

## **Online Appendix B to the PS 400**

### The Methodological Appendix

As noted in the text, there are various ways to measure reputation and professional visibility. Miller, Tien, and Peebler (1996: 73) assert that citation count data can provide “a useful and valid measure for determining standing in the profession.” We agree. We provide a roughly 15-year update of the citation data set in the Masuoka, Grofman, and Feld (2007) study, henceforth referred to as MGF. Like the MGF study, and its immediate predecessor, Klingeman, Campagna and Grofman (1989), our work is inspired by Somit and Tanenhaus (1967).

Like MGF we have restricted ourselves to those at political science departments in the U.S. that are Ph.D. granting. The two key restrictions here are, of course, limiting ourselves to only those faculty currently teaching in R1 departments, and limiting ourselves to faculty teaching in the United States. We know from personal knowledge that the first restriction eliminates some very highly cited faculty who are teaching at other types of units in the U.S. (e.g., Pippa Norris at the Kennedy School, or John Carey, at Dartmouth). The second restriction eliminates everyone teaching outside the U.S., and that is a huge restriction. Once upon a time, there were relatively few scholars with citations in English language journals or with English language books who taught outside the U.S. or other English speaking countries, and U.S. scholars had a marked edge in SSCI citations even compared to those in other English speaking countries. But the political science discipline worldwide has been experiencing immense changes and U.S. predominance is no longer something that can be taken for granted, as signaled most recently, by a change of editorial team of a major U.S. journal to a location in Germany --something which the senior author of this article views as having been inconceivable even as recently as a decade ago. Nonetheless,

all we can do is acknowledge these limitations in our data set. To actually try to create a data base for the 10,781 APSA members or even to extend our data set to Ph. D. departments worldwide, was a task well beyond us. (We would note, however, as a point of interest, that 20% of APSA's members are based outside the US.)

The third restriction, to those with primary appointments in political science, is also not that trivial. Consider, for example, scholars like Bruce Ackerman or Alec Stone Sweet, who are widely cited in political science, but who will be omitted from the data set because their primary appointment is in a Law School with a secondary appointment in political science. We would note, though, that a handful of faculty who fall into this category are being tallied as an Emeritus (e.g., John Ferejohn, now at NYU Law School, but listed as Emeritus at Stanford).

MGF looked at lifetime citations from 1960-2005, with data on faculty affiliations taken mostly from 2002, but broke down citations by other groupings, such as faculty cohort, as do we. However, unlike MGF, we have opted for use of Google Scholar data rather than the Social Science Citation Index data they made use of. We agree with Samuels (2013) that Google Scholar provides a better way than SSCI to include important citations to books, and not just to journal articles. While citations to books can (with considerable effort) be located in the SSCI data set, that data base is primarily oriented to identifying citations that are found in the 1000+ journals that it identifies as established journals in the social sciences. Since, many scholars have their books as their most highly cited works, despite some issues of over-inclusiveness in the Google Scholar counts as compared to a count consisting almost entirely of citations to articles found in the academic journals listed in SSCI, following Samuels (2013), we regard under-inclusiveness of book citation in MGF as the more important problem to correct. We have citation data from Google Scholar that spans over 50 years.

Our full data set on faculty at R1 universities has information on 4,089 faculty, of whom 3412 are currently employed in tenured or tenure track positions. Because we are missing the date of Ph.D for 320 scholars, for some analyses the N is 3769. We began with an advanced search by name, using university affiliation to deal with the potential problem of identical names, though in one case we needed to look at article topics to distinguish two scholars with the same name at the same institution. If the scholar has a Google Scholar Profile, the total number listed in it is what is used. If the scholar does not have a Google Scholar Profile, we added up citations by hand as these were shown on Google Scholar webpages. Of the 4089 faculty in our data, 1658 had Google Scholar profiles, while 2153 did not, and we could not find Google Scholar information for 278 faculty due to missing information, whom we omitted from the data base, leaving us with an N of 3,811 for much of the citation data analysis. In an Appendix available on-line we provide additional information on the characteristics of the faculty who had created Google Scholar Profiles. In addition we also excluded from some analyses, e.g., that by cohort, the 320 scholars for whom we were unable to track down date of Ph.D. The various kinds of missing data means that the n for the various tables, while similar, is not identical.

Like MGF we treat all citations as equal in weight to one another. In other words, we do not distinguish the prestige level of the citing source, or whether the citation is to a book or to a research note or to an article; and we treat all citations to a publication as equal regardless of whether or not the publication is co-authored (or co-edited), or how many co-authors there might be, or where in the list of co-authors the cited author's name appears. Given the size of the data set we are dealing with, and the amount of hand coding involved, it was simply not possible to implement any of these complications -- even were we to think them desirable. We have done our best to assure the accuracy of the data we report, but there are a number of factors making the task

difficult, such as hyphenated names and multiple name variants (e.g., use of middle names and middle initials, nicknames, and just initials), and changes in name. The problem is made more acute for faculty with common names and for those who moved universities, and those who publish in multiple subfields, or who have citations to work that is published in non-political science journals.

Thus, we have no doubt that there are some errors remaining. We should also note that the citation count data in the present study is primarily from the first four months of 2018. Because coding data on around 4000 faculty, primarily by going page by page by hand over website pages to count citations for the 56 per cent of scholars who did not have an official verified Google Scholar count (a Google Scholar Profile), while making sure that all citations were counted but that citations to of those with similar names were not inadvertently being counted, is remarkably time consuming, and cannot be delegated to a scraping algorithm, it required a full quarter to collect all the data. Because the data collection was not simultaneous, the relative rankings for those with very similar rank might be affected by when data was collected for each, so the ranks we report later in the paper should be regarded as approximate, not exact.

We broke faculty down by 5 year cohorts (e.g., 1970-1974), and by subfield of principal (using six subfield categories: American, Comparative, International Relations, Methodology, Political Theory, and a combined category of smaller cross-disciplinary fields: Public Policy, Public Administration, Public Law, and Political Psychology). Much of the data on fields of interest was taken from the previous MGF dataset, supplemented by information on faculty and university websites about main research interests/ subfields, or inferable from publications listed on Google Scholar or books listed online. In some cases, when a subfield identification listed in the MGF data set was suspect, we sent a follow-up email asking the scholar to identify their

primary field of interest. Our last resort was to look at the classes they had taught to get a general idea of their field, though this method was almost never used.

In addition, we collected data on gender. Gender is estimated on the basis of first names, with a look at university or personal websites in cases of uncertainty. If the name is gender neutral (e.g. Jaimie, Jean, Taylor, Rui, Robin, and Leslie), we looked up the faculty member's listed biography and used the third-person gender pronoun listed there to determine gender.

A limitation of the field of interest coding is that we are listing only a single category. Some scholars have multiple strong interests, For example, one of the authors of this essay had a very hard choice as to whether to put American politics or comparative politics as principal research area, and he previously had coded himself as a political theorist to reflect his strong interests in democratic theory. But he has also written extensively in the field of public law, especially with respect to race. Indeed, based on many other things he has written, including formal and statistical modeling, some of it work published in journals such as *Political Analysis*, and the fact that he regularly teaches introductory and intermediate level statistics, he could equally well call himself a methodologist, and that is the way we coded him. Thus he is listed as the 9<sup>th</sup> most cited among 129 methodologists, rather than, say, the 20<sup>th</sup> most cited among the 1001 living students of American politics, or the 25<sup>th</sup> most cited among the 827 students of comparative politics, or the 4<sup>th</sup> most cited political theorist, or the 5<sup>th</sup> most cited among those working in one of the applied interdisciplinary areas (including law). We should note, too, that we deliberately sought to purge the political theory category of those who would be called, a la Bill Riker's terminology, "positive political theorists," rather than falling into a more traditional interpretation of what constituted work in political theory. These scholars were coded either according to substantive interests, or as methodologists.

We also collected data on emeritus faculty at R1 departments. Anyone who became emeritus or emerita at one department while subsequently taking a tenured or tenure track position at another R1 political science department is listed as a faculty member at that second department rather than under the emeritus category. While most emeriti are no longer teaching anywhere, some emeriti have transferred out of U.S. political science R1 departments to other types of units within the U.S., e.g., Schools of Public Policy, Law Schools, interdisciplinary units, non-Ph.D. programs; or they now have appointments outside the U.S. In such cases we generally treat them as emeritus at their last R1 political science department. However, one scholar who would have been among the 25 most highly cited emeriti, but who transferred to an interdisciplinary unit prior to his emeritus status, preferred not to have his previous affiliation used, and his name is omitted from the list. There are a total of 677 emeriti in political science Ph.D.-granting institutions in the U.S. (for 77 of these, date of Ph.D. is missing years are missing, so we only have 600 emeriti for whom we have full data). The bulk of all emeriti, and of all emeriti who would have been in the PS 400 were it not for their emeritus status come from the period 1965-1974 (Tables omitted for space reasons). Emeriti who would have been in the PS400 are similar to all emeriti in other respects, e.g., in subfield (tables omitted for space reasons). Out of all 677 emeriti, only 76 are female, since the oldest cohorts are very heavily skewed toward male scholars. Out of those who would have been in the PS400 were they not emeriti, only five out of 69 are female. Out of the top 100 active emeriti, only 7 are female.

#### Aggregate Level Features of the Dataset

Table A1a shows the breakdown of our data set by rank, placing emeriti into a separate but still tallied category. We see that full professors are the largest category, at 39.7% of the entire faculty list at the R1 political science departments, and 47.6% of the non-emeriti.

**Table A1a. Breakdown of the Data by Rank**

Titles	Total Faculty	Share of Faculty
Assistant Professor	754	18.44
Associate Professor	1035	25.31
Professor	1623	39.69
Professor Emeritus	677	16.56
	4089	100

Table 1b shows the breakdown of our data set by rank and gender.

**Table A1b. Breakdown of the Data by Rank and Gender**

Titles	Women Faculty as Share of Faculty at that Rank
Assistant Professor	40.98 (309/754)
Associate Professor	35.56 (368/1035)
Professor	23.54 (382/1623)
Emeritus	11.23 (76/677)

Comparing this table to Table A1a, we see that women are more highly represented at the assistant professor ranks and associate professor ranks than they are at the full professor rank. If most associates become full professors, as is suggested to be likely to be true for R1 institutions, the present gender imbalance will be considerably reduced, but still not eliminated. Some of this imbalance reflect the over-time changes in the proportions of men and women receiving Ph.Ds in political science.

Table A1c shows the breakdown of our data set by subfield, but here we have omitted emeritus faculty to better reflect the current state of the discipline. As such, it includes a total of 3412 faculty members.

**Table A1c. Breakdown of the R1 Data by Subfield (excluding emeriti)**

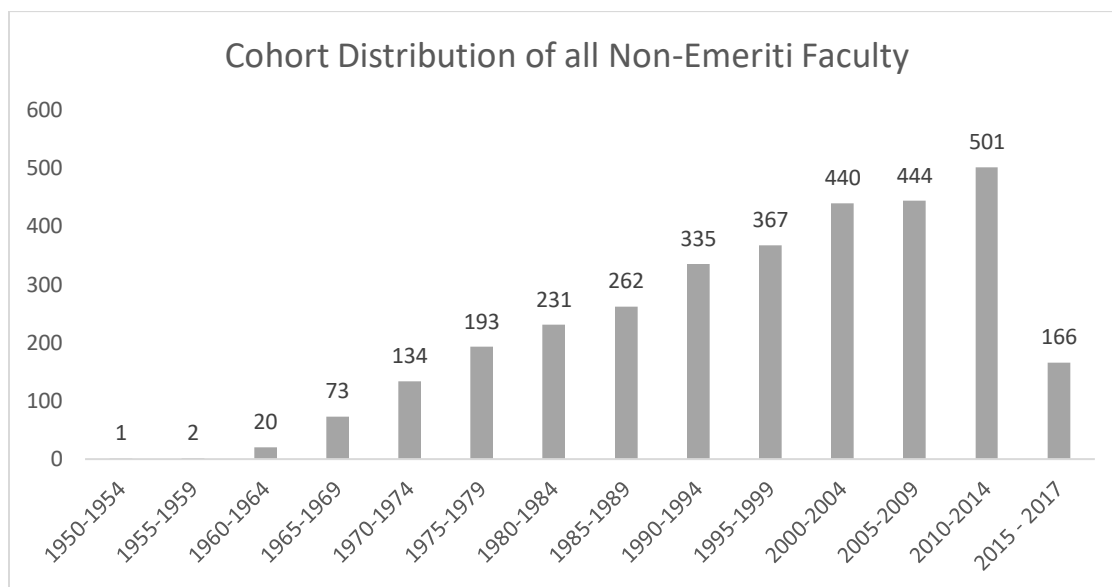
Field	Faculty	Share of Faculty
American	976	28.60
Comparative	827	24.24

International Relations	778	22.80
Methods	129	3.78
Theory	346	10.14
Public Policy	180	5.28
Public Administration	58	1.70
Public Law	56	1.64
Political Psychology	36	1.06
Race and Ethnicity	26	0.76

We see from this data that the majority of faculty are in American and Comparative Politics. Out of the major subfields, the subfield of methods remains the smallest with 3.78%. Smaller subfields such as Public Administration and Public Law only include 1-2% of the faculty population.

Figure A1 shows the breakdown of our data set by 5 Year Cohort identified by date of Ph.D.

**Figure A1. Cohort Distribution of All Non-Emeriti Faculty**



We see from this data that a high proportion of all R1 faculty come from recent cohorts.



Figure A2 provides the distribution of citation counts for the non-emeritus faculty in the full data set.

**Figure A2 – Distribution of Cumulative Citations for Non-Emeritus Faculty at U.S. Ph.D. Granting Departments**

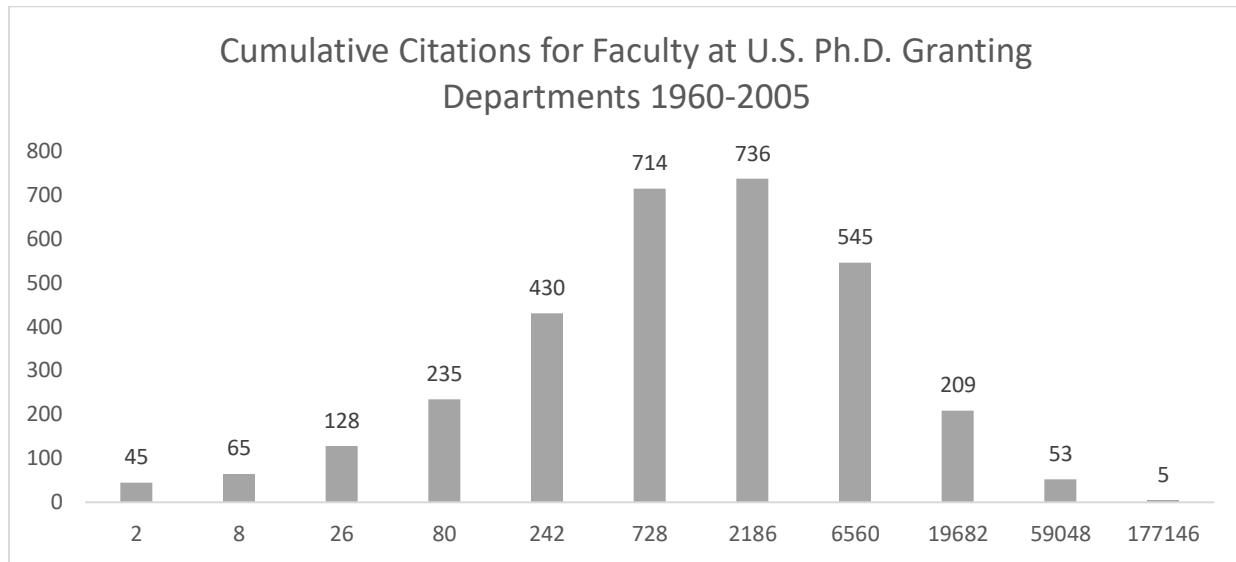


Figure A2 shows that the overall citation data has a nearly lognormal distribution.<sup>1</sup> Put simply, those with high citation counts have exponentially ( $\log 3$ ) more citations than those with few citation counts. Moreover, while there are a number of faculty who have very few citations, the majority of non-emeriti faculty have a citation count between 728 and 2186.<sup>2</sup> One part of the difference is due to cohort effects, since recently minted Ph.D.s are likely to have fewer citations.

#### Aggregate Level Features of the PS 400

Figure A3 breaks shows the proportion of scholars the political science 400 who come from each of the 5-year cohorts. As we would expect, this is greatly skewed toward older cohorts,

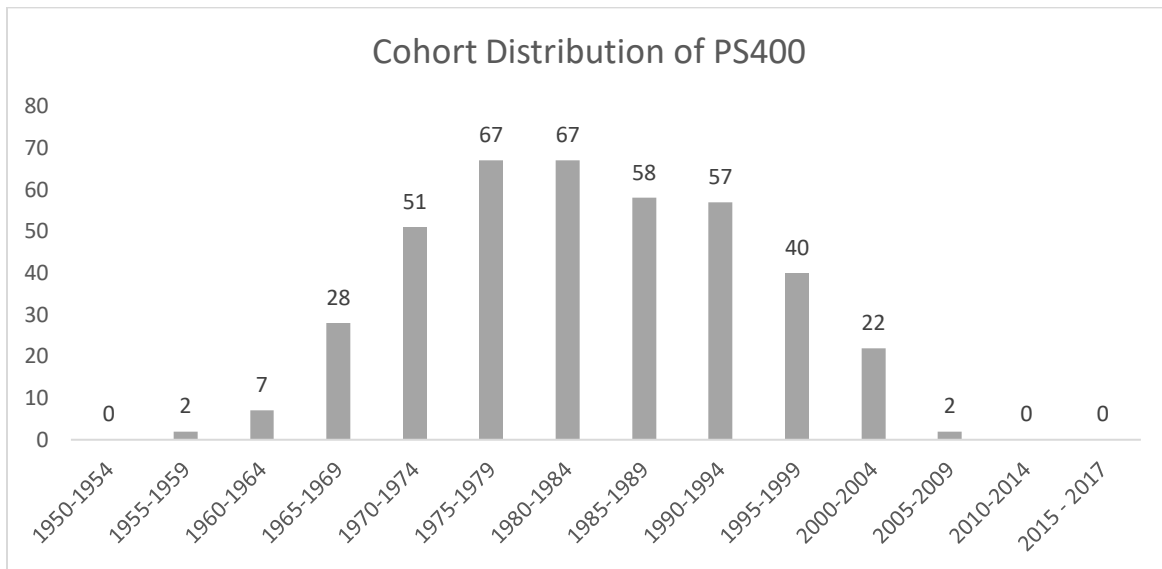
---

<sup>1</sup> MGF found a similar result for SSCI data, but there the exponent was 2.

<sup>2</sup> The range is 0 to 175198.

as compared to the overall data presented in Figure A1. Here we would note the continued prominence of scholars who got their Ph.D.s in the 1970s.

**Figure A3. Cohort Distribution of Non-Emeriti Political Science 400**



Comparing this distribution to that for the full data set, the skew toward older cohorts is obvious – which is why we also presented data sorted by cohort.