## Research Notes

# Faculty Publishing Productivity: An Institutional Analysis and Comparison with Library and Other Measures

John M. Budd

This paper addresses the level of publishing productivity of faculty for the years 1991 through 1993 at institutions with membership in the Association of Research Libraries (ARL). The sources of data are the three citations indexes produced by the Institute for Scientific Information. Both raw and normalized data are presented. In addition, these measures are compared with some library-related information, as well as some other institutional data such as numbers of doctorates awarded. Rank-order correlation is employed to examine relationships between variables. In addition, goodness-of-fit tests are used to test hypotheses regarding the relationship between the publishing data and the other variables.

There is no doubt that faculty at research universities must be concerned with publishing productivity. The literature on the subject of publishing requirements, pressure to publish, and the ties of the academic reward structure to publishing is far too voluminous to trace here. To note just one source that emphasizes the use of quantitative measures of publishing activity, an entire issue of *New Directions for Institutional Research*, entitled "Measuring Faculty Research Performance," contains several essays addressing the use of counts of various sorts in evaluating faculty publishing productivity at universities.[1] The question remains: how much do faculty actually publish? That question forms the basis of the present study. Beyond that, ancillary questions concern the relationship between publishing activity and other institutional variables, many of which are library based.

### Faculty Publishing

The pressures exerted on faculty to publish are recognized by several writers in the library field, Charles Osburn among them. Although his substantive study was published in 1979, many of his observations still apply: faculty are part of a complex research dynamic that is also composed of the academic reward structure; a large, and mainly public, pool of

*John M. Budd is Associate Professor in the School of Library and Informational Science at the University of Missouri at Columbia; e-mail: libsjmb@mizzou1.missouri.edu.*

funding is available to support research; and a multifaceted publishing industry is responsive to the need and desire for expanded outlets for the communication of research. This dynamic places pressure on libraries to both supply the raw information materials to be used in the research process and to serve as a communication medium of the products of faculty research.[2] This further complicates an already complex set of motives underlying the phenomenon of publishing. As Herb White observes:

> The purpose of publication is, after all, a twofold one. The first and the most immediately recognized purpose is the communication of findings, sometimes to an eager audience and sometimes to a disinterested one. The former is preferable, but even the latter is acceptable, because the other purpose of scholarly publication is the achievement of academic credit. Unfortunately . . . , credit depends less on the quality and more on the quantity of activity in today's academic marketplace.[3]

If quantity is so important, then how much are faculty publishing? The period 1991 through 1993 is examined to determine publishing rates by faculty at institutions that are members of ARL. The sources of publishing data are the three indexes produced by the Institute for Scientific Information (ISI): *Science Citation Index*, *Social Sciences Citation Index*, and *Arts & Humanities Citation Index*. It is recognized that these tools are limited, that they cannot account for the totality of publications (particularly nonjournal publications), but they do cover the three broad subject areas and include the contents of approximately 5,700 journals. Furthermore, the ISI databases allow for the searching of the Corporate Index, so that publications emanating from ARL institutions can be identified.

This is not the first examination of institutional publishing patterns, nor is it the first study to employ ISI databases. In 1978 Richard C. Anderson, Francis Narin, and Paul McAllister published a comparison of ratings in ten scientific fields using the Corporate Index of *Science Citation Index*.[4] Nine years later, John A. Muffo, Susan V. Mead, and Alan E. Bayer used the ISI databases to focus on five universities.[5] The authors note that instances of multiple authorship, affiliation with multiple departments, and interinstitutional publication present some problems. Also, only the affiliation of the first author of a multiauthored work is included. They urge that readers view these data not as absolute facts, but as indicators of activity. These caveats and recommendations apply as well to the present study. This research is on a larger scale than previous studies. All ninety-four United States universities that have membership in ARL are included.

The author searched the three databases for the time period in question, using the Corporate Index. This means that only the main campuses of the institutions are included. More significantly, in some instances, a university's medical school is included if it is attached to the main campus. Attempting to eliminate medical schools proved problematic, so they are included when they are part of the main campus. This creates some discrepancy between these universities and those without medical schools or with medical schools in locations apart from the main campus. One reason for this strategy is that ARL library statistics include medical school or health science collections when they are part of the main campus, but not when they are physically separate. Consistency of data collection allows for comparison with the library statistics. Beyond the medical school dilemma, the ARL data are accepted as presented. This may result in some inclusions of more than just the main campuses, but a reconciliation of the ARL data with in-

formation from the ISI databases is very difficult, if not impossible, due to the nature of the reporting mechanism.

At the most basic level, data are gathered on total numbers of publications. *Publication* is defined according to ISI's designation of an item as an "article." This results in the elimination of such works as book reviews, editorials, letters, and notes. Only "articles" are counted as publications in this study. The mean number of publications per institution is 4,595.8 (SD=3,089.9). The range extends from 669 publications at the low end to a high of 16,945. Table 1 presents a ranked list of institutions by number of publications.

It may come as no surprise that Harvard ranks first in total number of publications. The remaining nine institutions in the top ten are also ones that have reputations for prestige. It stands to reason that those universities with the largest faculties may produce the greatest numbers of publications. One way to normalize this measure is to compute per capita publication. The number of faculty for each university is taken from the 1991–1992 *ARL Statistics*.[6] This source provides the head count of faculty for each of the universities. For each institution, the total number of publications is divided by the number of faculty to arrive at a per capita figure. The mean per capita number of publications is 3.56 (SD=2.48). The lowest is 0.50 publications per capita and the highest is 12.71. The universities ranked by this measure are noted in table 2.

These rankings include no surprises either, with the possible exception of the inclusion of the University of California, San Diego. Seven of the top ten universities in table 1 also appear in the top ten in table 2. Even among these top ten, however, there is some separation. There is a gap between the first four institutions and the next six. Using rank-order correlation to examine the relationship of total publications and per capita publications yields a correlation coefficient of .793, which is quite a high positive correlation.

### Publishing Output and Library Measures

While these basic measures provide general institutional comparisons, they also offer an opportunity for comparison with other factors. It is often said that faculty and librarians are in a partnership when it comes to the production of research and scholarship. There is an assumed interdependence between information collections and the services of the university and the faculty, who are both the producers and consumers of that information. The publishing activity, therefore, can be compared with some key library-related variables. These include: total number of volumes held by the institutions' libraries, the libraries' total expenditures, materials expenditures, and the number of professional librarians on their staffs. These aspects of research libraries are chosen because they relate most directly to collections and services that may be of benefit to faculty. In addition to these library-based variables, the publishing data are compared with the number of doctorates produced by the universities in 1992. These data (library statistics and number of doctorates) are derived from the *1991–1992 ARL Statistics*. The final comparison is with one subjective measure—the rating of graduate schools as published in the latest edition of the *Gourman Report*.[7] Comparison with the Gourman ratings is not intended in any way to imply approval of his methods or ratings. In fact, many researchers frequently criticize Gourman for not being forthcoming with information regarding his methods of evaluation and for unclear

**It is often said that faculty and librarians are in a partnership when it comes to the production of research and scholarship.**

**TABLE 1**
**Institutions Ranked by Number of Publications**

| Rank | Institution | Number | Rank | Institution | Number |
|---|---|---|---|---|---|
| 1 | Harvard | 16,945 | 48 | Princeton | 3,803 |
| 2 | UCLA | 12,566 | 49 | Virginia Tech | 3,660 |
| 3 | MIT | 11,788 | 50 | Iowa State | 3,520 |
| 4 | Michigan | 10,907 | 51 | Cincinnati | 3,516 |
| 5 | U. of Washington | 10,645 | 52 | UC Santa Barbara | 3,442 |
| 6 | Cornell | 10,518 | 53 | Missouri | 3,439 |
| 7 | UC Berkeley | 10,378 | 54 | Indiana | 3,408 |
| 8 | Minnesota | 10,304 | 55 | Emory | 3,279 |
| 9 | Stanford | 9,723 | 56 | SUNY, Buffalo | 3,165 |
| 10 | Wisconsin | 9,663 | 57 | Brown | 3,100 |
| 11 | Johns Hopkins | 9,636 | 57 | Georgia | 3,100 |
| 12 | Pennsylvania | 8,636 | 59 | Arizona State | 3,068 |
| 13 | Illinois | 7,884 | 60 | Wayne State | 3,020 |
| 14 | Columbia | 7,824 | 61 | Massachusetts | 3,004 |
| 15 | Yale | 7,779 | 62 | Louisiana State | 2,986 |
| 16 | UC San Diego | 7,732 | 63 | Kansas | 2,974 |
| 17 | UC Davis | 7,621 | 64 | Kentucky | 2,953 |
| 18 | Ohio State | 7,155 | 65 | Washington State | 2,687 |
| 18 | Pittsburgh | 7,155 | 66 | Georgetown | 2,662 |
| 20 | Penn State | 6,925 | 67 | Tennessee | 2,638 |
| 21 | Arizona | 6,551 | 68 | New Mexico | 2,487 |
| 22 | Duke | 6,467 | 69 | Houston | 2,457 |
| 23 | Chicago | 6,216 | 70 | Oklahoma | 2,347 |
| 24 | Southern California | 6,025 | 71 | Dartmouth | 2,279 |
| 25 | Washington U. | 5,901 | 72 | Connecticut | 2,265 |
| 26 | Iowa | 5,837 | 73 | Delaware | 2,228 |
| 27 | Texas | 5,798 | 74 | Miami | 2,200 |
| 28 | Texas A&M | 5,784 | 75 | Nebraska | 2,163 |
| 29 | North Carolina | 5,782 | 76 | UC Riverside | 2,124 |
| 30 | Northwestern | 5,490 | 77 | Temple | 1,994 |
| 31 | Maryland | 5,475 | 78 | Florida State | 1,935 |
| 32 | Purdue | 5,341 | 79 | South Carolina | 1,898 |
| 33 | Florida | 5,335 | 80 | Notre Dame | 1,857 |
| 34 | New York U. | 4,850 | 81 | Tulane | 1,855 |
| 35 | Virginia | 4,700 | 82 | Colorado State | 1,726 |
| 36 | Michigan State | 4,554 | 83 | Hawaii | 1,717 |
| 37 | Rutgers | 4,464 | 84 | Oregon | 1,714 |
| 38 | Utah | 4,340 | 85 | Syracuse | 1,640 |
| 39 | Case Western Reserve | 4,262 | 86 | SUNY, Albany | 1,608 |
| 40 | Colorado | 4,241 | 87 | Alabama | 1,379 |
| 41 | North Carolina State | 4,209 | 88 | Oklahoma State | 1,332 |
| 42 | Rochester | 4,164 | 89 | Rice | 1,256 |
| 43 | Boston U. | 4,015 | 90 | Georgia Tech | 1,211 |
| 44 | Illinois, Chicago | 3,965 | 91 | Southern Illinois | 1,142 |
| 45 | SUNY, Stony Brook | 3,918 | 92 | Brigham Young | 1,041 |
| 46 | Vanderbilt | 3,853 | 93 | Kent State | 866 |
| 47 | UC Irvine | 3,823 | 94 | Howard | 669 |

**TABLE 2**
**Institutions Ranked by Per Capita Publications**

| Rank | Institution | Number | Rank | Institution | Number |
|---|---|---|---|---|---|
| 1 | Johns Hopkins | 12.71 | 48 | Kentucky | 2.76 |
| 2 | Harvard | 11.46 | 49 | Oklahoma | 2.74 |
| 3 | MIT | 11.26 | 50 | North Carolina | 2.71 |
| 4 | Washington U. (MO) | 10.24 | 51 | North Carolina State | 2.68 |
| 5 | UCLA | 7.51 | 52 | Vanderbilt | 2.64 |
| 6 | UC San Diego | 7.34 | 53 | Texas A&M | 2.63 |
| 7 | UC Berkeley | 7.06 | 54 | Case Western Reserve | 2.52 |
| 8 | Stanford | 6.92 | 55 | Massachusetts | 2.49 |
| 9 | Minnesota | 6.90 | 55 | Oregon | 2.49 |
| 10 | Cornell | 6.81 | 57 | Penn State | 2.45 |
| 11 | Brown | 5.79 | 57 | Wayne State | 2.45 |
| 12 | Princeton | 5.46 | 59 | Virginia Tech | 2.44 |
| 13 | Chicago | 5.16 | 60 | Indiana | 2.43 |
| 14 | Southern California | 5.04 | 61 | SUNY, Albany | 2.42 |
| 15 | UC Davis | 4.96 | 62 | New York U. | 2.30 |
| 16 | Virginia | 4.82 | 63 | Connecticut | 2.29 |
| 17 | Utah | 4.79 | 63 | Illinois, Chicago | 2.29 |
| 18 | Michigan | 4.64 | 65 | Florida | 2.26 |
| 19 | Maryland | 4.61 | 66 | Rutgers | 2.18 |
| 20 | Pennsylvania | 4.61 | 66 | Tennessee | 2.18 |
| 21 | Yale | 4.57 | 68 | Michigan State | 2.17 |
| 22 | UC Santa Barbara | 4.34 | 69 | SUNY, Buffalo | 2.16 |
| 23 | Wisconsin | 4.34 | 70 | Georgetown | 2.15 |
| 24 | Duke | 4.27 | 71 | Washington State | 2.07 |
| 25 | Arizona | 4.16 | 72 | Emory | 2.02 |
| 26 | Colorado | 3.97 | 73 | Tulane | 2.00 |
| 27 | Boston U. | 3.84 | 74 | Louisiana State | 1.99 |
| 28 | Illinois | 3.78 | 75 | Georgia Tech | 1.97 |
| 28 | Purdue | 3.78 | 76 | Miami | 1.95 |
| 30 | UC Riverside | 3.71 | 77 | Dartmouth | 1.87 |
| 31 | U. of Washington | 3.68 | 78 | Florida State | 1.86 |
| 32 | Columbia | 3.61 | 79 | Georgia | 1.79 |
| 33 | Iowa | 3.42 | 79 | South Carolina | 1.79 |
| 34 | Rochester | 3.41 | 81 | Arizona State | 1.76 |
| 35 | Kansas | 3.38 | 82 | Alabama | 1.74 |
| 36 | Cincinnati | 3.19 | 83 | Syracuse | 1.72 |
| 37 | Northwestern | 3.16 | 84 | Houston | 1.60 |
| 38 | UC Irvine | 3.09 | 85 | Hawaii | 1.59 |
| 39 | Rice | 3.07 | 86 | Nebraska | 1.41 |
| 40 | Notre Dame | 3.05 | 87 | Delaware | 1.28 |
| 41 | SUNY, Stony Brook | 2.99 | 88 | Temple | 1.24 |
| 42 | Pittsburgh | 2.95 | 89 | Southern Illinois | 1.12 |
| 43 | Missouri | 2.84 | 90 | Colorado State | 1.07 |
| 44 | Texas | 2.81 | 91 | Oklahoma State | 1.03 |
| 45 | New Mexico | 2.80 | 92 | Kent State | 1.01 |
| 46 | Iowa State | 2.78 | 93 | Brigham Young | 0.74 |
| 46 | Ohio State | 2.78 | 94 | Howard | 0.50 |

statements of his process.[8] This measure is used solely because the Gourman reports are widely read and may be seen as influential.

One set of comparisons is based on a series of hypotheses concerning the relationship between the measures of publishing activity and each of the remaining variables. It should be noted that these hypotheses are not overstated; that is, there is no pretense whatsoever that whether a hypothesis is rejected or not is an indication of a causal relationship. It is most likely that both publishing and library measures are simultaneously affected by a complex of factors that includes historical mission, administrative impetus, and legislative, governmental, or political influences, among others. This series of hypotheses has to do with the goodness of fit of the raw publishing data with each of the library and other variables, and of the per capita publishing data with the same variables. In other words, each hypothesis is focused on whether the pairs of variables are independent or not, whether the two variables in each pair vary independently of one another.

Stated as null hypotheses, there is no statistically significant relationship between the publishing measures and the other variables; the assumption is that the pairs are independent. For example, there is no significant relationship between raw publishing activity and the number of volumes in the libraries. As part of a cross-tabulation function created by SPSS/PC+, the chi-square test is performed on each pair of variables: publications by volumes, publications by total expenditures, per capita publications by volumes, per capita publications by total expenditures, etc. In no instance does the computed chi-square value result in a probability equal to or less than .05, the customarily accepted decision threshold, so none of the null hypotheses can be rejected. Stated differently, there is no statistical evidence that the pairs of variables are not independent. Moreover, the chi-square test is applied to raw publications with per capita publications. As is true of the other pairs, the null hypotheses cannot be rejected in this instance either. So, even with the two publishing measures, there is some independence.

These pairs of variables can be examined in another way. For each variable—publications, volumes, doctorates awarded, etc.—the ninety-four institutions can be ranked from highest to lowest. Because of this, the rankings can be compared. Specifically, the publishing measures can be compared with the other variables. Table 3 presents these comparisons.

Rank-order correlation is employed to arrive at the correlation coefficients. For instance, the ranked list of the ninety-four universities by numbers of volumes held is correlated with the ranked list of the universities by raw publications. The resulting correlation coefficient is .678. The coefficients in table 3 indicate the correlation between each pair of measures (volumes with per capita publications, total expenditures with raw publications, etc.). As is evident from the table, the correlation coefficients are higher in each case for raw publications than for per capita publications. To re-

**TABLE 3**
**Rank-Order Correlations:**
**Publishing Measures by Other Variables**

| | Raw Pubs. | Per Capita Pubs. |
|---|---|---|
| Volumes | .678 | .416 |
| Total Expenditures | .803 | .523 |
| Materials Expenditures | .737 | .470 |
| Professional Staff | .746 | .438 |
| Doctorates Awarded | .794 | .483 |
| Gourman Rating | .767 | .754 |

The number in each instance represents the correlation coefficient.

iterate, comparing the two publishing measures results in a correlation coefficient of .793. While this is high, it is not a perfect direct correlation. The divergence between these two measures, along with the differences with regard to the other variables, indicates that, where the two measures are different, the difference is exacerbated when comparing the publishing measures with the rest of the variables. These data do not indicate that raw publications, as a phenomenon, provide an explanation for the rankings of library and other measures, or vice versa. It simply means that the direct relationship is stronger between raw publications and each of the other variables than it is between per capita publications and the variables. Perhaps the factors that influence the number of publications also affect the library and other measures. It is most likely that all of these variables are elements of a complex set of interrelated factors. It is interesting to note that the factor with the highest correlation with per capita publications is the Gourman rating. Because Gourman does not disclose his criteria for the determination of the rankings, there is no way to tell if the criteria include, or are related to, per capita publications by the faculty at the institutions.

A few things should be noted about rank-order correlation. First, since the data are ordered, rank-order correlation does not necessarily examine assumptions regarding linearity, as does product-moment correlation. For example, the interval between the first- and second-ranked cases may be much greater than the interval between the second and third. For this reason, some measures related to product-moment correlation have no relevance to rank-order correlation. The coefficient of determination cannot be applied to rank-order correlation, since it is applied to the linearity of the relationship between variables. Next, while it is possible to apply tests of statistical significance to the results of rank-order correlation, such a measure is based on the assumption that the sample used is random and independent. These assumptions do not reflect the present sample; the data used here are purposively selected.

## Summary

Although the analysis presented in this paper indicates that there are some relationships between publishing activity and other variables, care should be taken not to impart too much significance to these relationships. As is noted above, it is likely that there is a complex dynamic at work in higher education that affects faculty publishing activity. The variables examined here undoubtedly reflect that dynamic, but there is no evidence that any causal relationship exists. Rather, it is apparent that the complexity of the university manifests itself in many ways. At the most basic level in the research university, the dynamic encompasses bigness; large faculties produce large numbers of publications and libraries spend large amounts of money and have large collections. Some of these variables could be seen as inputs; these include the library measures. The others may be viewed as outcomes, such as the publishing measures and the number of doctorates awarded. When these variables are examined together, as is the case in the present study, it might be expected that the relationship exhibits a relatively high correlation. With these data, the correlations are higher for raw publishing data and the other variables than for per capita publishing and the other data.

**An extension of this study might explore a larger population of institutions, perhaps including ACRL's university library data, and adding other measures, such as citations and internal and external funding levels.**

One possible explanation why the correlation coefficients are not higher for

both measures is that the sources of publishing data used in this study concentrate on the journal literature. Because of this, publishing in the sciences and, to a lesser degree, the social sciences will be more heavily represented. Perhaps because of this, an institution such as the Massachusetts Institute of Technology is ranked high in both of the publishing activity categories, but its rank is lower for most of the library-related measures. A more normalized set of publishing data may help to eliminate any bias that might result from focus on the journal literature. Such an approach may affect any goodness-of-fit tests. It would further be expected that the correlation coefficients would be even higher. The conclusion that can be drawn here is that there is an interdependence among the array of inputs and outcomes in higher education. An extension of this study might explore a larger population of institutions, perhaps including ACRL's university library data, and adding other measures, such as citations and internal and external funding levels. This study is intended to be an initial investigation of how these higher education variables relate to one another. It remains to be seen if a more inclusive study will have similar results. It also must be noted that this study focuses on traditional academic publishing—that is, print publication. As electronic communication presents more possibilities, and as the reward system in higher education reacts to these possibilities, the dynamics of publishing and the relationships among variables may, in time, be altered.

## Notes

1. John W. Cresswell, ed., "Measuring Faculty Research Performance," *New Directions for Institutional Research* 50 (June 1986): entire issue.
2. Charles B. Osburn, *Academic Research and Library Resources: Changing Patterns in America* (Westport, Conn.: Greenwood Pr., 1979).
3. Herbert White, "Scholarly Publication, Academic Libraries, and the Assumption That These Processes Are Really Under Management Control," *College & Research Libraries* 54 (July 1993): 295.
4. Richard C. Anderson, Francis Narin, and Paul McAllister, "Publication Ratings versus Peer Ratings of Universities," *Journal of the American Society for Information Science* 29 (Mar. 1978): 91–103.
5. John A. Muffo, Susan V. Mead, and Alan E. Bayer, "Using Faculty Publication Rates for Comparing 'Peer' Institutions," *Research in Higher Education* 27, no. 2 (1987): 163–75.
6. *1991–1992 ARL Statistics* (Washington, D.C.: Association of Research Libraries, 1993).
7. Jack Gourman, *The Gourman Report: A Rating of Graduate Schools in the United States* (Los Angeles: National Education Standards, 1993).
8. See, for instance, David Webster, "Who Is Jack Gourman and Why Is He Saying All Those Things About My College?" *Change* 16 (Nov./Dec. 1984): 14–19, 45–56.