

# Bugs in the NRC's Doctoral Program Evaluation Data: From Mites to Hissing Cockroaches

**Nelson C. Dometrius**, *Texas Tech University*

**M. V. Hood III**, *Texas Tech University*

**Kurt A. Shirkey**, *Texas Tech University*

**Quentin Kidd**, *Christopher Newport University*

In 1995, the National Research Council (NRC) released *Research Doctorate Programs in the United States, Continuity and Change*. This report was described as containing an "extensive, comprehensive, and systematic assessment of the quality, effectiveness (in training scholars), faculty productivity, and Ph.D. production of over 3,600 doctoral programs in 41 disciplines at 274 Universities" ("Departmental Rankings" 1996, 144). The study offers a rich array of objective and subjective data that its authors expected to be useful to policymakers, academic administrators, faculty, and prospective graduate students (NRC 1995). Scholars quickly took advantage of the data to examine sources of high reputational evaluations, or to compare reputational scores with objective measures (Katz and Eagles 1996; Jackman and Siverson 1996; Lowry and Silver 1996; Miller, Tien, and Peebler 1996).

Early studies, however, relied on the data as published in the report. Any massive project of this nature is bound to suffer some problems related to data quality or interpretation—perhaps minor, perhaps

severe. About a year ago, the NRC made available on CD-ROM the raw data used to generate the report. We examined the data in two ways. The first was primarily an exercise in reverse engineering—probing the data to uncover the guidelines followed in their collection, coding, and presentation. Second, we compared the NRC data with our own file of publications in eight leading journals to identify any differences and, if found, determine their source. Our study reveals a number of problems with the NRC data, ranging from coding decisions that may be inappropriate for some purposes, to some outright errors in the publication and citation data. Each problem identified renders suspect a number of the conclusions drawn by the NRC and later secondary analyses.

## The Survey of Graduate Faculty

The Survey of Graduate Faculty was the data source for the report's reputational assessments. It consisted of a series of five questions

about the quality of a program and its faculty and how familiar the rater was with both the program and the faculty (Table 1), that was sent to a randomly selected group of graduate faculty at programs participating in the study. Each rater was given evaluation forms for 50 programs. The NRC's goal was to have a sufficiently large pool (200 in the case of political science) to produce 100 raters evaluating each program. The NRC did end up with 100 raters that were sent evaluation forms for each program, but response rates varied widely. The published report provided a "visibility score" for each program, presented as the percent of evaluators who had some knowledge of a program. The lowest visibility score reported for any political science program was 45%, with most scores at 50% or more, suggesting that every program in the study had at least a fair sample of raters producing its evaluations.

We believe the visibility score is quite susceptible to misinterpretation, giving the impression of greater reliability in the data than actually exists. The operational definition of "some knowledge" that produced the NRC's visibility score was excessively broad. Raters were asked about their knowledge of each program (questions 1 and 3 in Table 1), but these responses did not constitute the visibility score. Instead, for a program to be considered "not visible" under NRC's coding system, a rater would have to answer "little or no familiarity" to both questions 1 and 3, and respond "Don't know well enough to evaluate" to questions 2, 4, and 5. Any other answer to any of the five questions ensured that the program was counted as visible—a generous interpretation.

A stricter definition gives a quite

---

**Nelson C. Dometrius** is professor and past chair of the department of political science at Texas Tech University. His past research has appeared in *American Journal of Political Science*, *Journal of Politics*, *Political Research Quarterly*, *Public Administration Review*, and other journals.

**M. V. Hood III** is a graduate of, and currently a visiting assistant professor in, the department of political science at Texas Tech University. He has published previously in *Political Research Quarterly*, *Social Science Quarterly*, *Political Behavior*, and *Legislative Studies Quar-*

*terly*, and has an article forthcoming in the *American Journal of Political Science*.

**Kurt Shirkey** is a doctoral student in the department of political science at Texas Tech University. He has published previously in *Political Research Quarterly*.

**Quentin Kidd** is a graduate of the department of political science at Texas Tech University and is currently an assistant professor in the department of government and public affairs at Christopher Newport University. He has published previously in *Social Science Quarterly* and has an article forthcoming in the *American Journal of Political Science*.

**Table 1****NRC 1993 National Survey of Graduate Faculty**

1. Familiarity with work of Program Faculty.
  - Considerable familiarity
  - Some familiarity
  - Little or no familiarity
2. Scholarly Quality of Program Faculty
  - Distinguished
  - Strong
  - Good
  - Adequate
  - Marginal
  - Not sufficient for doctoral education
  - Don't know well enough to evaluate
3. Familiarity with Graduates of this Program
  - Considerable familiarity
  - Some familiarity
  - Little or no familiarity
4. Effectiveness of Program in Educating Research Scholars/Scientists
  - Extremely effective
  - Reasonably effective
  - Minimally effective
  - Not effective
  - Don't know well enough to evaluate
5. Change in Program Quality in Last Five Years
  - Better than five years ago
  - Little or no change in the last five years
  - Poorer than five years ago
  - Don't know well enough to evaluate

Source: NRC (1995, 21)

different picture. Counting as visible only those programs about which raters said they had “some” or “a great deal” of knowledge, the visibility of program faculties in political science ranges from a high of 99% to a low of 11%. For knowledge of a program's graduates, the less expansive definition produces visibility scores ranging from a high of 95% to a low of only 4%. In fact, for only 30% of the programs did one-half or more of the raters claim to have “some knowledge” or better of a program's graduates.

This is not to say that the reputational evaluations are inaccurate. They may or may not be. But it does point to the bane of quantitative analysts; systematic measurement error. The nature of the problem becomes clear upon separating programs into thirds based on their faculty quality scores (Table 2). For schools in the top third, virtually all the evaluations they received—of their faculty, their program, or change in their program—came from raters with some knowledge about them. For schools in the second third, 22 to 25% of their program and program change evaluations came from raters with little knowledge of them. Schools in the bottom third fared the worst, with almost 30% of their faculty quality evaluations, and about 40% of their program evaluations, coming from those with little knowledge about them.

A program's reputational score is affected by its visibility. This is sometimes sensible, and sometimes not. For a rater to assume that not recognizing many of a program's faculty justifies a low faculty quality evaluation is reasonable. If the faculty were actively involved in the scholarly enterprise, most raters should know something about them.

Somewhat more problematic, however, is the apparent working assumption that knowing little about a program's graduates means the program's quality is also low. New programs cannot expect to break immediately into the national scene, and small programs, new or old, will probably market most of their graduates regionally, not nationally.

Thus, while it is not completely unreasonable to equate visibility with program quality, the connection between the two is more tenuous than it was for visibility and faculty quality. There is a good probability of monotonically increasing measurement error in assessments of program quality as visibility declines.

Finally, for question 5, we see no reason why an absence of knowledge provides any guidelines whatsoever for determining whether a program has become better or worse over the last five years. For all except the most visible programs, we consider responses to this question to be es-

entially meaningless, and analysts should use them at their peril.<sup>1</sup>

### Objective Data: Collection and Coding Decisions

The NRC supplemented the Survey of Graduate Faculty with a number of objective indicators, including publications, citations, and awards garnered by a program's faculty. Such a rich data source is undoubtedly tempting to many scholars but, again, caution is needed. The NRC's descriptions of the publications data can be readily misinterpreted, possibly leading some to see things in the data that are not really there.

To begin with, the NRC describes its measures as including “both a count of papers published in reviewed journals and monographs printed by recognized publishers” (NRC 1995, 143). The publication data set for political science is massive, containing nearly 8000 entries in over 1,000 different journals. Indeed, its sheer size raises questions about the term “reviewed journals.” Most items in the NRC publications file seem to fit within the general academic meaning of “journal,” but also included are such periodicals as *New York Review of Books*, *New Republic*, *BYTE*, and *Popular Computing*. Without knowing the publication requirements of each journal, it is hard to determine if we are looking at minor and sporadic problems or something more significant. Either way, scholars should explore the data file carefully before assuming that “reviewed” means “refereed.”

From inclusion to exclusion, some of the NRC's comments may lead scholars to assume the publications file is more comprehensive than it actually is. The data are described as “the publication record of the faculty” (NRC 1995, 143), or “publication/citation patterns for all 78,000 faculty members who were involved in doctoral training in fall 1992” (NRC 1995, 37). Neither is precisely accurate.

The NRC obtained the publication data from the Institute for Scientific Information (ISI), publishers of the *Science Citation Index*, *Social Science Citation Index*, and the *Arts and*

*Humanities Citation Index*. To provide the highest quality data, the NRC matched the names of authors from the ISI database with the names of faculty provided by the Institutional Coordinators (ICs) for each institution in the study. After the names were matched, they were further screened by matching the ZIP codes of authors in the ISI data with the ZIP codes of the participating institutions, using an algorithm that allowed ZIP codes near the institution to be accepted as well, in case the author used a home address.

While the NRC should be commended for engaging in such a major effort to provide clean data, their cleaning process has also produced a publications file with clear limitations. This is pointed out, though not completely, in the introduction to the codebook for the publications

at whatever institution they were affiliated with in 1992.

Some examples will make the point, and potential problems, clear. As part of a related project, the authors collected data on publications in the *American Political Science Review*, the five regional association journals—*The Journal of Politics*, *Political Research Quarterly*, *Social Science Quarterly*, *American Journal of Political Science*, and *Polity*—as well as *Comparative Politics* and *World Politics*. From these data, we selected two highly prolific publishers during the 1981–92 period: Lee Sigelman, with 42 publications, and Susan Welch, with 28 publications. In the NRC publications file, Professor Sigelman is credited with only 5 publications; those produced after he arrived at George Washington University in 1991. His 25 publications while he was at Kentucky in

publications and citations only between 1988 and 1992. However, truncating the reported time period brings its own problems since publications seldom flow from scholars at a constant rate each year, and a lengthier time period may be necessary to validly estimate publication productivity. To return to the beginning, one can certainly not fault the NRC's desire to produce a data set as free from error as possible. Researchers must be aware, however, that the cleaning process used means the data in the written report the report measures neither publications from an individual scholar or from a program's faculty, and it is hard to tell exactly what it does measure.<sup>3</sup>

### Objective Data: Data Errors

With our data available for comparison, we explored the accuracy of the NRC's publication file. We selected 116 scholars,<sup>4</sup> followed the NRC's practice of including only publications produced while at their 1992 institution, and were left with 684 publications in our eight-journal data file that could be compared with NRC entries for the same scholars and journals. We found a number of differences and proceeded to track them down.

First, there were some occasional extra entries in the NRC file. Upon looking them up, we found the NRC file included rejoinders as separate articles. Our data included articles and research notes only, neither comments nor rejoinders. However, the added rejoinders were few in number, not producing sizable differences between the two data sets.

More worrisome, 156 of the 684 entries in our file (22.8%) did not have corresponding entries in the NRC file.<sup>5</sup> We first thought we had resolved the disparity, for the NRC file seemed to exclude research notes, but, upon continuing the comparison, we found the NRC's treatment of research notes to be schizophrenic. For example, only three of the *APSR*'s 15 research notes appeared in the NRC file, but over one-third (20 of 57) of the *JOP*'s research notes were included, while eight of *Polity*'s 15 research notes

**Table 2**

#### Percent of Evaluations from Raters not Familiar with Faculty or Program

Faculty Quality	Faculty Rank	Program Quality	Program Change
Top Third	1.1%	6.4%	0.0%
Middle Third	7.2%	22.2%	25.2%
Bottom Third	29.7%	40.3%	39.3%

file: "In the matching process if a paper was published in one year and the author moved to a new institution the following year, the paper would not be credited to the faculty member, since the former institution would not have listed the individual and the location of the latter institution would not match with the data in the ISI file" (NRC n.d., 1). This description suggests there might be some occasional one-year gaps in the data as faculty move from institution to institution, but, in fact, the NRC had available only faculty names and locations provided by the ICs for 1992. Thus, the publications information lists neither all publications of a single scholar, nor all publications emanating from the faculty in a particular department, during this time period. It instead contains only publications of faculty while

the early 80s, and his 12 publications when he was a dean at Arizona in the late 80s are not counted either for Professor Sigelman or for those two institutions since the author ZIP codes for those articles did not match Professor Sigelman's current location. Similarly for Professor Welch. The NRC publications file contains only five publications credited to her, those produced after she accepted a deanship at Penn State in 1992. Her earlier 21 publications are lost to both Professor Welch and the University of Nebraska, her prior institution. Further, the calculated "losses" are based on only the eight journals in our data file, though each scholar published in a variety of other outlets as well.<sup>2</sup>

The extent of this problem was mitigated somewhat in the published report in which the NRC reported

were counted by the NRC. This inconsistent treatment makes interpreting the NRC data problematic for it is neither fish nor fowl—neither a measure of all publications nor a count of regular articles only. The exclusion of research notes was a major source of differences between the two data sets, with 91 of our 131 research notes not counted by the NRC.

Looking at regular articles only, we counted 65 of our 553 articles (11.8%) not included in the NRC file. Part of this disparity is due to errors in transcription by either the ICs or the NRC, which left a handful of authors without any publications. Three examples are: Mark Roelofs (NYU), who is included in the NRC file as *M.* Mark Roelofs, instead of *H.* Mark Roelofs; Michael Lewis-Beck (Iowa), who is included in the NRC file as *Lewis-Back*; and Clyde Wilcox (Georgetown) who is in the NRC file as *W.* Clyde Wilcox. Because, in each case, the name in the NRC file did not match the name on a publication, these three prolific scholars were credited with zero publications during the 1981–92 period. However, these misidentifications only account for 14 of the missing 65 articles. Other missing entries seemed to occasionally affect particular individuals, such as Richard Niemi (Rochester), who was credited with only 5 of the 13 articles in our data file, and James Enelow

(Stony Brook), who only had 4 of the 11 articles we had for him included in the NRC file. For the most part, however, most missing entries were scattered throughout the list.<sup>6</sup>

Measurement errors, to the extent they are random, will not necessarily subvert analysis, for their primary impact should attenuate, but not necessarily bias, any calculated correlations. Using individual publication records as the unit of analysis should be avoided since errors here can be quite severe, but larger aggregations, such as entire departments or all publishers in a particular journal, might be reasonable. Even here, analysts should be cautious in drawing inferences, for the sum total of errors and coding decisions can be sizable. Expanding from our sample of 116 scholars, we compared all entries in our data set from the institutions examined by the NRC to the entries recorded in the NRC file. As shown in Table 3, numerous articles in various journals are not included in the NRC data set, from a low of 37% for the *APSR* to nearly 47% for the *JOP*.<sup>7</sup>

Considerably more troublesome than missing entries were a number of publications in the NRC file that did not belong there. We are not talking here about the rejoinders mentioned previously, but instead a number of publications in hard science, engineering, or medical jour-

nals, such as *Journal of Acoustics and Sound*, *Plant Physiology*, *Journal of Fluid Mechanics*, *American Journal of Nephrology*, *Journal of Biological Chemistry*, *Molecular Cell Endocrinology*, *Gynecological Oncology*, *Journal of Vascular Surgery*, and *Journal of Urology*.<sup>8</sup> Now, political scientists are a highly eclectic group who might well team up with a biologist, chemist, or physician to write on health policy, bioethics, or the like, but the number of entries similar to those described above raise doubts about this as an explanation.

To do a thorough check of the accuracy of these entries would require something we do not have; complete publication records for every political scientist included in the data set. The best we could do was perform some spot-checking. We took the names and university affiliations of about 75 scholars listed as authors of articles in the questionable journals and searched department and faculty web sites. We found about 50 individuals who had at least brief biographies on their web sites that included research interests and selected publications, including four individuals whose web sites contained extensive *vitae*. We looked to see if there was anything in the individuals' research interests or samples of publications to indicate whether the journal we questioned might be a reasonable publication outlet for them. We found only four such possibilities.<sup>9</sup> In most cases, the biographical information made publication in the questionable journals highly doubtful, such as an individual interested in formal models of international interactions listed by the NRC as publishing in the *Journal of Neurochemistry* and *Experimental Eye Research*. We conclude that most of the "hard science" journal entries are errors.

Some examples from cases where the evidence is quite clear should help demonstrate our confidence in this conclusion.

Professor Francine Frankel (Pennsylvania) is interested in political economy, international security, and South Asia. Nothing on her web site profile, which includes sample publications from 1988

**Table 3**  
Articles and Research Notes Missing from the NRC Publications File, 1981–1992

Journal	Our Count <sup>a</sup>	NRC Count	Percent Missing
<i>American Journal of Political Science</i>	462	281	39.2%
<i>American Political Science Review</i>	535	337	37.0%
<i>Comparative Politics</i>	123	73	40.7%
<i>Journal of Politics</i>	504	269	46.6%
<i>Political Research Quarterly</i>	336	190	43.5%
<i>Polity</i>	180	97	46.1%
<i>Social Science Quarterly</i>	218	129	40.8%
<i>World Politics</i>	119	65	45.4%

Notes: This comparison includes articles absent for any reason: counting errors, misspelled names, deletion of research notes, or deletion of articles because the author moved to a different institution.

<sup>a</sup>This count includes only articles from faculty at graduate institutions included in the NRC Study.

through 1995, would lead us to expect her to publish, as the NRC shows her doing, in *Molecular Cell Endocrinology*, *Journal of Steroid Biochemistry*, or *American Journal of Pathology*.

Professor Stephen Cohen (Princeton) has a complete *vita* attached to his web page. That *vita* does not include NRC-credited publications in *The Journal of Biological Chemistry* or *Geophysical Review*.

Professor Mark Schneider (Stony Brook) also has a complete *vita* online. The NRC entries include most of the items on Professor Schneider's *vita*, but also credits him with works in *Journal of Neurochemistry*, *Journal of Biological Chemistry*, and *Health Psychology*.

Professor Susan Welch has an abbreviated *vita* online, covering the years from 1992 to 1996. Her *vita* does not include the NRC entry for *Environmental Entomology*. The NRC also credits Welch with an economic journal publication, and the ZIP code of the author apparently matches that of Penn State, but the article was published in 1988 when Professor Welch was still at Nebraska.

Professor Sung-Chull Lee (UC-Irvine) has a number of items in the NRC file from 1988 back to 1986, in journals such as *Developments in Brain Research*, *Medical Physics*, *Clinical Electrophysiology*, and *Neuropsychology*. However, Professor Lee's interests are mathematical modeling and East Asian politics, and prior to 1988 he was still completing graduate studies at Kansas and would have been unlikely to submit articles with an Irvine ZIP code.

Most striking, perhaps, is Professor Gary King of Harvard. Professor King also has a complete *vita* online stretching back to 1986. Of the 28 publications credited to him in the NRC database, only 12 match those on his *vita*. The 16 items in the NRC file but not on his *vita* include publications in *Endocrinology*, *Metabolism*, *Journal of Cell Biology*, and *American Journal of Physiology*. Seven of the 16 doubtful entries, though having a Harvard ZIP code, were published while Professor King was still on the faculty at NYU.

It is easy to guess at the probable source for these errors. There were

two individuals with similar names publishing at a particular institution or in a similar geographical area between 1981 and 1992.<sup>10</sup> One was at a department included in the NRC study and the other was either with a program not examined by the NRC (medical, law, and business schools, for example, were not part of the study), or had left the institution/region sometime before 1992. Thus, the NRC's data cleaning protocol found one and only one name/zip code match between the ISI data base and names submitted by the ICs, and attributed all publications to that one person.<sup>11</sup>

To estimate the quantity of errors introduced here, we divided all the journals in the NRC data file into three groups. There was a "highly unlikely" group (138 journals, 460 publications), consisting of journals similar to those mentioned earlier in this narrative. There was a "maybe" group (43 journals, 98 publications), consisting of journals where publications by political scientists were unlikely, but still plausible. Examples of journals in this group were a number of nursing journals (e.g., *Journal of Advanced Nursing*), some highly specialized psychology/psychiatry journals where we doubted even political psychologists would publish (e.g., *Psychotherapy and Treatment*, or *Journal of Pediatric Psychology*), and a smattering of others (e.g., *Journal of Modern Greek*). Finally, there was the "reasonable" category (868 journals, 7,081 publications), including all social science, humanities, business, math, and law journals, along with any journals whose title mentioned policy, organizational management, or had a multidisciplinary flavor. We assumed that 90% of the "highly unlikely" group constituted errors, and so did 50% of the "maybe" group. We accepted all other entries as accurate.

From this rough approximation, we estimated that 6.1% of the NRC entries credited to political scientists actually belonged to individuals in other disciplines. This should be a conservative estimate given that most entries fell into the "reasonable" category which we accepted as completely accurate even though some are surely not (recall the eco-

nomics journal example concerning Susan Welch). Now, 6.1% is not an overwhelming number, but add these errors of commission to the errors of omission described earlier, and the reliability of the NRC publications file continues to crumble.

More troubling is that these erroneous article attributions have an impact beyond their numbers. They are not scattered randomly throughout the file, but instead tend to group around particular individuals; those credited with the publications of others tend to be credited with a lot of them. As shown earlier, 57% of the publications credited to Gary King were actually written by someone else.

The greatest impact is on the citation patterns. Citation practices differ across fields. Excepting psychology, scholars in the social sciences tend to cite each other rather sparingly. In some other fields, the rule seems to be: If it has been published, cite it—repeatedly. The average number of citations per political science and sociology article is 3.54 and 3.40 respectively. In contrast, the average number of citations to articles in medical science and biology/chemistry journals is 16.22 and 17.51 respectively.<sup>12</sup> Consequently, the rather modest error rate of 6.1% for erroneous publication attributions produces an error rate for citations recorded in the NRC study of 19.2%. Two examples will demonstrate the severity of this impact. First, we return again to Gary King of Harvard. In its 1995 publication, the NRC reported, along with its reputational measures, the total number of citations garnered by a program's faculty between 1988 and 1992.<sup>13</sup> For Harvard, this number was 592. However, if we add up the number of citations for the erroneous articles published between 1988 and 1992 that were attributed to Professor King, they total 143, or nearly 25% of the total citations attributed to Harvard's political science program.<sup>14</sup> More severe is the case of UC-Irvine, where the erroneous publications attributed to Professor Sung-Chull Lee constitute 68% of all the citations recorded for that university's political science program.

Both King and Lee are competent

and recognized scholars who have notable research records, even without additions from NRC mistakes. Their programs are both well respected and sizable; Harvard has 48 political science faculty members and UC-Irvine has a closer-to-average faculty of 23. Yet, citations to Harvard's large faculty were severely distorted by errors associated with only one individual, and the distortion was even more severe for the more modest-sized UC-Irvine. Somewhat smaller programs, such as Rochester or SUNY-Stony Brook (18 faculty members each), or almost any program in the bottom third of the rankings, where faculty sizes seldom exceed 20, are clearly susceptible to substantial measurement error in the citation data. Given the severity of the possible distortions, the citation data are probably far too unreliable to permit reasonable analysis.

## Conclusions

No data generation effort is problem free, and the NRC should not be held to such a standard. For continuous research projects, errors are commonly known and described by either data creators or later researchers so scholars can compensate for them. With the American National Election Study series, for example, researchers are aware that the institutionalized tend to be under-sampled and respondents tend to over-report their voting behavior. The NRC studies are close to comprising a series, and our goal is to refine future analyses, not discourage them. However, we cannot completely avoid doing the latter.

While some may challenge the validity of reputational data, we did not delve into that issue here, nor do any of our findings directly challenge the variable garnering most attention in the NRC report; the reputational evaluations of program faculty. We do, however, see the

reputational evaluations of program quality as tainted with serious measurement error, and deem the reputational assessments of program change to be often meaningless.

Scholars want to go beyond the mere reporting of poll results to understand the interplay of forces that produce reputations. This requires measuring independent variables validly and reliably, and here we have grave reservations. The coding quirks and errors in the publication and citation data raise questions about what, if anything, is being measured by these data, and whether any conclusions drawn based upon them must now be considered suspect, through no fault of the authors. Consider, for example, the symposium articles presented in *PS* (Katz and Eagles 1996; Jackman and Siverson 1996; Lowry and Silver 1996) along with a summary of the NRC report. Each was seeking to identify sources of rater subjective evaluations, and included the NRC's publication data, citation data, or both, among their predictor variables. That is, one of the theoretical linkages being examined was the match between rater evaluations and objective measures of scholarly performance. Unfortunately, the NRC data don't operationalize the theory well enough to provide an appropriate empirical test.

In making their subjective assessments, raters were given only lists of each program's faculty; no information on their publications or citations. Raters thus dredged out of their own memories information about the scholarly performance of the individuals listed. In doing this, raters surely did not forget all the publications of a listed individual that were produced before she joined this particular department, nor did they sift through their memories to be sure and incorporate all comments and rejoinders while forgetting all research notes. Similarly, the raters recollections most definitely did not include either the er-

roneous hard science attributions or their quality evaluations (citation counts). The NRC data simply do not represent either the quality or quantity of scholarly performance as most of the raters (and we) conceptualized it. Further, given the likelihood of some collinearity between publication performance and other predictor variables (such as faculty size, percentage of full professors, and overall institution quality), unacceptable measures of the former could bias coefficients for the latter variables as well.

Scholars and policymakers will need to tread cautiously. The NRC touted the "objective" data in their report as compensation for any ephemeral character to the reputational rankings ("Grad School Rankings" 1995). If the objective data are themselves unreliable, then no such compensation exists.

Anyone who has ever tried to develop a publications data file can sympathize with the NRC. It is a phenomenally arduous and time-consuming task. Our own data file, limited to just eight political science journals, has been 18 months in preparation and we are still constantly cleaning the data to comb out occasional errors that are discovered. To put together a file covering multiple journals and disciplines is a Herculean task indeed. We applaud, not fault, the NRC for its efforts. However, in its projections for the future, the NRC report discussed adding new variables—possibly surveys of employer satisfaction with program graduates or using international scholars to prepare reputational evaluations. Perhaps these would be useful additions, but the NRC's goal of providing a fount of useful information can be better achieved by addressing some of the problems in the current data before adding further information that may be equally error-prone.

## Notes

1. These are the reputational evaluations of a program's improvement or decline. Change measures calculated by comparing 1982 scores with those from 1992 should be more reliable.

2. The NRC's citation information was tied to specific articles, not authors. The deletion of these articles from the database deleted any citations to them as well.

3. Miller, Tien, and Peebler (1996), for example, argued the best way to measure a program's faculty quality was to count all publications of scholars currently on a faculty, even when some were produced during an individual's time at a prior institution.

4. The scholars examined were selected to cover a variety of institutions. The list included: Alan Abramowitz, Paul Abramson, John Aldrich, David Austen-Smith, Robert Axelrod, Jeffrey Banks, Lawrence Baum, Paul Beck, Henry Bienen, Charles Bullock, Gregory Caldeira, James Campbell, Edward Carmines, William Claggett, Harold Clarke, Pamela Conover, Albert Cover, Gary Cox, Stephen Craig, Robert Dix, William Dixon, John Dryzek, James Enelow, Richard Engstrom, John Ferejohn, Steven Finkel, Morris Fiorina, Richard Fleisher, James Garand, Robert Grafstein, Bernard Grofman, John Gruhl, Timothy Hagle, Paul Hagner, John Hibbing, Jon Hurwitz, Ronald Inglehart, Helen Ingram, Robert Jackman, John Jackson, Malcolm Jewell, Kathleen Kemp, Patrick Kenney, John Kessel, Gary King, Tom Lauth, Michael Lewis-Beck, Milton Lodge, Nicholas Lovrich, David Lowery, Robert Luskin, Gregory Markus, Kenneth Meier, David Morgan, Edward Muller, Michael Munger, Max Neiman, David Nice, Richard Niemi, Helmut Norpoth, Karen O'Connor, Joe Oppenheimer, John Orbell, T. Wayne Parent, Mark Peffley, Stephen Percy, John Peters, John Pierce, G. Bingham Powell, Larry Preston, George Rabinowitz, Brian Roberts, Mark Roelofs, David Rohde, Shawn Rosenberg, C. K. Rowland, Robert Salisbury, Grace Saltzstein, Arlene Saxonhouse, Kay Schlozman, Mark Schneider, Donald Searing, Jeffrey Sedgwick, Jeffrey Segal, Robert Shapiro,

Elaine Sharp, Steven Shull, Lee Sigelman, Brian Silver, Barbara Sinclair, Randolph Siverson, Duncan Snidal, Paul Sniderman, Donald Songer, Priscilla Southwell, Harold Spaeth, Peverill Squire, Robert Stein, Charles Stewart, Walter Stone, Jeffrey Stonecash, John Sullivan, Neal Tate, Sue Thomas, Carole Uhlaner, Eric Uslaner, Kenneth Wald, Michael Wallerstein, Stephen Weatherford, Susan Welch, Frederick Whelan, Kenny Whitby, Clyde Wilcox, Rick Wilson, Gerald Wright, Thom Yantek, and John Zaller.

5. Unlike some ISI publications, the data file provided to the NRC presumably included all authors, not just first authors. Also, citations to an article were not a limiting factor since many articles in the NRC file had zero citations attributed to them.

6. An exception, which has been noted previously (Miller, Tien, and Peebler 1996), is the University of Houston, which is listed as having zero publications in the printed report. We were unable to explain this since Houston faculty were included in the NRC faculty file, and there were even some publications listed for them in the publications file, though an unreasonably small number.

7. The NRC publications file does include a variable identifying the type of publication, though what is distinguished here is regular paper versus proceedings paper versus poetry, etc. In the comparison of our data with the NRC's, we did find a number of articles in the NRC file with a "U" (proceedings paper) code, but we could find no rhyme nor reason why this code was applied to these articles. Every one we looked at was a normal journal article.

8. Journals are identified in the NRC file only by an 11-character abbreviation, the same abbreviation used in ISI publications. In about two-thirds of the cases, the full journal name was obvious. For the remaining cases, we matched the abbreviations against periodical listed in our library's database. However, journals in the ISI database are continually changing, and about 100 journals, constituting

some 280 entries, were not identifiable. They were treated as missing data for all calculations presented in this paper. We did not include as a hard science or medical science journal any whose name suggested it might deal with policy, such as health or environmental policy, or organizational concerns, such as hospital administration or public health service delivery.

9. For example, one individual had a brief *vita* showing he definitely did publish in *Medical Science*, and another had an interest in space policy that made publication in the *Journal of Atmospheric Science* plausible.

10. The geographical areas apparently do not need to be that close. Ken Meier moved during the time of this study from the University of Wisconsin at Madison to the campus at Milwaukee. This was apparently close enough that his publications at both institutions were included in the data file. This ZIP code laxity is surely necessary to avoid deleting the publications of faculty at urban institutions where their residence might be a considerable distance from campus, but it is likely to also introduce problems in these same urban areas that are likely to contain multiple research universities.

11. It is not clear how close the name match had to be, because the NRC report does not fully describe the data-coding process. At one point (NRC 1995, 143), the study description suggests the NRC had only last names plus initials, not full first and middle names, in the ISI data.

12. The research producing these citation averages is based on the data in the NRC citations/publications file, and is forthcoming elsewhere.

13. The NRC's 1988 to 1992 citation count is for items published during those years only. A citation in 1990 to an item published prior to 1988 would not be counted.

14. There are similar errors for other Harvard faculty, which leads to about half of Harvard's total citations in the printed report being probable errors.

## References

- "Departmental Rankings: Much Ado About Something?" 1996. *PS: Political Science and Politics* 28(June): 144-48.
- "Grad School Rankings Rankle." 1995. *Science* 269(September): 1660-62.
- Jackman, Robert W., and Randolph M. Siverson. 1996. "Rating the Rating: An Analysis of the National Research Council's Appraisal of Political Science Ph.D. Programs." *PS: Political Science and Politics* 28(June): 155-60.
- Katz, Richard, S., and Munroe Eagles. 1996. "Ranking Political Science Programs: A View from the Lower Half." *PS: Political Science and Politics* 28(June): 149-54.
- Lowery, Robert C., and Brian D. Silver. 1996. "A Rising Tide Lifts All Boats: Political Science Department Reputation and the Reputation of the University." *PS: Political Science and Politics* 28(June): 161-67.
- Miller, Arthur H., Charles Tien, and Andrew A. Peebler. 1996. "Department Rankings: An Alternative Approach." *PS: Political Science and Politics* 28(December): 704-17.
- National Research Council. 1995. *Research Doctorate Programs in the United States: Continuity and Change*. Washington DC: National Academy Press.
- . n.d. "Pub\_cit.doc" <computer file>. On *Research Doctorate Programs in the United States: Data Set* <CD-ROM>. Washington DC: National Academy Press.