scores have positive effects on the reputation of IHE. A 100-point increase in average SAT scores would lead to an increase in reputation of .4, or almost one-half point on the 4-point reputation scale, while a 10% increase in average publications would yield an increase of only .04 points. Of the institutional characteristic variables, however, we found that only two had a significant effect on reputation. First, the percentage of degrees awarded in the sciences is negatively related to reputation, indicating that schools that specialize heavily in the sciences pay a price in terms of reputation. Second, reputation was found to be positively related to the number of faculty at the institution. For every 100 person increase in the size of the faculty, reputation score increases by .05 points. Thus, the determinants of reputation for national universities are relatively simple: the research productivity and size of the faculty, the academic quality of an institution's student body, and the instructional emphasis of the school.

Turning to the research productivity equation, we can see that institutions with better reputations have on average more publications per faculty member; a one-point increase in reputation score is associated

[2]The first-stage results for the models in Table 3 and 4are summarized as follows. The R^2 range from .75 to .79 for the universities and .46 to .83 for the colleges, and indicate that the instruments are strongly correlated with the endogenous regressors. In addition, for the universities the F statistics are in 30–38 range, above the cutoff value of 10 recommended by Staiger and Stock (1997, p. 557). For the colleges, one equation (publications) has an F statistic that is near the cutoff value of 10, but is not less. In general, the first-stage regressions indicate that there is a reasonable degree of correlation with the endogenous variables.

We also tested for heteroskedasticity in the first-stage regressions, with a significant result ($p<.05$) only for the reputation equation for the universities. We estimated robust standard errors for the 2nd stage equation for this model and compared them to the regular standard errors. The standard errors are very similar, with only one change in the results (age no longer significant at $p<.10$). We conclude from these results that heteroskedasticity is not an issue for these two sets of cross-sectional data.

Table 2
Coefficient estimates for models explaining institutional reputation, ratio of publications to faculty, and average student quality in 1996: National universities

	Three-stage least squares			Two-stage least squares		
	Reputation score	Ratio of pubs. to faculty	Average student quality	Reputation score	Ratio of pubs. to faculty	Average student quality
Intercept	−2.361**	1.255	623.958**	−1.258	0.937	655.783**
	(0.482)	(1.227)	(59.421)	(0.829)	(1.277)	(65.142)
Reputation score		0.587 +	274.776**		0.329	242.459**
		(0.315)	(36.753)		(0.336)	(39.449)
SAT score	0.004**	−0.002		0.002**	−0.001	
	(0.000)	(0.001)		(0.001)	(0.002)	
Average pubs. per faculty (logged)	0.448**		−131.266 *	0.677**		−106.057
	(0.129)		(56.666)	(0.198)		(69.924)
Age of institution	−0.000			0.001 +		
	(0.000)			(0.001)		
Public institution	0.088	0.017	−22.670	0.006	0.028	−29.289
	(0.072)	(0.096)	(17.581)	(0.099)	(0.099)	(21.590)
Hospital or medical school	−0.092+	0.146**	26.846 +	−0.172**	0.138**	18.282
	(0.049)	(0.043)	(14.531)	(0.065)	(0.044)	(15.223)
HBCU	0.046	−0.351*	−14.747	0.060	−0.384 *	−21.322
	(0.174)	(0.178)	(45.053)	(0.174)	(0.182)	(48.750)
New England region	0.002			0.013		
	(0.026)			(0.062)		
Top 10 sports team	0.067		−20.496	0.119*		−27.149 +
	(0.050)		(13.903)	(0.054)		(14.660)
Carnegie Research I or II	0.092	0.127	−23.405	0.062	0.176 +	−18.774
	(0.061)	(0.090)	(16.056)	(0.066)	(0.095)	(19.097)
% degrees awarded in sciences	−0.425**	0.392**	114.902**	−0.317*	0.272 +	112.370**
	(0.113)	(0.135)	(27.886)	(0.125)	(0.141)	(28.593)
% graduate students	−0.016	−0.252 +		0.202	−0.343*	
	(0.088)	(0.145)		(0.172)	(0.165)	
Books in library (100,000's)		0.005*			0.006*	
		(0.002)			(0.002)	
Average fellowships ($1000's)			−0.604			−0.524
			(3.042)			(6.419)
% spending on instruction			5.104			29.296
			(24.888)			(56.673)
Expenditures per student		0.010**	0.128		0.012**	0.578
($1000's)		(0.002)	(0.391)		(0.002)	(0.818)
Number of faculty (1000's)	0.518**	−0.475*	−141.778**	0.523**	−0.351	−116.220*
	(0.173)	(0.209)	(50.623)	(0.175)	(0.216)	(52.291)
Number of faculty squared	−0.038	0.015	10.020	−0.029	0.003	8.624
(1000's)	(0.074)	(0.071)	(18.825)	(0.075)	(0.072)	(19.090)
% female faculty		−0.053			−0.295	
		(0.358)			(0.401)	
Adjusted R^2	0.88	0.74	0.75	0.87	0.75	0.77
N	203			203		

Note: Standard errors are shown in parentheses; $p<.01$ **, $p<.05$ *, $p<.10$+.

with an 80% increase in publications per faculty member. Conversely, as the quality of the student body increases, research productivity decreases, although this result is not statistically significant. The reputation result is not surprising, as respected institutions should have an easier time attracting top faculty. To the extent that students with higher SAT scores may demand more and better quality

Table 3
Coefficient estimates for models explaining institutional reputation, ratio of publications to faculty, and average student quality in 1996: National liberal arts colleges

	Three-stage least squares			Two-stage least squares		
	Reputation score	Ratio of pubs. to faculty	Average student quality	Reputation score	Ratio of pubs. to faculty	Average student quality
Intercept	−4.873**	0.991	870.513**	−4.568**	0.214	902.055**
	(0.992)	(0.603)	(36.318)	(1.088)	(0.763)	(42.725)
Reputation score		0.196	100.752*		0.075	125.058*
		(0.124)	(45.867)		(0.172)	(49.791)
SAT score	0.006**	−0.001 +		0.005**	−0.000	
	(0.001)	(0.001)		(0.001)	(0.001)	
Average pubs. per faculty (logged)	3.029**		−45.696	3.545**		19.167
	(1.119)		(205.006)	(1.252)		(207.158)
Age of institution	−0.000			0.000		
	(0.000)			(0.001)		
Public institution	−0.268 +	0.071*	−5.488	−0.265 +	0.056	−12.850
	(0.145)	(0.034)	(26.356)	(0.154)	(0.037)	(26.872)
HBCU	0.874**	−0.170	−134.275**	0.861**	−0.035	−148.636**
	(0.320)	(0.122)	(45.908)	(0.324)	(0.152)	(46.697)
New England region	0.020			−0.095		
	(0.039)			(0.104)		
% degrees awarded in sciences	0.307	−0.121	14.180	0.233	−0.122	25.123
	(0.424)	(0.087)	(63.437)	(0.438)	(0.088)	(63.715)
% graduate students	−0.428 +	0.138 +		−0.591*	0.063	
	(0.220)	(0.073)		(0.239)	(0.089)	
Books and journals (100,000's)		0.002			0.011	
		(0.005)			(0.007)	
Average fellowships ($1000's)			−3.854 +			−5.984 *
			(2.251)			(2.600)
% spending on instruction			100.905			53.552
			(83.912)			(100.894)
Expenditures per student		0.004	2.702		0.004	−0.166
($1000's)		(0.006)	(3.593)		(0.008)	(4.216)
Number of faculty (1000's)	−0.766	0.533	314.736	−1.030	0.524	54.253
	(2.541)	(0.788)	(462.767)	(2.573)	(1.002)	(500.947)
Number of faculty squared	1.569	−0.977	−1194.552	2.602	−1.349	−747.577
(1000's)	(9.064)	(2.460)	(1452.698)	(9.210)	(2.861)	(1510.116)
% female faculty		−0.103*			−0.070	
		(0.049)			(0.085)	
Adjusted R^2	0.82	0.41	0.79	0.80	0.50	0.74
N	143			143		

Note: Standard errors are shown in parentheses; $p<.01$ **, $p<.05$ *, $p<.10$ +.

teaching, the student quality results seen here suggest a possible negative relationship between research and teaching.

An institution with a hospital or medical school has on average more publications per faculty member, and HBCUs have lower publications per faculty member. Not surprisingly, schools oriented towards the sciences have higher levels of faculty productivity. Interestingly, schools with larger proportions of graduate students in the student body have fewer publications per faculty member. Given that these schools are all major research universities, this variable may be measuring differences in graduate teaching and mentoring workload. Two of the resource variables, number of volumes in the library and expenditures per student, are positively correlated with publications per faculty member. Surprisingly, the human capital variables were not statistically significant, with the exception of faculty size. Although research at the departmental level indicates returns to scale for faculty productivity (e.g., Dundar & Lewis, 1998; Jordan et al., 1988), the results here indicate the opposite. In part this may be due to faculty preferring smaller universities over larger ones. And contrary to much of the research on faculty research productivity at the individual level, the proportion of female faculty does not appear to significantly affect institution-level research productivity, although the coefficient for this variable is negative.

Finally, the average student quality equation indicates that higher reputation contributes to a higher quality student body, and that faculty research productivity has a negative impact on student quality. A one-point increase in the reputation score yields a 275-point increase in average SAT scores, a large increase considering the IHE averages range between 885 and 1505 in this sample. Students with high SAT scores also appear to prefer schools where the faculty are not as productive in terms of research. Of the remaining variables, higher-quality students appear to be attracted to science-oriented schools and smaller IHE. Interestingly, we found no relationship between average student quality and athletic team success, implying that institutions with successful teams do not have more success recruiting high-quality students, contrary to findings by Murphy and Trandel (1994) and Tucker and Amato (1993).

The results in Table 3 for the national liberal arts colleges are quite interesting in that the models accounted for about 80% of the variations in reputation and average student quality, but only about 41% of the variation in the ratio of publications to faculty. In addition, fewer variables were found to be statistically significant compared to the results for the national universities. For liberal arts colleges, average student quality and faculty research productivity appear to be the driving force behind reputation, with HBCUs having higher reputations after controlling for other institutional characteristics. Note that the impact of publications per capita on reputation is much larger for a liberal arts college than for a university. In part this may be due to the fewer publications per capita typically produced by liberal arts college faculty; thus when a faculty member publishes, it has a larger marginal impact on reputation.

Per-capita research productivity is affected by several variables. SAT scores negatively affect research productivity, with a 100-point increase in average SAT scores leading to an 11% decrease in publications per faculty member. Clearly for liberal arts colleges there is a tradeoff between teaching and research. Public liberal arts colleges have higher per-capita research productivity. Interestingly, the effect of graduate students for liberal arts colleges is positive and the opposite of the universities. Finally, we see that a 10% increase in the proportion of female faculty results in a 3% decrease in research productivity, similar to other research (Creamer, 1999; Gander, 1999). The poor explanatory power of this equation may indicate the difficulty in explaining faculty productivity at liberal arts colleges using traditional models. Almost all of the research on faculty research productivity has focused on research universities, and the process at liberal arts colleges is not well understood by analysts. Finally, average student quality is only affected by the institution's reputation and whether the institution is a HBCU.

To illustrate some of the long-run tradeoffs that universities face when attempting to maximize all three outputs in our model, we used the reduced form estimates for the 3SLS model to calculate the overall impact of a one standard deviation change in selected independent variables on the three output variables. We found that increasing expenditures by $16,000 per student leads to a modest increase in reputation (1/10 of a point on the four-point scale), a 20% increase in publications per faculty member, and a slight drop in average SAT scores. More spending leads to increased productivity, but because quality students appear to prefer institutions with lower faculty productivity, student quality declines. Increasing the number of faculty by 420 has no effect on institutional reputation, as the gains are wiped out by the negative effect of faculty size on productivity and student quality. Increasing an institution's proportion of science degrees awarded by 18% points yields a small increase in reputation, even though the direct effect of science degrees on reputation is negative. Productivity also slightly increases, with a large gain in average SAT scores. Together, these calculations illustrate the complexity faced by institutions seeking to change any of the three outputs.

In sum, after taking into account the endogeneity of reputation, per-capita research productivity and average student quality, we find that reputation has a positive impact on per-capita research productivity for universities, while student quality appears to have a negative impact for liberal arts colleges. In addition, productivity at liberal arts colleges is not well explained. Finally, some of the human capital variables that affect productivity at the individual level in previous research do not have an impact at the institutional level.

5. Summary and discussion

Although faculty research productivity is an important component of an IHE's mission, relatively little research has been conducted on how average student quality and other institutional characteristics of an IHE affect faculty research productivity. This paper is an attempt to further understand this issue. Because an IHE's outputs are interrelated and simultaneously determined, we chose a modeling approach that takes into account the relationships between an institution's reputation, faculty research productivity and quality of its undergraduate student body. This approach yields some interesting conclusions that differ from what a standard non-recursive approach would show.

In particular, we found that institutions may face a tradeoff between having a high quality research faculty and a high quality undergraduate student body. Much of the literature on faculty research productivity has focused on the nexus between teaching and research (e.g. Fox, 1992; Neumann, 1996). The 3SLS results indicate that the quality of an institution's student body, as measured by average SAT scores, is significantly—but negatively—related to faculty research productivity. If students with top academic credentials are more demanding than other students of faculty, then our results would indicate that research and teaching are competitive rather than mutually reinforcing activities.

Finally, our paper illustrates the importance of treating universities and liberal arts colleges separately when studying reputation and productivity. While our model appears to explain IHE outputs for national universities, our results for national liberal arts colleges were more limited: only a few variables were statistically significantly correlated with our three outputs. It is possible that these are the only variables driving these outputs. However, it is more likely that these processes are very different for liberal arts colleges, and merit a closer look. There are certainly many caveats and qualifiers that should be noted when attempting a study such as this. First and foremost is that there are many limitations in the data that are available for analysis at the institutional level. For example, it is not clear what is actually being measured in the reputation score obtained from *USNWR*. Likewise, the number of publications attributed to an institution is an output measure that does not reflect the quality of the research being conducted by any institution, except to the extent that the studies passed through the peer review system. Nonetheless, since policymakers and higher education stakeholders routinely point to available metrics and use this information for decision-making and policy formation, it is crucial that the debate around measuring performance be expanded to include other major facets of the goals and missions of higher education institutions.

Acknowledgements

Earlier versions of this paper were presented at the 2002 annual meetings of the Western Economic Association, Seattle, WA, and the Association for Institutional Research, Toronto, Canada. We would like to thank participants at the annual meetings of the Association for Institutional Research as well as Richard Butler for their comments on an earlier draft of this study. We also would like to thank Doris Spiller, Cherry Danielson, and Paula Hollis for their assistance in obtaining and compiling the dataset used in this study.

References

Baird, L. L. (1991). Publication productivity in doctoral research departments: Interdisciplinary and intradisciplinary factors. *Research in Higher Education*, *32*(3), 303–318.

Basmann, R. L. (1960). On finite sample distributions of generalized classical linear identifiability test statistics. *Journal of the American Statistical Association*, *55*(292), 650–659.

Becker, G. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education*. Chicago: The University of Chicago Press.

Bell, J., & Seater, J. (1978). Publishing performance: Departmental and individual. *Economic Inquiry*, *16*(4), 599–615.

Bentley, R., & Blackburn, R. T. (1990). Changes in academic research performance over time: A study of institutional accumulative advantage. *Research in Higher Education*, *31*(4), 327–354.

Bowen, H. (1980). *The costs of higher education*. San Francisco: Jossey-Bass.

Bray, N., Braxton, J., & Smart, J. (1996). Research activity and the support of undergraduate education. In J. Braxton (Ed.), *Faculty teaching and research: Is there a conflict? New Directions for Institutional Research*, *90*, 23–30.

Buchmueller, T. C., Dominitz, J., & Hansen, W. L. (1999). Graduate training and the early career productivity of Ph.D. economists. *Economics of Education Review*, *18*(1), 65–77.

Cook, P., & Frank, R. (1993). The growing concentration of top students at elite schools. In Clotfelter, & Rothschild (Eds.), *Studies of supply and demand in higher education*. Chicago and London: The University of Chicago Press.

Creamer, E. G. (1999). Knowledge production, publication productivity, and intimate academic partnerships. *Journal of Higher Education*, *70*(3), 261–277.

deGroot, H., McMahon, W., & Volkwein, J. F. (1991). The cost structure of American research universities. *The Review of Economics and Statistics*, *73*(3), 424–431.

Dolan, R. C., & Schmidt, R. M. (1994). Modeling institutional production of higher education. *Economics of Education Review*, *13*(3), 197–213.

Dundar, H., & Lewis, D. R. (1998). Determinants of research productivity in higher education. *Research in Higher Education*, *39*(6), 607–631.

Ehrenberg, R., & Hurst, P. (1998). The 1995 NRC ratings of doctoral programs: A hedonic model. *Economics of Education Review*, *172*(2), 137–148.

Fox, M. F. (1992). Research, teaching and publication productivity: Mutuality versus competition in academia. *Sociology of Education*, *65*(4), 293–305.

Gander, J. P. (1999). Faculty gender effects on academic research and teaching. *Research in Higher Education*, *40*(2), 171–184.

Gerrity, D., & McKenzie, R. (1978). The ranking of southern economics departments: New criterion and further evidence. *Southern Economics Journal*, *45*(2), 608–614.

Graves, P., Marchand, J., & Thompson, R. (1982). Economics departmental rankings: Research incentives, constraints, and efficiency. *American Economic Review*, *72*(5), 1131–1141.

Grunig, S. (1997). Research, reputation, and resources: The effect of research activity on perceptions of undergraduate education and institutional resource allocation. *Journal of Higher Education*, *68*(1), 17–52.

Hossler, D., Braxton, J. M., & Coopersmith, G. (1989). Understanding student college choice. In J. C. Smart (Ed.), *Handbook of higher education*. New York: Agathon Press.

Hughes, R. (1925). *A study of the graduate schools of America*. Oxford: Miami University Press.

Jordan, J. M., Meador, M., & Walters, S. J. K. (1988). Effects of department size and organization on the research productivity of academic economists. *Economics of Education Review*, *7*(2), 251–255.

Keith, B. (1999). The institutional context of departmental prestige in American higher education. *American Educational Research Journal*, *363*(3), 405–445.

Kennedy, P. (2003). *A guide to econometrics*. Cambridge: MIT Press.

Lawrence, J. H., & Blackburn, R. T. (1988). Age as a predictor of faculty productivity: Three conceptual approaches. *Journal of Higher Education*, *59*(1), 22–38.

Long, J. S., & McGinnis, R. (1981). Organizational context and scientific productivity. *American Sociological Review*, *46*(4), 422–442.

McPherson, M., & Shapiro, M. (1998). *The student aid game: Meeting need and rewarding talent in American higher education*. Princeton, NJ: Princeton University Press.

Milem, J. F., Berger, J. B., & Dey, E. L. (2000). Faculty time allocation: A study of change over twenty years. *Journal of Higher Education*, *71*(4), 454–475.

Murphy, R. G., & Trandel, G. (1994). The relation between a university's football record and the size of its applicant pool. *Economic of Education Review*, *13*(3), 265–270.

Neumann, R. (1996). Researching the teaching-research nexus: A critical review. *Australian Journal of Education*, *40*(1), 5–18.

Paulsen, M. (2001). Economic perspectives on rising college tuition: A theoretical and empirical exploration. In M. Paulsen, & J. Smart (Eds.), *The finance of higher education: Theory, research, policy and practice* (pp. 193–263). New York: Agathon Press.

Porter, S. R., & Umbach, P. D. (2001). Analyzing faculty workload using multilevel modeling. *Research in Higher Education*, *42*(2), 171–196.

Schmitz, C. C. (1993). Assessing the validity of higher education indicators. *Journal of Higher Education*, *64*(5), 503–521.

Singell, L. D., & Lillydahl, J. H. (1996). Will changing times change the allocation of faculty time? *Journal of Human Resources*, *31*(2), 429–449.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586.

Tien, F. F., & Blackburn, R. T. (1996). Faculty rank system, research motivation, and faculty research productivity. *Journal of Higher Education*, *67*(1), 2–22.

Toutkoushian, R. K., Dundar, H., & Becker, W. E. (1998). The National Research Council graduate program ratings: What are they measuring? *Review of Higher Education*, *21*(4), 427–443.

Toutkoushian, R. K., Porter, S. R., Danielson, C., & Hollis, P. (2003). Using publication counts to measure an institution's research productivity. *Research in Higher Education*, *44*(2), 121–148.

Tucker, I. B., & Amato, L. B. (1993). Does big-time success in football or basketball affect SAT scores? *Economics of Education Review*, *12*(2), 177–181.

Venti, S., & Wise, D. (1982). Test scores, educational opportunities, and individual choice. *Journal of Public Economics*, *18*(1), 35–63.

Volkwein, J. F., & Carbone, D. (1994). The impact of departmental research and teaching climates on undergraduate growth and satisfaction. *Journal of Higher Education*, *65*(2), 147–167.

Wanner, R., Lewis, L., & Gregorio, D. (1981). Research productivity in academia: A comparative study of the sciences, social sciences, and humanities. *Sociology of Education*, *54*(4), 238–253.

Webster, D. (1986). *Academic quality rankings of American colleges and universities*. Springfield, IL: Charles C. Thomas.

Webster, T. J. (2001). A principal component analysis of the U.S. News & World Report tier rankings of colleges and universities. *Economics of Education Review*, *20*(3), 235–244.

Wetzel, J., O'Toole, D., & Peterson, S. (1998). An analysis of student enrollment demand. *Economics of Education Review*, *17*(1), 47–54.

Zheng, H. Y., & Stewart, A. C. (2002). Assessing the performance of public research universities using NSF/NCES data and data envelopment analysis technique. *AIR Professional File*, *83*, 1–2.