Citation versus Reputation:
Assessing Political Science Journals

Will H. Moore
Department of Political Science
The Florida State University
Tallahassee, FL 32306-2230
whmoore@garnet.acns.fsu.edu

July 30, 2000

Draft: comments appreciated.

I would like to thank Mitch Sanders for helpful discussions about this project, and Sara Mitchell for helpful comments on an earlier draft of the paper.

Abstract

This study tackles the problem of assessing the forhcoming and recently published work of one's colleagues. Arguing that this is best conceived as a forecasting problem, I contend that the most useful concept to measure for forescasting the likely quality of a journal article is the quality of the journal in which it is (to be) published. I then distinguish between direct and indirect measures of journal quality: citation counts and reputation. I evaluate two measure of each type, arguing that (direct) citation count measures are perferable to (indirect) reputational measures. I also compare a new cumulative 10 year citation count per article measure with the Institute for Scientific Information's impact factor measure of citation counts, and conclude that the new measure is better for political scientists.

1 Introduction

This study addresses a situation all political science faculty face during their careers: how should one make comparative judgments about one's colleagues' work? It was motivated by my own experience on an annual merit raise committee that had been charged with judging the recent research, teaching, and service accomplishments of our faculty. More specifically, I embarked on this analysis in response to a debate that ensued when a colleague asked me to share my evaluation of a specific field journal with other journals. By way of response I argued that the most sensible way to do so was to rate the journal in question relative to six field journals: the *American Political Science Review*, the *American Journal of Political Science*, the *Journal of Politics*, *Polity*, *Political Research Quarterly*, and the *Southeastern Political Science Review*. This claim (along with my own ranking of the six journals) sparked the type of discussion that is the bane of the chair of such a committee.

The various positions espoused around the table are surely familiar to anyone who has taken part in such a discussion: some agreed with me that journals can be compared with one another across subfields; others objected that while comparison within subfield was valid, one could not compare across subfields; still others expressed unease with comparing subfield journals with general journals; and the view that surveys of a sample of the discipline, even if dated, were superior to the rankings of the considerably smaller sample formed by the committee was also aired (e.g., Giles, Mizell and Patterson 1989).

Can one make valid comparisons across journals in an effort to evaluate the work of political scientists? While these types of questions are generally addressed in the literature, Norris and Crewe (1993) use the results from a survey of British political scientists to evaluate the extent to which comparisons can be made across journals. Identifying two categories, which they label the consensus and pluralist views, Norris and Crewe (1993:22) conclude that their survey "reveals a broad consensus in the profession about the overall quality of journals" in political science. Their consensus view endorses the conclusion that "it is

therefore legitimate to use the quality of a journal as a proxy measure of the quality of the articles within its covers" (p. 7).

I thus accept Norris and Crewe's consensus view that relative comparisons of all political science journals are both necessary and valuable. Yet, little information one might use for this puprose is available. Michael Giles (Giles and Wright 1975; Giles, et al. 1989) has provided the discipline with two reputational surveys for this purpose. In addition, Christenson and Sigelman (1985) proposed a citation-based measure as a superior alternative to reputational surveys. The purpose of this study is to further explore the debate opened by Christenson and Sigelman: Are citation based measures more useful for the purpose of evaluating one another's work than reputational measures?

The study proceeds in four sections. In the next section I discuss the specific problems raised by the need to evaluate recently published and forthcoming articles and define concepts that are useful to that task. In section three I describe some operational measures of these concepts, and present rankings based on the different measures. In section four I report the findings from my analysis of some regressions. The purpose of these analyses is to determine the relationships among the measures and, following Christenson and Sigelman, to identify the extent to which each journal's reputation exceeds or falls shy of its citation count. In the conclusion I consider the uses to which we might usefully put this information.

2 Evaluating Recent Work

A useful distinction can be drawn between the problem posed when (1) evaluating articles[1] that were published several years ago and (2) when evaluating articles that are either forthcoming or have only recently been published. In the former circumstance one can both survey scholars in the field, soliciting their opinions of the work and one can determine the

[1]Because I have no new information about the reputations or citations accumulated by the presses that publish books, I focus on articles in this study. The most recent ranking of presses with which I am familiar is Goodson, Dillman and Hira (1999).

number of citations to the article. Indeed, when making tenure and promotion decisions, where one evaluates an entire body of work, both practices are standard operating procedure in most departments with which I am familiar (and, I strongly suspect, most PhD granting departments). However, in the latter circumstance, neither of these pieces of information are available. In effect, we are put in the position of having to make a forecast of the impact of forthcoming or recently published articles. This raises the question of what information gives one the best forecast. The best focal point is the quality of the journal in which the article is (to be) published. Yet, it is not obvious how one ought to determine the quality of the journal: several options present themselves, and I discuss each in turn.

If the problem is to forecast the quality of an article by using the quality of the journal in which it is published as a proxy, then we need to determine how we might assess the quality of a journal. The quality of a journal might be defined as the value scholars place on that journal when conducting their own research. That is, quality reflects the extent to which scholars have used articles published in the journal when conducting their own work, or, put differently, the extent to which the scholarship in the journal has influenced future scholarship on the topic.

Having thus defined quality, we can turn our attention to how it might best be measured. A major distinction can be made between indirect and direct measures. Define reputation as the beliefs of the community of scholars about the quality of a journal. A reputational measure (e.g., Giles et al. 1989, Garand 1990) is an indirect measure of journal quality because rather than directly determining the value scholars place on a journal when conducting their own work, one asks determines scholars beliefs about what the value they place on different journals. In contrast, a direct measure of quality would endeavor to assess the value scholars place on a journal without seeking their opinions. The best known method for directly measuring journal quality is to count the number of citations to the journal in question (e.g., Christenson and Sigelman 1985).[2] Thus, two types of indciators present themselves

[2]For a parallel discussion that focuses on how knowledge is accredited, see Christenson and Sigelman

as measures of journal quality: reputational (indirect) measures and citation count (direct) measures.

A few political scientists have debated the relative value of direct and indirect measures. For example, Christenson and Sigleman (1985) cite Cole's (1983) conclusion that citation counts are a valid measure of the quality of scholarship, argue that reputations are influenced by irrelevant factors and resistant to change, and present suggestive evidence in support of their contentions. They conclude that, despite some shortcomings, "impact data [i.e., citation counts] seem to us to provide a firmer foundation for assessing the quality of sociology and political science journals than any other method devised to this point" (p. 973). Lester (1990) is critical of both approaches, and advocates a systems approach that takes into account input measures (e.g., the number of articles submitted to the journal), decisional measures (e.g., characteristics of the review process and the lag from acceptance to publication), and outcome measures (e.g., circulation and the number of articles published in each volume). Finally, Norris and Crewe (1993:22) briefly raise the issue of reputational versus citation measures, but conclude that the debate remains open "for discussion by the profession as a whole."

The present study cannot settle this issue, but like Christenson and Sigleman's study, it does offer some suggestive evidence to support the view that direct measures are superior to indirect ones, but some readers may prefer to side with Lester and conclude that both types of measures are important. In any case, the study can contribute to the professional discussion Norris and Crewe envisioned.

3 Some Operational Indicators and Rankings

In this section I describe four measures of journal quality: reputation (Giles and colleagues), reputation weighted by readership (Garand 1990), the Institute for Science Information's

(1985:964-5).

(ISI) Impact Factor citation score (Garfield 1972), and a cumulative citations per article, 1990-2000, measure that I am introducing. After reviewing how each of these indicators is created, I rank the journals by the cumulative citations per article, 1990-2000, measure. I then compare the measures, examining bivariate correlations and our ability to make distinctions among the journals. I contend that the (direct) citation measures are superior measures of quality to the (indirect) reputational measures, noting that our ability to distinguish among broad ranges of values in the reputational scores is very limited. I also argue that the cumulative citations per article, 1990-2000, measure is superior to ISI's impact factor score because it captures a larger portion of the citations likely to accrue to an article.

Michael Giles and his colleagues have defined the standard for a reputational ranking of political science journals (Giles and Wright 1975, Giles et al. 1989). Their mail survey is conducted by drawing a random sample of American political scientists as listed in the *APSA Guide to Graduate Study* and sending them instructions to rate a specified set of "journals in terms of the general quality of its articles on a scale from 0 to 10 with 0=poor, 2=fair, 4=adequate, 6=good, 8=very good, and 10=outstanding" (Giles et al, 1989:613). They instructed respondents to rank only those journals with which they were familiar. That decision not only enhances the validity of the ranking by excluding uninformed opinions, but also provides information about the cohesion/fragmentation of the discipline[3]

James Garand (1990) argued that the Giles et al. survey lacked face validity: he identified a handful of journals that were highly rated, but were familiar to a very small percentage of the respondents. Arguing that it is important to not only determine the reputational quality of a journal, but the number of people likely to read the journal, Garand proposed a new measure which takes account of readership: reputation + (reputation*familiarity). Because both pieces of information were published in Giles et al., Garand was able to calculate his new measure and compare the two. Most journals retain their relative standing using the

[3]These issues are discussed in Crew and Norris 1991 and Norris and Crewe 1993. The mean percentage of respondents familiar with a journal in the most recent Giles and colleagues survey is 37, with a standard deviation of 24.

Garand measure, but several move, some quite a bit.[4]

A direct measure of journal quality was proposed by Christenson and Sigelman (1985) who use the Institute for Science Information's (ISI) Journal Citation Reports (JCR) 'Impact Factor' to operationalize citations to a journal. ISI publishes an annual JCR analysis of its Social Science Citation Index (SSCI) database. The impact factor is

> a measure of the frequency with which the 'average article' in a journal has been cited in a particular year.... the impact factor of Journal X would be calculated by dividing the number of all current citations of source items published in journal X during the previous two years by the number of articles journal X published in those two years (Garfield 1991:9).

Thus, the 1990 impact factor for the *APSR* is the number of citations found in the 1990 issues of journals in the ISI's SSCI database to articles published in the 1988 and 1989 issues of the *APSR*, divided by the total number of articles published in the 1988 and 1989 issues of the *APSR*.[5]

While the impact factor is a useful measure, it is not without its critics (see Porta, 1996; Harter and Nisonger, 1997; Rousseau and Van Hooydonk, 1998; Garfield, 1998; and Moed, Van Leeuwen and Reedijk, 1999). In addition to the limitations imposed by any finite database (particularly that not all journals, and no books, are in the ISI database; see Christenson and Sigelman, 1985:972-3), I contend that the ISI's impact factor uses a subomptimal time frame for our discipline. Prior to becoming aware of the existence of the JCR and the impact factor, I realized that I could access ISI's SSCI database via the webofscience.com web site and calculate the total citations to date to a given year's issue of a given journal. This impressed me as a useful direct measure of a journal's quality, and I embarked on a project to collect such data for the journals listed in the Giles et al. (1989)

[4]Giles et al. report that the correlation between reputation ('evaluation' in their study and Garand's) and familiarity is .39. Thus, it is not surprising that the rankings are similar for most journals.

[5]For a more detailed discussion, see Garfield (1972) or visit the ISI web site: http://www.isinet.com

survey. I selected the year 1990. That decision hinged on three considerations: I wanted to let enough time elapse that articles would have accumulated an appreciable number of citations;[6] I did not want the time elapsed to be so long as to make the results meaningless for the present; and I wanted to be able to compare the results to the Giles et al. survey. The year 1990 maximizes this set of criteria.

Of course, as a measure of quality the total number of citations is biased in favor of journals that publish more articles than the average journal. I thus used the CARL Uncover database to determine how many articles were published in 1990 in each of the journals on which I collected data. I then divided the total number of citations to articles in the 1990 issue of the journal by the number of articles published in the 1990 issue to arrive at the average number of cumulative citations, 1990-2000, per article to the 1990 issue of the journal. Table 1 lists the journals ranked by my calculation of citations per article. It also lists the SSCI's impact factor, Giles' et al.'s reputation score, and Garand's impact score.[7]

(Insert Table 1 about here)

A question that immediately comes to mind is: What relationships exist among these variables? Table 2 reports the bivariate correlations between the four measures, and introduces several others recommended by Lester (1990), each of which can be found in Martin and Goehlhert (1997):[8] the number of manuscripts reviewed annually, the acceptance rate, the average number of months it takes to review a submission, the percentage of articles that are solicited, the lag from acceptance to publication, the circulation of the journal, and the year it was founded. Table 3 displays the descriptive statistics of the various measures.

[6]I calculated the scores for 1989, 1990, 1994, and 1997 for the *APSR*, *AJPS*, and *JOP*. The results showed that the variance collapses as one reduces the elapsed time. These results are available on panel d of the QuattroPro (and Lotus 123) spreadsheet containing all of the data described in this study. It is available at the Research page on my web site: `http://garnet.acns.fsu.edu/~whmoore/research.html`.

[7]Readers may be interested in seeing this table in a spreadsheet so that they can resort the data along a different variable. It is available at the web site described in the preceding note.

[8]The 1997 volume of the Martin and Goehlhert data is less useful for this study than the 1990 volume. I have requested a copy of the 1990 volume through interlibrary loan, and will use that data once I have received it.

(Insert Tables 2 and 3 about here)

Table 2 reveals that the reputational and citation measures are the most highly correlated: Lester's suggested measures exhibit low correlations with the reputational and citation measures, and also exhibit fairly low correlations with one another. The two reputational measures have a bivariate correlation of .77 whereas the two citation measures have a bivariate correlation of .82. Thus, both pairs are similar. Across pairs, the table reveals that the ISI impact factor is more strongly related to the reputational measures (.71 and .61) than is my citations per article measure (.54 and .44). One might interpret this finding as suggesting that the ISI's impact factor is a more valid measure of journal quality as it shares more in common with the reputational measures. While I would concede that it is a more valid measure of reputation, I contend that for political science my measure is superior to the ISI's impact factor journals, but I make my case by appeal to other conceptual and empirical criteria.

The validity conundrum should not be resolved by noting the most similar correlations among a set of measures. Doing so cannot avoid the potential pitfall of foresaking a superior measure in favor of a set of inferior ones. A better idea is to turn to conceptual grounds and consider what it is we are trying to measure. One issue that I have not yet raised is the length of time it takes for a typical article to secure its citations. The ISI's *JCR* annuals (e.g., Garfield 1991) report the half-life of the citations for each journal in their database.[9] While I did not analyze the half-life data, casual observation was consistent with my prior belief that few articles in political science receive a substantial portion of their eventual total citations in the first two years from publication. Yet, this is precisely what the ISI's impact factor measures: the average number of times articles in a journal published in a given two year period are cited in the following year. This is a sound measure in a discipline that moves quickly, but a less useful measure in a discipline that moves less rapidly. I selected the 10

[9]The cited half-life of a journal is defined as "the number of journal publication years going back from the current year which account for 50% of the total citations received by the cited journal in the current year" (Garfield 1991:9).

year window and calculated the cumulative citations during that period precisely because I felt doing so struck a balance between (1) producing a measure that captured a substantial portion of the citations an article is likely to receive and (2) producing a measure that is outdated.

Another question worth is asking is whether one can distinguish among journals that receive similar scores on a given indicator. The spread (or variance) of the reputational scores turns out to be rather limited. As Norris and Crewe (1993:12) noted in their survey of British political scientists[10]

> evaluations of... general quality [etc.] generally fell within two points of the mean, in the 5 to 7 point range of the 10-point scale... [the] ratings were bunched and skewed towards the upper end of the scale. The standard deviations showed little variation between journals, although there was a stronger consensus about journals occupying the top rather than the bottom quartile.

The ratings from Giles et al.'s survey of American political scientists exhibit a similar distribution: all but the top six rated journals fall within one standard deviation (2.11) of the mean (5.23), and less than 1 point on the scale separates the 20th and 50th ranked journals (78 were ranked). While none of those who have written on the subject have raised this question, one ought to ask: what is our confidence in these distinctions?

It is, of course, well known how to determine a confidence interval in survey responses. Given that there were roughly 6,000 political scientists in Giles et al.'s population[11] and they received 215 survey respondents, the confidence interval in the sample is somewhere in the neighborhood of .7 (given a 95% confidence level and the 0-10 scale).[12] As such, we cannot have much confidence in the reputational difference between journals that have scores that

[10]They used Giles' reputational mail survey as an instrument.

[11]This figure is taken from the APSA's 1995-97 *Graduate Faculty and Programs in Political Science*, which lists over 6,400 political scientists. Giles et al. conducted their survey in 1988, so there were probably somewhere near 6,000 political scientists listed in the volume they used to draw their sample

[12]I should determine a more precise value. This figure was taken from Manheim and Rich (1995:428) which lists the $\pm$ 5% sample size as 375 and the $\pm$ 10% sample size as 98 for a population of 6,000.

are within $\pm$ 0.7 points of one another.

With respect to Giles' most recent survey, this implies that the top ranked journal, *World Politics*, with a score of 7.9, cannot be meaningfully distinguished from the eighth ranked journal, *Soviet Studies*, which has a score of 7.2. More dramatically, one cannot confidently distinguish the 20th ranked journal, *Journal of Conflict Resolution*, with a score of 6.4 from the 55th ranked journal, *Far Eastern Survey*, which has a score of 5.7. Our limited confidence in the differences found by conducting reputational surveys should at a minimum dissuade us from using ordinal rankings (indeed, Giles et al. do not report the ordinal rankings; they report only the scores, though Garand does report ordinal rankings). Yet it also raises questions about the value of the ratings themselves for making judgments about the relative likely impact of our colleagues' forthcoming and recently published articles. Some distinctions can be drawn between journals with a fairly substantial gap between them, but by and large the reputational survey seems to suggest that there isn't much difference where one publishes ones work.

One might object that this criticism is a bit unfair as it is one-sided: it is not as obvious how one could construct a confidence interval for the citation data. Thus, one cannot say anything about one's confidence in the difference between the scores of two journals on the impact factor or cumulative citations per article measures. However, one can observe that there is a great deal more variation in the citation measures by plotting the two against one another. Figure 1 plots the cumulative cites per article measure against the Giles et al. reputation measure. The plot is bowed with a host of journals with less than an average of less than 2.5 citations per article receiving a wide variety of reputation, with a more linear looking relationship for journals that have an average greater than 2.5. This suggests that while the citation and reputation measures are somewhat consistent for journals that garner more cites than the average number of citations, the measures are not very consistent with one another for journals that are less heavily cited. The converse, however, is not so: virtually all of the journals with a reputation score below 5.5 have less than 2.5 citations per

article, whereas those with a reputation score greater than 6 have citation per article scores ranging from below 2.5 to above 20.

(Figure 1 about here)

These considerations lead me to conclude that the direct citation measures are superior to the indirect reputational measures as indicators of the quality of a journal. In addition, I contend that the 10 year cumulative average citations per article is the most valid indicator of journal quality available. Yet, more information is available by conducting multivariate regressions, and I turn to that task in the following section.

4 Some Multivariate Analyses

More information about the relative merits of the indicators can be gathered by conducting multivariate regression analyses. In this section I present the results from the estimation of four sets of regressions, including a replication of the Christenson and Sigleman study.

The dataset created for this study makes it possible to explore several questions of interest. First, can we account for the variance in the submission rates across political science journals (i.e., the number of manuscripts reviewed at each journal)? Second, can we account for variance in political scientists' familiarity with different journals? Third, can we account for variance in the reputations and citation counts of political science journals? Each issue is addressed in turn.

I begin with an exploration of the determinants of the demand scholars exhibit for the pages of a given journal: the number of submissions a journal reviews. It seems reasonable to surmise that the demand for a journal's pages might be influenced by (1) the circulation of the journal, (2) the quality of the journal, (3) the lag from acceptance to publication, (4) the annual number of articles the journal publishes, (5) the acceptance rate, (6) the length of the review process, and (7) the number of years for which the journal has been published.

The dependent variable, number of submissions reviewed by each journal, is from a survey that was presumably conducted in 1996.[13] I have four measures of journal quality: Giles et al.'s reputation rating (1988), Garand's reputation weighted by familiarity (1988), ISI's impact factor (1988-89 issues cited in 1990), and cumulative citations for 10 years for the 1990 issues.[14] Finally, the other variables are from the Martin and Goehlert (1997) survey.

The bivariate correlations reported in Table 2 reveal that the Giles et al. and ISI factor impact measures have indistinguishably strong relationships with demand for journal space (.37 and .36), and the Garand and cites[15] measures have higher correlations (.51 and .58, the latter of which is not reported in the table). Circulation (.74) has the strongest relationship with demand for space, followed by the number of articles published annually (.57, not listed in Table 2), the year the journal began publishing (-.38), the acceptance rate (-.36), the lag between acceptance and publication (-.26), and the length of the review process (-.23).

Table 4 reports the estimated regression coefficients, associated standard errors and summary statistics on the regressions described above. I estimated several regressions, substituting each of the four journal quality measures. As the table indicates, the only measure of quality that failed to produce a statistically significant parameter estimate was the ISI's impact factor. This is interesting for it is consistent with Christenson and Seligson's (1985) contention that while reputational measures are resistant to change, the ISI's impact factor is responsive. Both reputational indicators and the impact factor measure the 1988-89 period, yet knowing which journals were held in high regard in the late 1980s helps us explain the variance in the demand for pages in journals roughly eight years later whereas knowing how many citations that journaled garnered over a two year window is not, ceteris paribus, very helpful roughly eight years later. Finally, the cumulative number of cites also performs

[13]The data were published in 1997, but Martin and Goehlert do not identify when they conducted their survey.

[14]Because the number of articles published annually is used as a separate control variable, I used the total number of citations rather than the number of citations per article in this analysis.

[15]Because the number of articles published may have an independent impact on demand, I use the total cumulative 10 year citation count, not the cumulative 10 year citation count per article.

well in the regression, but unlike the other measures, it contains information for the period 1990-2000. That it performs well is consistent with my contention that 10 years is not too long a lag to make it dated–the measure is useful for helping explain the most current data available on the demand for space in a journal.

(Insert Table 4 about here)

Circulation, the number of articles published in each issue, and the accepatance rate also produce statistically significant parameter estimates. Alternatively, the lag between acceptance and publication, and the length of the review process, and the venerableness of the journal do not have separate impacts when the other variables are included in the regression. This implies that the larger the circulation of a journal, the higher its quality, the more articles it publishes, and the lower its acceptance rate, the higher will be the demand for space in its pages. The only result that is somewhat counter-intuitive is the sign of the acceptance rate parameter estimate: political scientists appear to collectively court rejection. One might argue that this finding is not surprising on the grounds that high quality journals have lower acceptance rates. While such a belief is certainly borne out in Table 2, it is not obvious why acceptance rate would have a negative relationship with demand once we have controlled for quality. This finding suggests that acceptance rate is capturing something beyond circulation, quality, and articles per annum.[16]

The next question to explore is: What accounts for the variance in political scientists' familiarity with different journals? It seems reasonable to surmise that familiarity may be driven by (1) circulation, (2) quality, and (3) the number of years in which the journal has been in print. The percentage of respondents familiar with each journal is not listed in Table 2, but circulation has a bivariate correlation of .28, the quality measures have bivariate

[16]One caveat is worth noting here: as is often the case when one collects data from several sources, the sample size suffers from missing data. The Martin and Goehlert (1997) volume reported data for only 42 of the 76 journals listed in Table 1. Other missing data eliminated an additional eight cases. Thus, these findings may suffer from unspecified selection bias.

correlations of .38 (Giles et al.), .34 (impact factor), and .22 (cites/article), and the year the journal was founded has a correlation of -.38 with familiarity.[17]

Table 5 contains the results from the regression of familiarity on the three indicators. Again, several models were estimated, substituting the different measures of journal quality. All three measures produce statistically significant parameter estimates and, interestingly, circulation failed to produce a statistically significant coefficient in any of the regressions while the year the journal was founded had a statistically significant coefficient in only one of the three (though including them improved both the fit of the model and reduced the standard errors associated with the quality coefficients). These results suggest that the primary determinant of a political scientist's familiarity with journals is the quality of the journal. That said, the explained variance is fairly low suggesting large random error or, more likely, incomplete specification.

The final question to explore is: What explains the variance in the different measures of quality? I begin with the Giles et al. reputational ratings. It seems reasonable to surmise that (1) circulation, (2) the age of the journal, (3) recent citations, and (4) familiarity are likely related to the reputation of a journal. A review of Table 2 indicates that neither circulation nor the birth date of the journal are strongly related to the Giles et al. measure. On the other hand, the ISI's impact factor (which is a measure of 1990 citations, so the temporal order is a bit suspect) has a strong relationship with reputation (.71). In addition, familiarity (.38), the number of manuscripts reviewed (.36), the acceptance rate (-.26), and the lag to publication (-.24) have moderate to weak relationships with reputation.[18] I thus regressed the Giles et al. reputational measure on impact factor, familiarity, the number of manuscripts reviewed, the acceptance rate, and the lag to publication. The results are contained in Table 6.

(Insert Table 6 about here)

[17]Because familiarity is included as part of Garand's measure, I do not use it in this regression.

[18]Several of these are 1997 measures and will be replaced with more contemporaneous 1990 measures when I receive the data.

The results suggest that the ISI impact factor, the percentage of political scientists familiar with the journal, and the lag between acceptance and publication all have a positive impact on a journal's reputation. While impact factor and familiarity have intuitive interpretations, publication lag does not.[19]

Given that I can essentially replicate Christenson and Sigelman's (1985) study by regressing Giles et al.'s reputation measure on impact factor, and that doing so increases the sample size as well, I chose to do so, and report the results in the second column of Table 6. Christenson and Sigelman further argue that one can use the residuals from such a regression to evaluate whether some journals reputations are over-rated relative to their impact. The correlation between the residuals from the regression and the dependent variable is .71 (Christenson and Sigelman report a correlation of .82 between the Giles and Wright (1975) reputational measure and the residuals from their regression), and a visual inspection of a scatterplot of the residual against the dependent variable confirms "that in each field high-status journals tend to have better reputations than their influence would warrant, while lower-status journals tend to have poorer reputations than their influence would warrant" (Christenson and Sigelman, 1985:971). Table 7 lists the 63 journals, ranked by their residuals. Those with positive residuals are over-rated in comparison with their impact factor and those with negative residuals are under-rated relative to their impact factor. Christendon and Sigelman noted that *World Politics* and the *Journal of Politics* were the two most over-rated journals by their findings. Both the *Journal of Politics* and *World Politics* are in the top 10 over-rated journals by this analysis. This finding is consistent with Christenson and Sigelman's contention that journal reputations are resistant to change.

(Insert Table 7 about here)

The final regression analysis seeks to account for the variance in cumulative citations from 1990 to 2000 in the 1990 issue of political science journals. It seems reasonable to

[19]Unfortunately, missing data again drive down the sample size.

conjecture that (1) the number of articles published per issue, (2) the number of citations to recent issues, (3) the reputation of the journal, and (4) the circulation of the journal will each affect the total number of citations that an issue of a journal will receive. The bi-variate correlations between these variables and 10 year cumulative citations (1990-2000) are .78 (impact factor, 1990), .48 (Giles et al. reputation), .40 (Garand reputation weighted by familiarity), .34 (number of articles), and .07 (circulation). I estimated a regression equation based on the above specification, and the results are reported in Table 8.

(Insert Table 8 about here)

The results in Table 8 suggest that each of the variables has a statistically significant impact on the number of cumulative citations a journal receives.[20] The implication is that reputation has an independent impact on the average number of citations per article a journal draws, even when controlling for the recent citations per article to that journal (as measured by impact factor). This finding provides some support for Lester's (1990) argument that any individual measure only captures a portion of the quality of a journal. The fourth column is included because dropping circulation almost doubles the size of the sample.[21]

What should one make of Lester's argument and these regressions? My own position is that these results do not provide much direction for producing the kind of systems measure envisioned by Lester. Few of the additional variables he calls for add much information about the quality of a journal. Further, the best method for producing a summary measure is less than obvious (though one could conduct a factor analysis and do it empirically). An obvious downside to following Lester's lead is that it is costly–doing so would require several

[20]The residuals in these regressions were not very well behaved. I report Huber-White standard errors in two of the regressions to account for heteroskedasticity. In addition, I re-estimated the equations without the articles variable, but doing so did not improve the residuals (nor did it lead to different inferences in the impact of the remaining variables). I also estimated negative binomial regressions using the 10 year total cumulative cites as the dependent variable, and included the number of articles on the right hand side. The regressions produced the same inferences as those reported in Table 8 (except that articles is statistically significant in those regressions).

[21]The reputation indicators were not included because each is too collinear with impact factor when included in this regression.

distinct data collection efforts. The citation count data, on the other hand, are provided by ISI (impact factor) or can be collected at `webofscience.com`in 5-6 hours. When one weighs the costs of data collection, the citation count measures are clearly superior. Whether they are sufficient to aid individual political scientists as they forecast the likely impact of their colleagues' work is a question individual political scientists will have to answer on their own.

5 Conclusion

We all have (financial as well as psychological) incentives to contend that the journals in which our work is published are the best in our field. Given that power and influence is not distributed uniformly in departments, the existence of this incentive has the potential to create inequities during intradepartmental evaluations. Departmental power brokers may seek to impose their personal reputational rankings on the evaluations of less influential colleagues. A major advantage of accepting the consensus view sketched by Norris and Crewe is that it dampens the ability of power brokers to do so: to the extent that the evaluation of one's recently published or forthcoming work depends on a measure of the quality of the journal in which it is published, it is less likely that one will be victimized by skewed rankings that poorly forecast the likely impact of one's work in one's field.

Michael Giles and his colleagues have twice provided the discipline with a valuable public good, and Norris and Crewe have followed their lead and produced the same public good for British political scientists. Yet, are reputational measures the best we can do. I have argued that they are not. Further, I have argued that though useful, the ISI's impact factor citation score is a less useful indicator of the quality of political science journals than is the 10 year cumulative citations per article measure reported here.

In closing, people will undoubtedly find that a journal in which they have interest is not included. I presently plan to update this measure (i.e., create a 10 year cumulative citations per article score for 1991) in May or June 2001. I encourage those interested in lobbying on

behalf of a journal not included in this study to contact me.

References

Christenson, James A. & Lee Sigelman. 1985. "Accrediting Knowledge: Journal Stature and Impact in Social Science." *Social Science Qaurterly* 66(4):964–975.

Cole, Stephen. 1983. "The Hierarchy of the Sciences." *American Journal of Sociology* 89(0):111–139.

Garfield, Eugene. 1972. "Citation Analysis as a Tool in Journal Evaluation." *Science* 178(0):471–479.

Garfield, Eugene. 1991. *1990 Journal Citation Reports: A Bibliometric Analysis of Social Science Journals in the ISI Database*. Philadelphia: .

Garfield, Eugene. 1998. "The Multiple Meanings of Impact Factors." *Journal of the American Society for Information Science* 49(8):768–768.

Giles, Michael W., Francie Mizell & David Patterson. 1989. "Political Scientists' Journal Evaluations Revisited." *PS: Political Science and Politics* 22(3):613–617.

Giles, Michael W. & Gerald Wright. 1975. "Political Scientists' Evaluations of Sixty-Three Journals." *PS: Political Science and Politics* 8(2):254–257.

Goodson, Larry P., Bradford Dillman & Anil Hira. 1999. "Ranking the Presses: Political Scientists' Evaluations of Publisher Quality." *PS: Political Science & Politics* 32(2):257–262.

Harter, S.P. & T.E. Nisonger. 1997. "ISI's Impact Factor as a Misnomer: A Proposed New Measure to Assess Journal Impact." *Journal of the American Society for Information Science* 48(12):1146–1148.

Lester, James P. 1990. "Evaluating the Evaluators: Accrediting Knowledge and the Ranking of Political Science Journals." *PS: Political Science and Politics* 23(3):445–447.

Moed, H.F., T.N Van Leeuwen & J. Reednijk. 1999. "Toward Appropriate Indicators of Journal Impact." *Scientometrics* 46(3):575–589.

Norris, Pippa & Ivor Crewe. 1993. "The Reputation of Political Science Journals: Pluralist and Consensus Views." *Political Studies* 41(1):5–23.

Porta, M. 1996. "The Bibliographic "Impact Factor" of the Institute for Scientific Information: How Relevant is it for Public Health Journals?" *Journal of Epidemiology and Community Health* 50(6):606–610.

Rousseau, R. & G. Van Hooydonk. 1998. "The Harter-Nisonger Proposal Concerning Impact Factors." *Journal of the American Society for Information Science* 49(6):577–578.

Table 1
Citations per Article and other Measures of Journal Quality

Journal	Cites/Art	Impact Factor	Giles et al.	Garand
Am Pol Sci Review	23.3	2.282	7.6	15.09
Intl Organization	19.5	2.102	7.1	9.78
Intl Security	18.5	1.694	NR	NR
World Politics	15.5	2.267	7.9	12.6
Signs	14	1.708	NR	NR
Am J of Pol Science	13.9	1.268	7.5	13.74
Political Theory	12.8	0.412	6.6	8.45
Social Forces	12.6	1.01	6.2	8.22
Public Opinion Qrtrly	11.3	0.656	6.4	10.27
J of Politics	11	0.701	7.4	14.14
Brit Jrnl of Pol Sci	10.2	0.564	6.8	10.75
Intl Stud Qrtrly	9.2	1.364	6.5	9.22
Political Behavior	8.8	NF	5.7	7.56
Theory and Society	8	0.481	NR	NR
Daedalus	7.8	0.784	6.3	10.24
J of Conf Res	6.9	0.781	6.4	10.08
Am Pol Qrtrly	6.8	0.652	5.8	9.22
Politics and Society	6.8	0.543	6	7.87
Social Science Qrtrly	6.8	0.351	6.1	9.56
Comp Stud in Soc & Hist	6.3	0.579	NR	NR
Pub Admin Rev	6.2	0.828	6.6	10.41
Am J of Intl Law	6.1	0.694	7.3	8.66
Behavioral Science	6.1	0.452	5.4	6.6
Lat Am Res Rev	6.1	0.757	NR	NR
Legis Studies Qrtrly	5.6	0.523	6.1	9.08
J of Asian Studies	5.4	0.529	5.8	6.5
Am Behavioral Scientist	4.8	0.191	5.3	7.39
Intl Affairs	4.8	1.057	6.2	7.81
Foreign Affairs	4.7	1.1	6.4	11.25
Policy Sciences	4.6	0.357	5.9	7.86
Comp Pol Stud	4.3	0.711	6.5	9.49
Comp Pol	4.1	0.455	7	10.87
Eur Jrnl Pol Res	3.9	0.507	6.3	7.62
Public Choice	3.9	0.497	6	7.98
Admin and Society	3.8	0.239	5.8	7.5
J of Peace Res	3.5	0.456	5.7	7.48
Public Interest	3.4	0.871	5.8	8.61
Judicature	3.2	0.39	5.8	6.94
Annals of Am Acad of P&S Sci	3.1	0.453	5.4	9.76
Political Studies	3.1	0.384	6.2	7.73
Soviet Studies	3	0.623	7.2	7.87
Can Jrnl Pol Sci	2.9	0.357	6.1	8.2
J of Latin American Studies	2.9	0.29	6.4	7.24
J of Modern African Studies	2.8	0.301	NR	NR
West Pol Qrtrly (now PRQ)	2.8	0.529	6	10.71
China Qrtrly	2.7	0.521	6.6	7.65
Political Sci Qrtrly	2.7	0.605	6	9.97
J of Developing Areas	2.6	0.213	5.8	7.16
Polity	2.4	0.203	6	10.15
Am Rev of Public Admin	2.3	NF	5.5	6.94
J of Democracy	2.3	NF	NR	NR
Conf Mgmnt & Peace Sci	2.2	0.5	NR	NR
J of Intl Affairs	1.9	0.122	5.9	6.83
Review of Politics	1.9	0.279	5.7	7.96
J of InterAm Stud & Wrld Aff	1.8	0.128	5.9	6.53
Asian Survey	1.7	0.254	5.5	6.68
Policy Studies J	1.6	0.219	5	7.54
Social Science J	1.6	0.141	5.2	6.24
Rationality and Society	1.4	NF	NR	NR
Simulation and Games	1.4	0.244	4.1	4.42
Intl Social Science J	1.3	0.192	5.3	6.39
Political Qrtrly	1.3	0.152	5.4	6.69
Slavic Review	1.3	NF	7	7.52
Govt and Opposition	1	0.132	5.8	7.71
J of Black Studies	0.9	0.145	4.3	4.84
Political Science	0.9	0.37	4.9	5.7
Intl Interactions	0.7	NF	5.6	6.49
Third World Quarterly	0.7	0.033	NR	NR
Dissent	0.5	0.135	5.5	7.33
Middle Eastern Studies	0.5	0.02	5.6	6.23
Women and Politics	0.4	NF	4.4	4.98
Orbis	0.3	0.04	5.2	7.31
Publius	0.3	0.395	5.8	8.51
PS: Pol Sci and Politics	0.2	0.5	5.4	10.09

Table 2
Bivariate Correlations

	Cites/ Article	ISI's Impact Factor	Giles et al.	Garand	# of Ms. Reviewed	Acceptance Rate	Review Period	% Solicited	Lag to Publication	Year Begun	Circulation
Cites/ Article	1.0										
ISI's Impact Factor	0.82*	1.0									
Giles et al.	0.54*	0.71*	1.0								
Garand	0.44*	0.61*	0.77*	1.0							
# of Ms. Reviewed	0.35*	0.37*	0.36*	0.51	1.0						
Acceptance Rate	-0.30*	-0.41*	-0.26	-0.36*	-0.41*	1.0					
Review Period	-0.11	-0.16	-0.04	-0.07	-0.23	-0.25	1.0				
% Solicited	-0.18	-0.01	-0.03	-0.03	0.05	0.57*	-0.31*	1.0			
Lag to Publication	-0.15	-0.24	0.06	-0.10	-0.26	-0.09	0.25	-0.40	1.0		
Year Begun	-0.09	-0.09	-0.14	-0.28*	-0.38*	0.30*	0.20	0.12	-0.04	1.0	
Circulation	0.05	0.22	0.13	0.26	0.74*	-0.22	-0.16	0.38*	-0.33*	-0.32*	1.0

Note: * indicates that the correlation is statistically significant at the .10 level.

Table 3
Descriptive Statistics

Variable	N	Minimum	Mean	Stand. Deviation	Maximum
Cites	79	1	227.6	395.9	2681
Articles	79	6	31.8	15.4	84
Cites/Article	79	0	6.9	9.8	62.3
Impact Factor	71	0.02	0.68	0.64	2.75
Giles et al	69	4.1	6.08	0.79	7.9
% Familiar	69	7	39	23	99
Garand	69	4.4	8.5	2.1	15.1
# of Ms. Reviewed	41	40	202	141	750
Acceptance Rate	42	0.03	0.18	0.13	0.90
Review Period	42	1	2.7	1.0	6
% Solicited	42	0	10.6	19.8	97.5
Lag to Publication	42	2	8.5	3.4	16.5
Year Begun	42	-14	57.6	21.5	90
Circulation	40	400	5611	15,802	100,000

Table 4
The Demand for a Journal's Space

Variable	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)
Circulation	0.005** (0.001)	0.005** (0.001)	0.004** (0.001)	0.003** (0.001)
Cites	0.219** (0.640)			
Impact Factor		24.37 (29.09)		
Giles et al.			32.47** (15.44)	
Garand				11.66** (5.48)
Lag to Publication	0.724 (3.604)	1.649 (4.910)	0.244 (3.581)	1.146 (3.543)
Articles per Annum	2.119** (0.922)		4.289** (0.882)	4.017** (0.900)
Acceptance Rate	-227.9** (97.8)	-701.7** (310.4)	-940.0** (209.5)	-915.1** (211.7)
Review Period	-14.44 (13.26)	-29.71 (18.08)	-18.20 (12.70)	-20.18 (12.72)
Year Begun	0.194 (0.586)	-0.756 (0.808)	0.222 (0.567)	0.375 (0.575)
Constant	135.5* (67.2)	354.2** (94.8)	47.0 (116.8)	134.6 (84.8)
N	39	37	36	36
F	19.46**	10.21**	20.89**	20.97**
Adj. R^2	0.77	0.60	0.80	0.80

Note: ** indicates statistically significant at the .05 level, * indicates statistically significant at the .10 level.

Table 5
Political Scientists' Familiarity with a Journal

Variable	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)
Circulation	0.000 (0.000)	0.000 (0.000)	0.000 (0.00)
Cites/Article	0.015** (0.005)		
Impact Factor		0.133** (0.055)	
Giles et al.			0.155** (0.374)
Year Begun	-0.002 (0.001)	-0.002 (0.002)	0.002* (0.001)
Constant	0.490 (0.098)	0.493** (0.099)	-0.371 (0.250)
N	37	36	37
F	5.49**	4.21**	9.21**
Adj. R^2	0.27	0.22	0.41

Note: ** indicates statistically significant at the .05 level, * indicates statistically significant at the .10 level.

Table 6
Giles et al.'s Journal Reputation Rating

Variable	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)
Impact Factor	0.827 (0.131)	0.86 (0.11)
Familiarity	0.806 (0.381)	
Manuscripts Reviewed	0.000 (0.001)	
Acceptance Rate	-0.385 (1.306)	
Publication Lag	0.051 (0.020)	
Constant	4.887 (0.383)	5.54 (0.10)
	N = 35 F (5, 29) = 15.83 Adj R^2 = 0.69	N = 63 F (1, 61) = 60.58 Adj R^2 = 0.49

Table 7
Over- and Under-Rated Political Science Journals

Journal	Residual	Reputation	Predicted Reputation
J of Politics	1.3	7.4	6.1
Am J of Intl Law	1.2	7.3	6.1
Soviet Studies	1.1	7.2	6.1
Comp Pol	1.1	7	5.9
Am J of Pol Science	0.9	7.5	6.6
Brit Jrnl of Pol Sci	0.8	6.8	6
Political Theory	0.7	6.6	5.9
China Qrtrly	0.6	6.6	6
J Latin American Studies	0.6	6.4	5.8
World Politics	0.4	7.9	7.5
Comp Pol Stud	0.3	6.5	6.2
Pub Admin Rev	0.3	6.6	6.3
Political Studies	0.3	6.2	5.9
Am Sociological Review	0.3	7.6	7.3
Eur Jrnl Pol Res	0.3	6.3	6
Public Opinion Qrtrly	0.3	6.4	6.1
Polity	0.3	6	5.7
Social Science Qrtrly	0.3	6.1	5.8
J of Intl Affairs	0.3	5.9	5.6
Can Jrnl Pol Sci	0.3	6.1	5.8
J InterAm Stud & Wrld Aff	0.2	5.9	5.7
J Conf Res	0.2	6.4	6.2
Govt and Opposition	0.1	5.8	5.7
Legis Studies Qrtrly	0.1	6.1	6
Am Pol Sci Review	0.1	7.6	7.5
Daedalus	0.1	6.3	6.2
J of Developing Areas	0.1	5.8	5.7
Admin and Society	0.1	5.8	5.7
Policy Sciences	0.1	5.9	5.8
West Pol Qrtrly (now PRQ)	0	6	6
Politics and Society	0	6	6
Public Choice	0	6	6
Middle Eastern Studies	0	5.6	5.6
Pol Sci Qrtrly	-0.1	6	6.1
Judicature	-0.1	5.8	5.9
Publius	-0.1	5.8	5.9
Review of Politics	-0.1	5.7	5.8
Foreign Affairs	-0.1	6.4	6.5
Dissent	-0.2	5.5	5.7
Am J of Sociology	-0.2	7.5	7.7
J of Asian Studies	-0.2	5.8	6
Social Forces	-0.2	6.2	6.4
Intl Stud Qrtrly	-0.2	6.5	6.7
J Peace Res	-0.2	5.7	5.9
Intl Organization	-0.2	7.1	7.3
Intl Affairs	-0.2	6.2	6.4
Asian Survey	-0.3	5.5	5.8
Political Qrtrly	-0.3	5.4	5.7
Am Pol Qrtrly	-0.3	5.8	6.1
Admin Science Qrtrly	-0.3	6.9	7.2
Orbis	-0.4	5.2	5.6
Am Behavioral Scientist	-0.4	5.3	5.7
Intl Social Science J	-0.4	5.3	5.7
Social Science J	-0.5	5.2	5.7
Public Interest	-0.5	5.8	6.3
Behavioral Science	-0.5	5.4	5.9
Annals of Am Acad of P&S Sci	-0.5	5.4	5.9
PS: Pol Sci and Politics	-0.6	5.4	6
Policy Studies J	-0.7	5	5.7
Political Science	-1	4.9	5.9
J of Pol Economy	-1.1	6.8	7.9
J of Black Studies	-1.4	4.3	5.7
Simulation and Games	-1.6	4.1	5.7

Table 8
Cumulative Citations per Article, 1990-2000

Variable	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)
Articles	7.1320 (1.374)	6.4305 (1.3603)	6.95 (1.80)
Giles et al.	98.519 (45.528)		
Garand		31.135 (12.172)	
Impact Factor	164.03 (54.20)	180.63 (44.10)	491.64 (44.37)
Circulation	-0.0020 (0.0012)	-0.0024 (0.0012)	
Constant	-736.52 (263.62)	-398.05 (99.75)	-312.97 (69.10)
	N = 36 F (4, 31) = 21.87 Adj R^2 = 0.70	N = 36 F (4, 31) = 23.42 Adj R^2 = 0.72	N = 71 F (2, 68) = 73.28 Adj R^2 = 0.67