Faculty Scholarly Productivity at American Research Universities

Lawrence B. Martin*
Graduate School, Stony Brook University

lmartin@notes.cc.sunysb.edu

Draft: Please do not cite without permission from the author

Prepared for the American Enterprise Institute Conference, "Increasing Accountability in American Higher Education"
November 17, 2009

The collected papers for this conference can be found at www.aei.org/event/100134.

* Statement of interests. Lawrence Martin is the Founder and President of Academic Analytics, LLC and has a significant financial interest in the company.

Discussions of higher education accountability among policy makers have tended to focus on the value provided by universities and colleges in terms of human capital, specifically the number of well-educated graduates that is being produced. To the extent that research is considered the discussion tends to focus on somewhat nebulous terms like “innovation” and “new technologies”, usually with a link to ideas concerning economic development. The university research mission is, quite properly, much broader than that and many of the great discoveries of the 20th century that drive the economy today were the result of discoveries from basic research driven by the pursuit of knowledge rather than the search for a better product. If accountability policies continue to develop in a way that reflects the prevailing interest in degree production and narrowly-defined research, the broader scholarly mission of higher education could be damaged and marginalized.

America’s research universities have come to be the envy of the World since their rapid expansion following World War II. Much of this stature results from the extraordinary contributions made by US-based scholars to the developing knowledge base as evidenced by the quantity of research publications and citations of those publications. In recent years, America’s ability to continue to dominate the development of knowledge has been questioned as the result of rapidly improving universities in Asia and increased emphasis on research excellence at universities in Europe. Past President of the National Association of Land Grant Universities and Colleges NASULGC, now re-named the Association of Public and Lad-grant Universities [APLU}), Peter Magrath and Association of American Universities (AAU) President Bob Berdahl have separately stated that they believe that the United States has too many research universities to sustain. The National Academies study entitled "The Gathering Storm" raised similar concerns about American competitiveness in STEM disciplines and recent reports in the Chronicle of Higher Education have suggested that the level of productivity of faculty at American Universities has declined in the last decade.

The research mission of universities should not be exempt from assessment and accountability as it is both expensive and vital to national competitiveness. However, it is critical to employ fair, accurate, nuanced metrics of research productivity that provide a basis to compare and assess performance while recognizing the breadth of the scholarly enterprise. In particular, it is important to pay close attention to cross-disciplinary differences in performance that require more sophisticated data on performance within disciplines as the basis for assessment. In most cases, such metrics need to be standardized (to standard deviation or z-score, units) to permit comparisons among disciplines. This chapter addresses the following questions:

1. How productive are faculty at American research universities?
2. How much variation in productivity is there among disciplines?
3. How is scholarly productivity best measured and compared across disciplines?

Background to the use of Faculty Scholarly Productivity Data to Evaluate Program Quality

In 1982 and again in 1995, the National Research Council published an assessment of Research doctorate (Ph.D.) programs in the United States. These studies contained information on faculty scholarship at the discipline level but their results for ranking programs derived from surveys of faculty. In the 1995 study, faculty members were asked to rate programs in their own discipline on the basis of their assessment of the scholarly quality of the faculty of the program on a scale from 0 (not sufficient for doctoral education) to 5 (distinguished). The average scores for programs were reported as 93Q. The fiercest critics of the use of "reputational", as opposed to "objective", metrics of quality were the late Hugh Graham and his collaborator Nancy Diamond who argued that few faculty had sufficient, accurate and current knowledge of 50 programs in the country to assess their quality on the NRC's six point scale. In their 1997 book, Graham and Diamond concluded that only per capita metrics should be used when comparing programs made up of different numbers of faculty members. Graham and Diamond also analyzed data from the 1995 NRC study and concluded in their 2001 paper that the best single metric to measure quality was the number of journal article citations per faculty member for the program.

Beginning with the release of the last National Research Council report on Research Doctorate Programs in the United States in 1995, I began to analyze data on faculty scholarship in relation to the perceived quality of Ph.D. programs in the United States. Initial analyses were based on data provided in the 1995 NRC book and in downloadable Excel files. These analyses led to the conclusion that some, but by no means all, of the metrics compiled by the NRC correlated well with the overall rating of programs using the NRC's summary statistic (93Q). Most importantly, aspects of scholarly performance on which most believe that academic quality and reputation is based (journal articles, citations of articles, Federal research grants, and prestigious honors and awards) were shown to be useful metrics of quality based on recent performance. Some programs had performance metrics that exceeded the perceived quality of the program (whose reputation might therefore be expected to improve in the future) while others had lower performance measures that their reputation would suggest and these programs might be vulnerable to declining reputation in the future. An early conclusion of this work was that size adjusted metrics such as per capita measures and ratios were much more useful in comparing programs that are the raw numbers.[1]

Available discipline level data were later supplemented by the NRC with the release of more complete data files made available on CD-ROM in 1997. At this stage my research effort was enhanced when I was joined by Anthony Olejniczak who increased the level of sophistication of our analyses of performance metrics in relation to the perceived quality of programs. With the addition of more complete data on faculty performance it was possible to develop an index of faculty scholarly productivity (FSP) for each of the 41 disciplines in the 1995 NRC study that covered journal publications, citations, grants and honors and awards.

The FSP results were of interest to many as ways to look at current performance as a means to identify strong programs that had not yet achieved national recognition and programs whose reputation was unlikely to persist unless scholarly performance was to improve. The overall conclusions of our analyses of the 1995 NRC data were that, while there is agreement that scholarly performance in terms of journal articles, citations, grants and honors and awards is well correlated with and probably the underlying basis of academic quality assessments, performance in these areas must be measured in the appropriate way for the particular discipline and that per capita or other size corrected measures have the most power to describe performance differences among programs. We also shared the prevailing view of the NRC Methodology Committee for the current NRC study that assessments of scholarly performance in the humanities and other lettered disciplines that did not include data on books published were incomplete and unreliable.

Beginning in 1995, Academic Analytics, LLC (AA) began to build annually a national database on faculty scholarly productivity (FSP) in American research universities and data from AA’s FSP 2007 database are used in this chapter. AA’s data are proprietary but the company has a policy to make data available to scholars of higher education for use in academic work (such requests may be directed to AA's CEO, Peter Maglione, peter@academicanalytics.com).[2] The initial goal for the database was to include all elements of faculty scholarly productivity included in the 1995 NRC study, with the addition of authored and edited books and with an increased coverage of scholarly disciplines. This approach led to the adoption of a taxonomy of Ph.D. disciplines that includes 172 disciplines. To facilitate comparisons of performance across universities employing different program groupings (for example, Biomedical Sciences as a general category in contrast with a finely divided set of disciplines) the disciplines are grouped into eleven broad fields that mirror the taxonomy commonly used by Federal Agencies, Universities and the NRC (for example, Social and Behavioral Sciences,; Biological and Biomedical Sciences). Following the NRC approach, AA assembled lists of faculty associated with Ph.D. programs in the United States and used these lists to compile data on journal articles published, citations of those articles, books published, Federal research grants awarded by the principal Federal funding agencies, and prestigious honors and awards received.

Data on Faculty Scholarly Performance used in the Current Study

For the FSP 2007 database, from which the data in this chapter are drawn, the data contained are as follows:

Data on journal articles and data on citations of those articles were obtained from Scopus (http://www.scopus.com), covering over 16,000 peer-reviewed journals from more than 4,000 publishers. Each year of journal article data represents a complete year's worth of scholarly activity captured by Scopus, amounting to more than 6.5 million journal articles per year. FSP 2007 includes journal publications from 2005, 2006 and 2007 that had been indexed by Scopus by May 2008). FSP 2007 includes four years of citations to journal articles that were published 2004-2007. All authors are recorded for journal articles and for books.

FSP 2007 employs book publication data from Baker and Taylor, Inc. (http://www.btol.com) whose academic library service offers a comprehensive catalogue of titles. Each year of initial books data contains approximately 266,000 titles, for a total of approximately 1,328,500 books over five years. FSP 2007 contains data on books published in 2003, 2004, 2005, 2006 and 2007 in six of the broad fields (Business; Education; Family, Consumer and Human Sciences; Health Professions Sciences; Humanities; and Social and Behavioral Sciences). A title is included once per person and all published works are weighted equally.

FSP 2007 includes three years (2005, 2006, 2007) of data on new federal funding from five agencies (National Institutes of Health [NIH] [170,000+ grants /05-07], National Science Foundation [NSF] [39,000+ grants/05-07], Department of Education [DoED] [18,100+ grants 2005-2007], National Oceanic and Atmospheric Association [NOAA] [3,100+grants /05-07], United States Department of Agriculture [USDA] [2,100+/05-07]) and also for two programs at the Department of Energy (DOE) (Office of Nuclear Physics and Office of Biological and Environmental Research) (357 grants/2005-2007). Only new, competitive awards (including competitive renewals at NIH) are collected; continuing awards, non-competitive grants or non-competitive renewals are not included in FSP 2007. Only Principal Investigators are counted for grants awarded as this is the common element for all of the data sources currently obtained (NIH only added co-PI's in 2007).

FSP 2007 includes data for 62,761 recipients of 2,517 honorific awards, issued by 295 governing held by faculty in 2007 (a complete list showing durations for which the data are compiled, which varies by award, may be obtained on request from Academic Analytics).

For FSP 2007, faculty are grouped into disciplines based on their listing as faculty of a Ph.D. program or a Department that offers a Ph.D. program. Data on faculty performance are combined for the individuals making up the faculty of a program and metrics of performance for the whole faculty are then added (percentage with a grant etc.). The individual data elements available in FSP 2007 are:

- The number of journal articles published by the faculty
- The number of faculty who authored or co-authored a journal article
- Journal articles/faculty
- Journal articles/author
- Authorship index (percentage of faculty who authored an article)
- The number of citations to all journal articles
- The number of faculty with a citation
- The number of authors of journal articles with a citation
- Citations/faculty
- Citations/author
- Citations/article
- The number of books published by the faculty in the discipline
- The number of authors of books among the faculty of the discipline
- Books/faculty
- Books/author

- Authorship index (percentage of faculty who authored a book
- The number of NEW grants awarded
- Average annualized funding for grants awarded
- The number of faculty who are PI's of grants awarded
- Grants/faculty
- Grants/author
- Research dollars/faculty
- Research dollars/author
- Research dollars/grant
- The number of awards held
- The number of faculty holding awards
- Awards/faculty
- Awarded faculty index (percentage of faculty holding an award or honor)

The Scale of the Scholarly Endeavor in American Research Universities

In 2007, 8,737 Ph.D. programs in the United States at the 383 accredited universities that award a Ph.D. listed a total faculty membership of 235,217 (programs that are delivered as distance programs and for which faculty lists are not available were not included). As many faculty members belong to more than one Ph.D. program these faculty counts reduce to a total of about 170,000 individual people. In the FSP 2007 database, each person listed counts fully in each program rather than being pro-rated across the various programs to which they belong. To avoid double counting of performance when comparing larger units (whole universities, broad fields of disciplines etc.) faculty records are de-duplicated at each level of aggregation of disciplines. This means that a faculty member who appears in the list for both Biochemistry and for Genetics at a particular university will count as a single person in the data for each discipline and also as a single person when Biomedical Sciences is considered as a broad field or when the performance of the entire university is measured. Table 1 shows the relative size of the faculty associated with the 172 disciplines included in FSP 2007.

In the period under study, 152,663 of these 170,000 scholars published 1,086,252 journal articles, which attracted 10,514,390 citations. 25,752 of the scholars authored or edited 54,878 books. They were awarded 89,755 grants with an estimated annual expenditure value in 2007 of $30,380,307,000. 20,205 of these scholars held 26,408 prestigious honors and awards.

The contribution of American research university faculty to scholarship is substantial but the costs to produce the scholarship are not fully documented or understood.

A significant component of the national investment in scholarship is the investment made by research universities in faculty resources to allow them to fulfill both teaching and scholarly commitments. As an initial estimate let's assume that at a typical four year college a faculty member may teach six courses per year (3+3) in the sciences and eight courses per year (4+4) in the humanities and social sciences. At a research university faculty will teach at most half of this load in the social sciences and humanities and perhaps one third of the load in the sciences. This means that at least half, and more likely about 60% of the costs associated with faculty at research universities are costs to allow them to be productive as scholars. If the average faculty member were to earn $80,000 per year (all levels and all disciplines at Ph.D. institutions) then the universities' investment in scholarship would amount to somewhere between 7 and 8billion dollars annually.

As the salary assumption levels are low for research universities these are the *lowest* likely costs for faculty effort. Given the current status of funding for higher education these investments demand some assessment of the return that is being received.

Publishing in Scholarly Journals.

The primary medium for dissemination of new knowledge in the Sciences is the scholarly article in a peer reviewed journal. The absolute number of such publications is obviously an important output measure but is less interesting for studying productivity than other metrics as the number of faculty involved in producing the output varies so much among disciplines (Table 1).

Table 2 shows the percentage of Ph.D. program faculty who authored or co-authored at least one article during a three year period. It is perhaps not surprising that no discipline has 100% of the faculty publishing but it is quite surprising that so few approach this figure. As much of the variance appears to be patterned by discipline, the results in Table 2 are grouped according to the eleven broad fields in the FSP taxonomy. Even within these groups there is a considerable difference between the high and low publication rate across disciplines (Agricultural Sciences 72%-52%, Biological and Biomedical Sciences 91%-68%, Business 63%-39%, Education 55%-25%, Engineering 84%-67%, Family, Consumer and Human Sciences 78%-29%, Health Profession Sciences 82%-48%, Humanities 48%-3%, Natural Resources and Conservation 77%-72%, Physical Sciences and Mathematics 89%-64%, Social and Behavioral Sciences 75%-45%).

While differences among these broad fields are intriguing some of them simply reflect the pattern of scholarly publishing so that humanists and social scientists may publish their work in books rather than journals. It is also the case that the Scopus database for the period concerned did not include complete coverage of the humanities, arts and lettered social sciences journals. That being said, the significant differences in publication rates within broad fields is striking as is the fact that over all broad fields (including those for which book publishing is the primary forum) only 65% of faculty publish an article in a three-year period.

Even more surprising, in most fields for which journal publishing would be expected fully 20% of the faculty associated in Ph.D. training programs have not authored or co-authored a single publication in one of the 16,000 journals indexed in a three year period. It is possible that these faculty members have been productive in other areas but one has to wonder how suitable they are to mentor the next generation of scholars based on this low level of productivity.

Returning to rough cost estimates, if 20% of Ph.D. faculty are not publishing as expected this represents something close to $2 billion in university investments in scholarship for which the return seems to be nil.

Another way to look at productivity is to examine the number of publications per faculty member and these also vary dramatically among disciplines (Agricultural Sciences 4.35-2.01, Biological and Biomedical Sciences 10.35-3.58, Business 2.74-1.06, Education 1.76-0.49, Engineering 9.24-3.85, Family, Consumer and Human Sciences 6.11-0.97, Health Profession Sciences 6.60-2.04, Humanities 1.03-0.09, Natural Resources and Conservation 4.52-3.79,Physical Sciences and Mathematics 9.86-2.67,Social and Behavioral Sciences 4.34-1.02). It is almost certainly a mistake to compare productivity levels among disciplines even within a broad field as this variance likely represents different patterns of scholarly output in different disciplines. However, it is a cautionary note for those who seek to compare or to evaluate performance levels in fields beyond those for which they have immediate expertise or appropriate comparative, discipline-level data. A significant effect on the numbers reported above could result from different proportions of faculty who publish books as compared to journal articles across the disciplines.

To reduce this effect, FSP 2007 also calculates the number of publications per author (that is, leaving out the faculty count for individuals to whom no article was matched). This number can be compared with less caution as it will have eliminated from the denominator both those who have been correctly identified as having zero publications and those whose publications were not properly matched to them as the result of errors caused by recent changes of address or name etc. (Agricultural Sciences 6.01-3.24, Biological and Biomedical Sciences 11.79-5.28, Business 4.37-2.31, Education 3.22-1.83, Engineering 11.34-5.54, Family, Consumer and Human Sciences 7.82-2.80, Health Profession Sciences 9.98-3.43, Humanities 3.65-1.27, Natural Resources and Conservation 6.28-5.02,Physical Sciences and Mathematics 12.09-4.18,Social and Behavioral Sciences 6.27-2.25).

The fact that for within most of the eleven broad fields there is a doubling of output from the low number to the high indicates that disciplinary standards for a publishable unit are quite different even among disciplines that one would expect to behave in a similar manner (for example, Physical Sciences and Mathematics). The fact that the lowest performance levels occur in broad fields for which significant book publishing activity takes place may also reflect the fact that scholars in these fields may publish both books and articles. Again, these results represent a cautionary tale for cross disciplinary performance evaluation and thus for many college, school or university-wide promotion and tenure committees.

Return on Investment of Federal Research Funding

In the sciences, great weight is placed on the ability to secure funding to support research so it seemed interesting also to look at the number of articles produced as the consequence of funding to support the production of scholarship. The broad fields of Business, Education and Humanities have been eliminated from this analysis as they are not areas for which a great deal of Federal funding for research exists. It is important to remind readers at this point that some Federal funding agencies are not yet included in the database (Department of Defense, Department of Energy and NASA being the most critical). This means that comparisons among broad fields should be treated as preliminary at this point as the missing agencies are disproportionately involved in funding of Engineering and Physical Sciences whose costs per publication and per citation may thus be underestimated). The comparisons of return-on-investment within the eleven broad fields should be much less affected and thus much more reliable.

The results of this return-on-investment analysis are really quite striking so the detail is presented in full in table 3. To summarize for the broad fields included the high and low figures for disciplines are: (Agricultural Sciences 241-55, Biological and Biomedical Sciences 110-14, Engineering 585-37, Family, Consumer and Human Sciences 371-27, Health Profession Sciences 60-22, Natural Resources and Conservation398-124, Physical Sciences and Mathematics 223-37,Social and Behavioral Sciences 758-25).

It is the range of the variance overall and within broad fields that is so striking. The fact that $1,000,000 of research funding in Pharmacology (the most expensive research field by this measure) results in 14 papers while the same amount produces 110 papers in Evolutionary Biology and 758 in Applied Economics was not something that I knew or expected! Some work is clearly much more expensive to do that others but I have been surprised to see the scale of this effect with, for example, physicists produce 92 journal articles per $1M while biochemists produce 21. This result does not appear to be the consequence of differences in what represents a publishable unit as biochemists publish more articles per author (8.95) than do their physicist colleagues (6.99).

The Impact of Journal Articles on the Field

Since the pioneering study by Cole and Cole (1967), citation analysis has become an important tool to understand the impact of publications. Most scholars can find some outlet for their work so citation data provide an important counterbalance to counts of raw output when these are not graded in terms of quality. The standard way to grade publications is to use the citation impact for the journal (that is the average citation rate for the papers in that journal over a defined period). Unfortunately, for interdisciplinary or broadly based journals, (such as Nature or Science there is a strong discipline influence on paper citation counts within the journal (Neuroscience papers in Nature have higher citation counts than do Geology papers in Nature) so that the impact factor is equally applicable to all articles. As FSP 2007 enables analysis of actual citation numbers per paper and per author it provides a more independent assessment of publication quality than a weighted count of publications.

Raw counts of citations for broad fields show a huge discipline effect that has generally been ascribed to the number of scholars working in (and thus available to cite other work) a discipline. For that reason, the standard metric for citations used by the NRC in 1995 and by Graham and Diamond[3] in their analysis of NRC data has been citations per faculty member. Within a tightly constrained discipline this may be a good metric but it is clearly influenced by some combination of the number of practitioners in the discipline, the number of papers in the discipline and discipline-particular citation practices. For the broad fields, the ranges of values are: (Agricultural Sciences 28.15-6.63, Biological and Biomedical Sciences 155.41-25.16, Business 9.67-6.38, Education 6.41-0.99, Engineering 67.32-9.90, Family, Consumer and Human Sciences 61.05-2.75, Health Profession Sciences 65.68-8.53, Humanities 4.09-0.04, Natural Resources and Conservation 28.36-16.14, Physical Sciences and Mathematics 113.26-9.67, Social and Behavioral Sciences 35.09-3.49).

One obvious impact is in the lettered disciplines but the dramatic differences in values within broad fields where journal articles are the norm demonstrate that this metric should be used with great caution across disciplines. To enable comparison of performance across disciplines the data should be transformed to standard deviation units (z-scores) for the programs or people in a discipline and comparisons across disciplines should use z-scores rather than raw data).

The impact of book publishing fields on the citations per faculty index can be eliminated by replacing faculty number as the denominator with the number of authors of journal articles. When this is done the broad field results are: (Agricultural Sciences 43.59-10.70, Biological and Biomedical Sciences 176.30-34.62, Business 17.37-11.87, Education 14.18-3.77, Engineering 83.06-14.25, Family, Consumer and Human Sciences 78.15-8.93, Health Profession Sciences 116.81-14.34, Humanities 22.92-0.45, Natural Resources and Conservation 39.41-21.38,Physical Sciences and Mathematics 138.79-15.17,Social and Behavioral Sciences 57.39-7.72). There clearly remains a substantial discipline and broad field effect on this metric, which renders it useless to compare performance across disciplines except in a statistically standardized space.

To address the co-linearity of data on counts of articles and counts of citations, the FSP 2007 database also measures citations per paper, which is analogous with the citation impact factor for journals but applied to individual articles. This index should be less influenced by the number of scholars in a field or by the number of articles published in the field and might be expected to provide a better basis for comparison across fields. The broad field results are: (Agricultural Sciences 8.72-3.30, Biological and Biomedical Sciences 17.35-6.09, Business 6.03-3.42, Education 5.01-2.02, Engineering 8.51-2.57, Family, Consumer and Human Sciences 9.99-2.83, Health Profession Sciences 11.71-4.18, Humanities 9.71-0.36, Natural Resources and Conservation 6.27-4.18,Physical Sciences and Mathematics 13.88-2.98,Social and Behavioral Sciences 9.15-3.14).

Clearly this index is the one that has the least problems in comparing performance among the broad fields but there is still a dramatic effect across disciplines that suggest that comparisons of citations/paper should only be made in statistically standardized space.

The final analysis of citation data is another return-on-investment approach that looks at the number of citations to journal articles that result from $1,000,000 of Federal Research funding from the agencies included in this study. The broad field results are: (Agricultural Sciences 1066-345, Biological and Biomedical Sciences 1071-167, Engineering 1848-273, Family, Consumer and Human Sciences 1835-265, Health Profession Sciences 405-130, Natural Resources and Conservation 1696-694, Physical Sciences and Mathematics 2047-433, Social and Behavioral Sciences 2492-171). As was the case for the analysis of publications in relation to research funding, the results of this comparison are so intriguing that the discipline detail in presented in full in table 4. Overall the most striking result is that it is so much more expensive to produce citations in biomedical science fields than in other laboratory sciences.

Federal Grants in Support of Research

Many universities pay close attention to the volume of federally sponsored research grants at their campus and these figures are often seen as a measure of the quality of the scholarly endeavor. While that is almost certainly the case when making comparisons among funding levels within a discipline the results in table 5 suggest that great caution should be applied in making comparisons among disciplines (or departments) within a university.
The data show that the average size of a grant varies significantly among disciplines, which suggests that comparisons of performance need to be adjusted to discipline level norms. The broad field results for the range of the average value of a three-year grant for each discipline are: (Agricultural Sciences $264,978-$124,596, Biological and Biomedical Sciences $550,167-$149,114, Engineering $346,418-$88,252, Family, Consumer and Human Sciences $366,293-$130,431, Health Profession Sciences $515,284-$276,915, Natural Resources and Conservation $187,290-$113,098, Physical Sciences and Mathematics $597,862-$82,381,Social and Behavioral Sciences $523,903-$59,419). While it has long been appreciated that it is unwise to compare funding levels between physical and social sciences, less attention has been paid to the very real differences in funding availability within broadly similar disciplines. These render comparisons of expenditure data among faculty members or departments almost meaningless unless they are adjusted to reflect disciplinary norms. On most campuses administrators and faculty fail to make such adjustments when discussing success in obtaining external support for research.

The total amount of sponsored research reflects not only the average size of a grant but also the number of grants held by faculty. This measure of availability of grant funding needs to be adjusted to reflect the number of people competing for the funds and this has been done by looking at the number of grants awarded per 1000 faculty members in Ph.D. programs in the discipline. These data show that not only are grants in some fields much larger but that they are also more plentiful. The broad field results for the number of grants awarded in a three year period per thousand faculty members are: (Agricultural Sciences 245.06-62.50, Biological and Biomedical Sciences 1476.08-233.46, Engineering 692.83-67.73, Family, Consumer and Human Sciences 628.91-38.82, Health Profession Sciences 530-71-106.55, Natural Resources and Conservation 194.54-72.75, Physical Sciences and Mathematics 559.22-192.66,Social and Behavioral Sciences 358.62-33.22).

Scholarship Published as a Book.

For this discussion, the data on books have been limited to the broad fields Business; Education; Family, Consumer and Human Sciences; Health Professions Sciences; Humanities; and Social and Behavioral Sciences where most of book publishing occurs. In these broad fields, 25,666 faculty authored or edited a book indexed by Baker and Taylor for the period 2003-2007. Among them they produced 54,738 titles. The same ten disciplines top the lists of the number of authors of books (History, English, Political Science, Sociology, Psychology, Anthropology, Philosophy, Economics, Business Administration and Religion/Religious Studies, in rank order) and the numbers of volumes produced (History, English, Political Science, Sociology, Psychology, Philosophy, Economics, Anthropology, Business Administration and Religion/Religious Studies, in rank order).

Although discussions of scholarship in the Humanities have long focused on the importance of books it is striking that five (or six depending how one views History) of the top disciplines for book publishing are Social and Behavioral Science fields. This reinforces the importance of including data on scholarly books in any assessment of productivity or quality. Results for the percentage of faculty to whom a book was attributed range from a high of 53% in Religion/Religious Studies to a low of 9% in Nutrition Sciences. These results suggest that a five year window for book publishing captures a large portion of the scholarly activity and is likely a representative sample.

Productivity can also be measured as the number of books produced per person and there are two reasonable denominators to use in such studies. First is the number of faculty, second the number of authors. The second of these has the advantage that fields in which faculty may either publish books or articles are not disadvantaged. Based on books per book author, productivity levels are impressive ranging from the highest average for a discipline of 2.55 books in Accounting to a low of 1.47 books in Environmental Health Sciences. The complete results for books per book author by discipline are presented in Table 6.

Honorific Awards in Recognition of Scholarly Performance.

The 1995 NRC study included data on prestigious honors and awards and these data proved valuable to measure program quality in the lettered disciplines (for which data on books were not available). FSP 2007 includes a much wider range of honors and awards than were compiled for NRC 2007 and these are designed to be used within disciplines as the availability of awards is by no means uniform across disciplines. For that reason, comparisons of success in gaining awards need to be made in standardized statistical space. The ten disciplines for which the largest numbers of faculty holding awards were found are History (1,066), Chemistry (636), Physics (592), English Language and Literature (564), Neurobiology/Neuroscience (550), Mathematics (549), Molecular Biology (489), Mechanical Engineering (443), Biology/Biological

Sciences (439). The ten disciplines with the largest numbers of awards held are: History (1,488), Chemistry (938), Physics, (843), Mathematics (743), English Language and Literature (741), Neurobiology/Neuroscience (710), Molecular Biology (647), Biology/Biological Sciences (587), Cell Biology (572).

Obviously these two results are strongly influenced by the scale of the discipline so success is perhaps better measured by looking at the percentage of faculty who hold a prestigious award. In this case the top ten disciplines are: Slavic Languages and Literatures (21%), History (21%), Art History and Criticism (21%), Classics and Classical Languages (20%), Chemical Sciences (18%), Ancient Studies (18%), Humanities/Humanistic Studies (18%), Comparative Literature (17%), Italian Language and Literature (16%), Asian Studies (16%). It is possible to examine the relative abundance of included honors and awards as a count of awards held per 1,000 faculty members. In this case the values range from a high of 337 awards per 1000 faculty in Slavic Languages and Literatures to a low of 19.15 awards per 1000 faculty in Counselor Education. The top ten disciplines in terms of availability of honorific awards in this database are: Slavic Languages and Literatures (337.02), History (290.80), Art History and Criticism (279.42), Classics and Classical Languages (261.64), Chemical Sciences, various (253.57), Comparative Literature (249.02), Asian Studies (225.95), Italian Language and Literature (225.23), Asian Languages (221.97), Ancient Studies (221.52). The dramatic differences in terms of the relative abundance of honorific awards by disciplined in standard compilations (e.g., the NRC study or Academic Analytics' compilation) demonstrates once again that performance comparisons among disciplines based on raw data are at best problematic and probably misleading and ill advised.

Conclusions

This chapter has examined data on faculty scholarly performance by discipline in terms of publication of scholarship (journal articles and books), citation data for journal articles, Federal research funding and the receipt of honors and awards. What is immediately clear is that few data elements lend themselves to performance comparison across discipline boundaries as it is abundantly clear that the style of scholarship in a discipline dictates strongly how scholarship is packaged for publication (books, long articles, reports etc.) so that raw counts mislead more than they inform. That being said, return on investment analyses reveal dramatic differences in the cost to produce publications of new knowledge and to achieve penetration into the field as indicated by citation of the work by others and these finding warrant further study.

For those who wish to use faculty scholarly productivity measures to compare performance within or across disciplines there are some important rules to follow:

1) Unless comparisons are being made among units with equal numbers of faculty (or narrowly constrained bands of department or program size) the metrics should be per capita.
2) The best metric to measure performance is the metric that answers the particular question that is most important to the stakeholder. This question must therefore be defined, which seems rarely to be done. I would recommend the selection of one per-capita metric in each of the key areas of faculty scholarly performance: book publications, journal publications, citations, research grants and honors and awards.
3) Performance comparisons among disciplines must be made in units standardized to reflect the different availability of resources and different scholarly practices among disciplines. Means and standard deviations for national datasets should be used to

generate z-scores that may then be used to compare relative success across disciplines (a gold medal is always better than a bronze whether won in swimming or in the marathon).

---

[1] Lorden and Martin,; Lorden and Martin,

[2]

[3] (2001)