



Replication, Replication

Author(s): Gary King

Source: *PS: Political Science and Politics*, Sep., 1995, Vol. 28, No. 3 (Sep., 1995), pp. 444-452

Published by: American Political Science Association

Stable URL: <https://www.jstor.org/stable/420301>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



American Political Science Association is collaborating with JSTOR to digitize, preserve and extend access to *PS: Political Science and Politics*

JSTOR

September symposium speaks to the purposes of *PS: Political Science & Politics*. The journal serves many ends. It is a journal of contemporary politics striving to publish intellectually engaging and accessible analyses and commentaries on current political issues and events. *PS* is the principal pedagogic journal within the profession offering articles on teaching, training, and curriculum development. With over 16,000 subscribers, *PS* is

also *the* journal of the profession. As illustrated by the symposium, *PS* is a forum in which political scientists from all regions of the country and world, from every type of degree-granting institution, as well as from all subfields can debate what it means to be a “political scientist” and where the rules and tools of our practice can be identified, debated, and ultimately agreed upon.

References

- Altman, Lawrence K., M.D. “Revisionist History Sees Pasteur as Liar Who Stole Rival’s Ideas,” *The New York Times*, 16 May 1995, sec. C.
- Geison, Gerald L. 1995. *The Private Science of Louis Pasteur*. Princeton, NJ: Princeton University Press.

Replication, Replication

Gary King,¹ *Harvard University*

Political science is a community enterprise; the community of empirical political scientists needs access to the body of data necessary to replicate existing studies to understand, evaluate, and especially build on this work. Unfortunately, the norms we have in place now do not encourage, or in some cases even permit, this aim. Following are suggestions that would facilitate replication and are easy to implement—by teachers, students, dissertation writers, graduate programs, authors, reviewers, funding agencies, and journal and book editors.

Problems in Empirical Political Science

As virtually every good methodology text explains, *the only way to understand and evaluate an empirical analysis fully is to know the exact process by which the data were generated and the analysis produced*. Without adequate documentation, scholars often have trouble replicating their *own* results months later. Since sufficient information is usually lacking in political science, trying to replicate the results of others, even with their help, is often impossible.

For quantitative and qualitative analyses alike, we need the answers to questions such as these:

How were the respondents selected? Who did the interviewing? What was the question order? How did you decide which informants to interview or villages to visit? How long did you spend in each community? Did you speak to people in their language or through an interpreter? Which version of the ICPSR file did you extract information from? How knowledgeable were the coders? How frequently did the coders agree? Exactly what codes were originally generated and what were all the recodes performed? Precisely which measure of unemployment was used? What were the exact rules used for conducting the content analysis? When did the time series begin and end? What countries were included in your study and how were they chosen? What statistical procedures were used? What method of numerical optimization did you choose? Which computer program was used? How did you fill in or delete missing data?

Producing a comprehensive list of such questions for every author to address, or deciding *ex ante* which questions will prove consequential, is virtually impossible. For this reason, quantitative analysts in most disciplines have almost uniformly adopted the same method of ascertaining whether enough information exists in a published work. The *replication*

standard holds that sufficient information exists with which to understand, evaluate, and build upon a prior work if a third party could replicate the results without any additional information from the author. The replication standard does not actually require anyone to replicate the results of an article or book. It only requires sufficient information to be provided—in the article or book or in some other publicly accessible form—so that the results could in principle be replicated. Since many believe that research standards should be applied equally to quantitative and qualitative analyses (King, Keohane, and Verba 1994), the replication standard is also appropriate for qualitative research, although the rich complexity of the data often make it more difficult.²

The process of reducing real-world phenomena to published work involves two phases: the representation of the real world by essentially descriptive quantitative and qualitative data, *and* the analysis of these data. Both phases are important components of the replication standard. Future scholars, with only your publication and other information you provide, ought to be able to start from the real world and arrive at the same substantive conclusions. In many types of research this is not possible, but it should always be at-

tempted. In principle, the replication standard can sometimes be met even without making public the data used in the analysis, provided that one's description of both phases of the analysis is sufficiently detailed. However, providing this level of detail without the data is difficult if not impossible for the author and much less helpful to future researchers. Moreover, it may not be possible to replicate the data collection phase, inasmuch as the world may have changed by the time a future researcher undertakes the duplication effort.

An excellent example of a recent study of adherence to the replication standard is by Dewald, Thursby, and Anderson (1986). One of the authors was the editor of the *Journal of Money, Credit, and Banking*. After accepting a year's worth of articles, they received an NSF grant to replicate the results from all the articles accepted. Their work is a revealing (and disconcerting) report of their extensive but largely failed attempts to replicate each of these articles. Their findings (p. 587–88) “suggest that inadvertent errors in published empirical articles are a commonplace rather than [a] rare occurrence.” Even when they found no errors, replication was often impossible even with the help of the original author—and help from the authors often was not provided. More important, when the editors started requiring authors to meet the replication standard, they (p. 589) “found that the very process of authors compiling their programs and data for submission reveals to them ambiguities, errors, and oversights which otherwise would be undetected.” Since political scientists collect far more *original* data, rather than following the economists' practice of relying primarily on *existing* data from government sources, the benefits of a replication policy in our discipline should be even more substantial than indicated in Dewald, Thursby, and Anderson's conclusions.³

As this rather striking example demonstrates, the widespread failure to adhere to the replication standard poses serious problems for any discipline. At its most fun-

damental, if the empirical basis for an article or book cannot be reproduced, of what use to the discipline are the conclusions? What purpose does an article like this serve? At a minimum, some protection should be afforded to keep researchers from wasting their time reading these works. At worst, vastly more time can be wasted in ultimately fruitless efforts to expand, extend, and build on a body of work that has no empirical foundation.

More generally, the replication standard enables scholars to better understand and evaluate existing research, and select more discriminatingly among this body of work in developing their own research agendas. Without complete information about where data come from and how we measured the real world and abstracted from it, we cannot truly understand a set of empirical results.⁴ Evaluation likewise requires at least as much information. Thus, reviewers and journal and book editors should be privy to sufficient information to replicate work submitted to them for publication. Perhaps most importantly, the replication standard is extremely important to the further development of the discipline. The most common and scientifically productive method of building on existing research is to replicate an existing finding—to follow the precise path taken by a previous researcher, and then improve on the data or methodology in one way or another. This procedure ensures that the second researcher will receive all the benefits of the first researcher's hard work. After all, this is why academics refer to articles and books as “scholarly *contributions*,” and such contributions are recognized with citations, acknowledgments, promotions, and raises. Such contributions are considerably more valuable when the cost of building thereon is as small as possible.

Reproducing and then extending high-quality existing research is also an extremely useful pedagogical tool, albeit one that political science students have been able to exploit only infrequently given the discipline's limited adherence to the replication standard. Moreover,

apart from these altruistic reasons to support the replication standard, there is an additional, more self-interested motivation: an article that cannot be replicated will generally be read less often, cited less frequently, and researched less thoroughly by other scholars. Few events in academic life are more frustrating than investing enormous amounts of time, effort, and pride in an article or book, only to have it ignored by the profession, not followed up by other researchers, not used to build upon for succeeding research, or not explored in other contexts. Moreover, being ignored is very damaging to a career, but being applauded, cited favorably, criticized, or even attacked are all equally strong evidence that you are being taken seriously for your contributions to the scholarly debate (see Feigenbaum and Levy 1993, citing Diamond 1988, and Leimer and Lesnoy 1982). Unfortunately, a recent study indicates that the modal number of citations to articles in political science is *zero*: 90.1% of our articles are never cited (Hamilton 1991; Pendlebury 1994)! An even smaller fraction of articles stimulates active investigation by other researchers.

This problem greatly limits our collective knowledge of government and politics. Academia is a social enterprise that is usually most successful when individual researchers compete and collaborate in contributing toward common goals. In contrast, when we work in isolation on unrelated problems, ignoring work that has come before, we lose the benefits of evaluating each other's work, analyzing the same problem from different perspectives, improving measurement techniques and methods, and, most important, building on existing work rather than repeatedly reinventing the wheel.

Proposed Solutions

Solutions to many existing problems in empirical political science are best implemented by individual authors. However, experience in many disciplines has shown that

some formal rules are also needed. Academics, administrators, reviewers, and editors can play an important part in encouraging or requiring adherence to the replication standard.

Authors

If individual authors wish to increase the probability that their work will be read, understood, and taken seriously in future research, following the replication standard is a very important step. (It is also an effective means of ensuring that researchers will be able to follow up on their own work after the methodological details have faded from memory.)

In practice, following the replication standard might involve putting more information in articles, books, or dissertations about the precise process by which information was extracted or data collected, coded, analyzed, and reported. Unfortunately, journals and books generally will not provide sufficient space to do this properly. Moreover, much of the material necessary is best communicated in electronic form rather than on paper. Fortunately, two of the discipline's best digital archives, described in more detail below, can readily be used to satisfy the replication standard: the collection of the Public Affairs Video Archive at Purdue University and the Inter-University Consortium for Political and Social Research at the University of Michigan.

The first step in implementing the replication standard is to create a *replication data set*. Replication data sets include all information necessary to replicate empirical results. For quantitative researchers, these might include original data, specialized computer programs, sets of computer program recodes, extracts of existing publicly available data (or very clear directions for how to obtain exactly the same ones you used), and an explanatory note (usually in the form of a "read.me" file) that describes what is included and explains how to reproduce the numerical results in the article. One need

not provide all information available, only the subset of variables and observations from a data set actually used to produce the published results.

The replication standard can be applied analogously vis-à-vis most qualitative research. A replication data set for qualitative projects should include detailed descriptions of decision rules followed, interviews conducted, and information collected. Transcripts of interviews, photographs, or audio tapes can readily be digitized and included in a replication data set. Adhering to the replication standard is more difficult in qualitative research and sometimes cannot be

Replication data sets include all information necessary to replicate empirical results.

completely followed. But because rich, detailed qualitative data is very informative, not adhering to the replication standard when it is not possible is still well worth the cost. It would also be worthwhile for qualitative researchers to begin to discuss collectively the appropriate applications or modifications of the replication standard (see Griffin and Ragin 1994).

Once a replication data set has been created, it should be made publicly available and reference to it made in the original publication (usually in the first footnote). One approach is to make the information available on request, but this can be inconvenient to both requestors and authors alike. Moreover, academics are rarely professional archivists. Their comparatively high degree of mobility also complicates self-distribution, in that affiliations indicated on earlier published articles will not remain accurate.

These problems are resolved by using professional data archives. Professional archiving entails routine backups on site, off-site duplicates in a different building or city,

continual verification of the digital medium on which the replication data set is stored, a commitment to permanent storage (which involves changing the storage medium as technology changes), frequent advertising so the author's contribution will remain widely known, fast and efficient methods for distributing this information to anyone who asks, and sufficient funding to perform each of these functions for the indefinite future.

As noted, two archives—the "Social Science Research Archive" of the Public Affairs Video Archive (PAVA) at Purdue University and the "Publication-Related Archive" of the Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan—will now accept replication data sets. PAVA is the more technically up-to-date archive. Staff will make data available within hours of submission. Replication data sets are instantly available via Internet through such servers as "gopher," "anonymous FTP" (file transfer protocol), and "Mosaic." Anyone, anywhere in the world, with an Internet account has free, unlimited access to these data.

The ICPSR is the older and better known of the two archives. Its staff will also keep and distribute data, and is presently able to distribute publications-related data via FTP or through the mail to other scholars. The ICPSR also offers other classes in which to deposit data, if the submitter is willing to provide additional documentation; for these, the ICPSR will provide various levels of data checking and additional advertising. Thus, PAVA has some technological advantages over the ICPSR, but the ICPSR is still the better known institution and also offers more options. Moreover, submission of data sets is free and relatively easy in both cases. There is little cost in submitting data to both institutions (as is my current practice).

Replication data sets can be submitted to either archive via disk or tape, mailed to PAVA (Director, Public Affairs Video Archive, Purdue University, 1000 Liberal Arts Building, West Lafayette, IN 47907-1000) and/or the ICPSR

(Director, User Support, ICPSR, P.O. Box 1248, Ann Arbor, MI 48106). An easier and quicker approach is to put data in a self-extracting archive file (with a utility such as PKZIP for the DOS operating system, TAR for Unix, or StuffIt for the Macintosh) and submit the data via anonymous FTP; the file name, article, book, or dissertation citation, and a brief paragraph describing the contents should also be included in an accompanying electronic mail message. To send to PAVA, FTP to **pava.purdue.edu** in directory **pub/incoming** and send electronic mail to **info@pava.purdue.edu**. To submit to the ICPSR, FTP to **ftp.icpsr.umich.edu** in directory **pub/incoming** and send electronic mail to **jan@tdis.icpsr.umich.edu**.

Once a replication data set is submitted and made available by the archive, it will be advertised by PAVA and ICPSR through their regular publications and catalogues and on the Internet. To maximize visibility, citations to the publication and corresponding replication data set will also appear in several newsletters distributed to the membership of the American Political Science Association (as described later).

Tenure and Promotion Review Committees

Tenure and promotion review committees are in the business of judging candidates for promotion in their contributions to the scholarly community. Adherence to the replication standard should be part of this judgment. Those who follow this standard are more likely to enjoy a wider scholarly audience and have their research better understood and extended; thus, they will be taken more seriously by their scholarly peers. In addition, however, candidates for tenure and promotion who submit their data to a national archive should be recognized for this contribution to the discipline. I recommend, for example, that scholars add an extra section to their curriculum vitae for "replication data sets archived." This important contribution would

then be more clearly noticed, and should be given substantial weight in committee, departmental, and university decisions about promotion and tenure. Outside letter-writers should also make note of these significant contributions to the scholarly community.

Graduate Programs

The design of graduate programs and specific research-oriented courses can also encourage adherence to the replication standard, which in turn can strengthen students' ability to learn the basics of academic research and ultimately conduct their own original research. The first professional "publication" for most political scientists is the Ph.D. dissertation. This is intended to be an original contribution to knowledge and to the scholarly community. To maximize the impact of thesis work, students are well advised to submit replication data sets for their dissertations. Graduate programs can also adopt rules that require dissertation students to submit replication data sets when appropriate. In doing so, graduate programs will further socialize and professionalize students into the standards of the discipline.

PAVA will accept replication data sets for dissertations and embargo them for a period of your choosing. In the Department of Government at Harvard University, students doing quantitative and, when applicable, qualitative dissertations must submit replication data sets as a requirement of the Ph.D. degree. (It is important that the student create and submit the replication data set when the work is fresh in his or her mind.) We embargo the data for up to five years, as determined by the student and his or her advisor, to give the student a head start at publication. In most cases since our policy has been adopted, students have opted for a short embargo or none at all.

As noted, having students replicate the results of existing articles has proven to be an effective teaching tool. Many economics graduate programs even require Ph.D. students to replicate a published article for their second-year paper.

This practice will be useful for students in political science as replication data sets become more widely available.⁵

Editors and Reviewers of Books and Journals

Editors of journals and university and commercial presses work hard to publish scholarship that makes important contributions to the political science discipline and has maximum impact in the profession. For the reasons described above, publications by authors who adhere to the replication standard are more likely to meet these criteria. Thus, editors can maximize the influence of their journal or book series by requiring adherence to a replication standard.

Possibly the simplest approach is to require authors to add a footnote to each publication indicating in which public archive they will deposit the information necessary to replicate their numerical results, and the date when it will be available. For some authors, a statement explaining the inappropriateness of this rule, of indeterminate periods of embargo of the data or portions of it, could substitute for the requirement. In this case peer reviewers would be asked to assess the statement as part of the general evaluative process and to advise the editor accordingly. I believe we should give maximum flexibility to authors to respect their right of first publication, the confidentiality of their informants, and for other reasons that are discussed below. However, these exceptions do not apply to the vast majority of articles and books in political science.

This policy is very easy to implement, because editors or their staffs would be responsible only for the existence of the footnote, not for confirming that the data set has been submitted nor for checking whether the results actually can be replicated. Any verification or confirmation of replication claims can and should be left to future researchers. For the convenience of editors and editorial boards considering adopting a policy like this, the following is a sample text for such a policy:

Authors of quantitative articles in this journal [or books at this press] must indicate in their first footnote in which public archive they will deposit the information necessary to replicate their numerical results, and the date when it will be submitted. The information deposited should include items such as original data, specialized computer programs, lists of computer program recodes, extracts of existing data files, and an explanatory file that describes what is included and explains how to reproduce the exact numerical results in the published work. Authors may find the "Social Science Research Archive" of the Public Affairs Video Archive (PAVA) at Purdue University or the "Publications-Related Archive" of the Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan convenient places to deposit their data. Statements explaining the inappropriateness of sharing data for a specific work (or of indeterminate periods of embargo of the data or portions of it) may fulfill the requirement. Peer reviewers will be asked to assess this statement as part of the general evaluative process, and to advise the editor accordingly. Authors of works relying upon qualitative data are encouraged (but not required) to submit a comparable footnote that would facilitate replication where feasible. As always, authors are advised to remove information from their data sets that must remain confidential, such as the names of survey respondents.

Some journals may wish to adopt stricter requirements. (Although these may be appropriate in some cases, I believe they are not usually necessary or desirable.) For example, some journals now verify that the data were actually deposited in a public archive. For authors who request embargoes, some journals might wish to require submission at the time of publication and have the archive do the embargoing so that the replication data set will be prepared when it is fresh in the mind of the investigator. The APSA Political Methodology Section has proposed a maximum allowable embargo period of five years. The Committee on National Statistics of the National Academy of Sciences even recom-

mends that data be made publicly available during the review process.

Finally, some journals might wish to experiment with asking an extra reviewer or perhaps a graduate student (acting as an editorial intern to the journal) to replicate analyses for accepted but not yet published articles. Reviewers of the replication data sets could then make suggestions to the authors that could be incorporated before publication to make replication easier or clearer for future scholars. These kinds of experiments would be very useful to journals, authors, readers, and future scholars.

The exact requirement should be left to the needs of individual journals and presses, although in political science the less restrictive version above will be more than adequate in implementing the replication standard. Moreover, it probably fits better with the norms of the discipline.⁶

Important Exceptions

Although the presumption should be that authors will provide free access to replication data, the editor, in combination with the author, will always have final say about applying general policies to particular instances. Exceptions are essential when confidentiality considerations are important, to guarantee authors rights of first publication and for a variety of other reasons. Important as these exceptions are, they will probably not apply in the vast majority of scholarly works in the discipline.

Confidentiality

To maintain confidentiality, survey organizations do not normally release the names and addresses of respondents. In these and related cases, authors relying on such information can comply with the replication requirement by releasing a subset of their data, stripped of identifying information. However, in some instances, providing any data would be inappropriate. For example, in a recent article on graduate admissions, my coauthors and I used data from Harvard's

admissions process (King, Bruce, and Gilligan 1993). Clearly, the grades and GRE scores of students in the program cannot be released. In fact, we even withheld regression coefficients in the paper, since we felt it would be inappropriate for prospective students to be able to calculate expected grades in our program. Publishing sufficient information to enable students to calculate the expected probability of admission would also have been an unpopular decision at the Harvard general counsel's office! However, cases such as these are the exception rather than the rule.

In some rare situations, confidentiality could not be protected if *any* data were made publicly available. For example, studies based on elite interviews among a small population, or other surveys based on a very large proportion of the relevant population, potentially pose this problem. In a subset of these cases, the author might be able to make the data available to individual scholars willing to restrict their use in very specific ways. (Analogously, specific data analysis rules have been adopted for U.S. Census data to avoid revealing individual identities. For example, cross-tabulations with fewer than 15 people per cell are not permitted.) These are important exceptions, but they too cover comparatively few published works in political science.

In some situations, data used in a published work cannot be distributed because they are proprietary, such as survey data from the Roper Center. However, most of these organizations allow data to be redistributed by other authors if they are modified in some way, such as by making extracts of the variables used or doing recodes. Wholly proprietary data are rare in political science.

Rights of First Publication

As indicated previously, to guarantee the right of first publication, it is appropriate to submit data to a public archive and request an embargo for some specified period of time. However, embargoes like this should be relatively rare, in that

the replication standard obligates one to provide only the data actually used in a publication. For example, if you conducted your own survey with 300 variables and used only 10 for an article, you need to provide only those 10. If you have a five-category variable and use only two of the categories, you could provide just the recoded variable with only the two categories. If you have 1,500 observations, and use only 1,000 of them in the article (perhaps by dropping the Southern states), you also need to submit only the 1,000 cases used in your analysis. Then you can save the remaining information for your future publications. You certainly could provide the rest of the data, which would probably make your work more valuable to the scholarly community, but the decision to do so would remain with you.

In some cases, authors might wish to embargo the subset of data used in an article to clean, document, and then publicly archive the larger data set from which it was extracted. This would be more convenient for the investigator, and might also benefit future researchers by encouraging them to wait for the more comprehensive version of the data. (In these cases, investigators should retain an old version of the data, or fully document any changes in the data since the publication of the article.)

Broadly speaking, the basic point of this proposal is to change authors' expectations, from the current situation of rarely taking any measure to ensure that their work can be replicated, to usually taking some steps in this direction. Exceptions are important, but they would not apply to the vast majority of articles published.

Support for These Policies

Replication Policy Adoptions

Formal support for these policies appears to be growing. Beginning last year, the *American Journal of Political Science*, under the editorship of Kenneth Meier, and *Political Analysis* (the discipline's methods journal), under the editorship

of John Freeman, now require footnotes about replication data sets to be included with all articles. The editors have encountered no resistance from authors, and the policy has required very little time and effort to implement. (Kenneth Meier also reports that 70% of the empirical articles he has accepted use original data collected by the author.) The *British Journal of Political Science* (David Sanders and Albert Weale, editors), and the *Policy Studies Journal* (edited by Uday Desai and Mack C. Shelley II) have adopted similar policies, and *International Interactions*, under the editorship of Harvey Starr, is in the process of doing so.

The new policy of the University of Michigan Press political science and law is "to expect routinely that all authors have a footnote in their books indicating where their replication data set is archived (and when it will be available, if appropriate)," although it is not an absolute requirement of publication. The Cambridge University Press series, "Political Economy of Institutions and Decisions," under the editorship of James E. Alt and Douglass North, has recently adopted a version of this policy. The Free Press and HarperCollins have done the same. Many other editors and editorial boards in political science have indicated support for a replication policy but are still in the process of considering the specific form of the policy they will adopt.

The National Science Foundation (NSF) Political Science Program's policy regarding replication data sets is clearly stated in their award letters: "All data sets produced with the assistance of this award shall be archived at a data library approved by the cognizant program officer, no later than one year after expiration of the grant." To enforce this rule, the NSF Political Science Program recently adopted several new policies. First, all journal and book publications that were prepared with the help of NSF funds must include a statement indicating in which public archive they will deposit the information necessary to replicate their numerical results and the date that it will

be submitted (or an explanation if it cannot be). Second, when a grant is awarded, the political science program officer will ask the prospective investigator to verify that he or she has allocated sufficient funds to fulfill the program's data-archiving requirement. Third, within a year of a grant's expiration date, principal investigators must inform the political science program officer where their data have been deposited. Finally, NSF program officials will consider conformity with their data-archiving policy as an important additional criterion in judging applicants for renewals and new awards.

Anyone receiving funds from the National Institute of Justice (NIJ) must now "deliver to NIJ, upon project completion, computer-readable copies and adequate documentation of all data bases and programs developed or acquired in connection with the research" (Sieber 1991, 9). The "Committee on National Statistics" of the National Academy of Sciences has also recommended policies similar to those suggested here (see Fienberg et al. 1985). Although there are many national differences in norms of data sharing, related policies and recommendations have been adopted or addressed by national governments, international organizations, academic and professional societies, granting agencies, other disciplines, and scholarly journals (see Boruch and Cordray 1985).

To help provide additional visibility for authors of replication data sets, *The Political Methodologist*, the newsletter of the APSA Political Methodology Section (edited by R. Michael Alvarez and Nathaniel Beck), has announced that scholars' data-set reference, a citation to the associated article, and a brief abstract all will be highlighted in subsequent issues. Similar citations will appear in other newsletters if the data are relevant to their substantive focuses; these newsletters include *Policy Currents* (edited by Laura Brown), *Law and Courts* (edited by Lee Epstein), *Urban Politics Newsletter* (Arnold Vedlitz, editor), the *Computer and Multimedia Section Newsletter* (edited by Bob Brookshire), *The Caucus for a*

New Political Science Newsletter (John C. Berg, editor), *Clio: The Newsletter of Politics and History* (Dave Robertson, editor), and *VOX POP*, the newsletter of the APSA section on political parties and interest groups (edited by John Green).

Replication Policy Discussions

Replication policies have been widely discussed in the political science community in recent years. Among political methodologists, support is enthusiastic and appears to be unanimous. Lengthy formal and informal discussions were held at the three most recent annual Political Methodology Group summer meetings (University of Wisconsin, Madison, 1994; University of Florida, 1993; and Harvard University, 1992). Well-attended panels at the last two meetings of the American Political Science Association have also been devoted in part (in 1993) or in full (in 1994) to this issue. The APSA Political Methodology Section unanimously passed a resolution in 1994 asking all journal editors in the discipline to require footnotes indicating where replication data sets are stored.

The APSA Comparative Politics Section held a discussion of this issue at the 1994 annual convention. After considerable debate at the convention focusing on the special concerns comparativists have about confidentiality and distributing “de-contextualized” data, the Section’s executive committee endorsed the idea in general terms. The committee subsequently wrote a proposed policy statement that reflects the special concerns of comparativists while still requiring the replication footnote. This proposal is now being distributed through the Section newsletter for general comment from the membership.

Wide-ranging discussions have also been held in meetings of many of our discipline’s editorial boards and section meetings of the American Political Science Association and, judging from the response in these forums, support for the replication standard is strong throughout the discipline. Many insightful

questions and issues have been raised about the specific methods for implementing replication policies. The proposals discussed in this paper have been greatly improved as a result.

Questions and Answers

In the course of numerous conversations about these issues, several questions have been raised and discussed. I list some of these here, along with the most common resolutions.

Will a replication standard reduce incentives for individual investigators to collect large and difficult data sets?

Investigators receive recognition for collecting data and making them available to the scholarly community. This recognition is in the form of citations to the data and to the author’s articles, acknowledgments for the author’s help, and promotion, tenure, and raises. Scholarly work is judged by its contribution, so making an article more important by contributing a replication data set can only enhance the recognition that the author receives. The risk is not having one’s ideas stolen, a familiar but largely unfounded fear most of us have experienced while writing dissertations; the much larger risk—indeed a risk with a high probability—is having one’s publications ignored. Submitting a replication data set can significantly decrease that probability.

Moreover, as discussed above, information gathered but ultimately not used in the article need not be included in the replication data set; only those variables and observations necessary to replicate published results need be submitted. If there happens to be new information in the variables and observations submitted, the author will have a substantial head start in extracting such information. In most cases, nearly two years elapse from completion of the article to final publication. If this is not sufficient lead time in a given instance, the author can still submit the data set and choose an embargo for a speci-

fied period, or even commit to submitting it at a specified future date.

Implementing the replication standard will make much more data available through public archives. Won’t an unintended consequence of this proposal be that future scholars will spend most of their time analyzing existing data rather than collecting new data, spending time in the computer lab rather than in the field?

Experience suggests just the opposite. When the ICPSR was founded, and later expanded, the amount of publicly available data increased dramatically. However, content analyses indicate that many more articles containing original data were published during this time (King 1991). Hence, it appears that increasing the availability of original data inspires other scholars to collect original data themselves. In fact, one learns so much by replicating existing research that collecting new data, by following the successful procedures developed in past research, should be made much easier.

Wouldn’t it be better if all journals and book series adopted exactly the same policy at the same time?

There might be some advantages to coordination, but the reason we have different journals in the first instance argues against waiting. Each journal has a different constituency, follows different style manuals, has different quality standards, different editorial boards, different editors, different reviewers, different methodological styles, different copyeditors, and encourages a different mix of substantive articles. It should not be surprising or troubling if different journals adopted slightly different policies regarding replication, or adopted them in accordance with different timetables.

If our journal requires adherence to the replication standard, won’t authors send work elsewhere or not publish articles and save their work until a book manuscript is ready?

This may be true for some authors, but it has not been the experience at the journals that have already adopted this policy. Moreover, many book presses are adopting the same policy, and no one can recall a scholar turning down NSF

funds to avoid this rule. Once the replication policy is adequately communicated and explained to authors, they will likely understand that it is in their interest. It is also clearly in the interest of journals to have their articles cited, and thus to follow the replication standard. Moreover, failing to follow this standard would be far more unfair to potential readers, and more damaging to the profession.

If I give you my data, isn't there a chance that you will find out that I'm wrong and tell everyone?

Yes. The way science moves forward is by making ourselves vulnerable to being wrong. Ultimately, we are all pursuing a common goal: a deeper understanding of government and politics. Thus, we must give others the opportunity to prove us wrong. Although being criticized is not always pleasant, it is unambiguous evidence of being taken seriously and making a difference. Again, being ignored—the fate of over 90% of all political science publications—is the much more serious risk.

Shouldn't editors collect replication data sets to guarantee that they have been submitted?

This is a possibility, but editors are no more professional archivists than are authors. Editors might as well avail themselves of PAVA or the ICPSR. I also do not think verification is necessary, since any discrepancies in the public record will be corrected by future researchers.

If everyone starts submitting replication data sets, won't archives rapidly begin to be filled with junk?

This is extremely unlikely. Unlike most data sets submitted to public archives, replication data sets will have been filtered by the peer review process, and will likely be verified by future researchers. Thus, public archives can harness the regular scientific process to build value into their collections. Moreover, the average size of a replication data set in political science is under a megabyte. Even if as many as 100 replication data sets were submitted to an archive each year, approximately \$600 for a gigabyte of hard disk space and a con-

nection to the Internet would easily accommodate all submissions for well over a decade.

If submitting replication data sets is in the interest of individual investigators, why do we need journals and book presses to require their submission?

We shouldn't need laws when custom will do, but experience in our discipline and most others indicates that this collective goods problem cannot be solved in practice without some policy change. See Dewald, Thursby, and Anderson 1985, Boruch and Cordray 1985; 209–210, and Fienberg et al. 1985 for more detailed justifications.

Why are we worrying ourselves with what might be called "duplication" of existing research? Isn't the more important question actual replication where the same measurements are applied to new substantive areas, countries, or time periods?

Good science requires that we be able to reproduce existing numerical results, and that other scholars be able to show how substantive findings change as we apply the same methods in new contexts. The latter is more interesting, but it does not reduce the necessity of the former. In fact, we can encourage scholars to pursue replication in new contexts if they can be more certain of present results. Better knowledge of existing results, through the distribution of replication data sets, will also enable easier adaptation of existing research methods and procedures to new contexts. Moreover, a standard practice in estimating causal effects is to make one change at a time so we are able to relate individual changes to specific effects and judge each effect in isolation.

Notes

1. This paper has benefited immeasurably from innumerable conversations I have had with many groups and individuals. For many helpful comments on previous versions of this paper, I am especially grateful to Jim Alt, Neal Beck, Robert X. Browning, John DiIulio, John Green, Matthew Holden, Gary Klass, David Laitin, Malcolm Litchfield, Ken Meier, Jonathan Nagler, Bob Putnam, Richard Rockwell, Phil Schrod, and Sid Verba. I also thank the National Science

Foundation for grants SBR-9321212 and SBR-9223637, and the John Simon Guggenheim Memorial Foundation for a fellowship.

2. In some cases, the replication standard refers to running the same analyses on the same data to get to the same result, what should probably be called "duplication" or perhaps "confirmation." For other articles, the replication standard actually involves what is more popularly called "replication"—going back to the world from which the data came and administering the same measurements, such as survey instruments. Since this involves different numerical results, due to a change in time, place, or subjects, we would not expect to duplicate the published results exactly; however, this procedure confers the scientific benefit of verifying whether the substantive conclusions are systematic features of the world or idiosyncratic characteristics of the last author's measurement. In this article, I follow the common current practice in the social sciences of referring to all of these procedures as "replication."

3. For other work on replication, from the perspectives of other social sciences, ethical considerations, the advantages to science, incentives of investigators, and other concerns, see Sieber 1991, Ceci and Walker 1983, Neuliep 1991, Fienberg et al. 1985, and Feigenbaum and Levy 1993.

4. It is worth mentioning that I doubt fraud is much of a problem in political science research. It probably exists to some degree, as it does in every other discipline and area of human endeavor, but I see no evidence that it is anything but extremely rare.

5. At present—before the replication standard has been widely adopted in the discipline—replicating published articles is frequently difficult or impossible. However, other procedures can be used in the interim by teachers of quantitative methods classes. For example, I try to have students submit a draft of their term papers about halfway through the semester (usually with data analyses but few written pages), along with a disk containing a replication data set. These are then given randomly to other students in the class. The next week's assignment is to replicate their classmate's project. In most cases, the replicator and the original author learn a lot about the data, methods, and process of research.

6. Another occasion the replication standard can be implemented is during the peer review process. Reviewer of journals and book manuscripts also should verify the existence of a footnote indicating in which archive a replication data set has been deposited. Since the footnote affects the magnitude of the contribution of the scholarly work, commenting on this is probably the reviewer's responsibility. Moreover, suggesting to authors in reviews that they include this footnote, and deposit their data in a public archive, will help remind authors and perhaps editors of this useful method of scholarly contribution. Journals also could include requests to evaluate the footnote on replication when they send out their request for a review.

References

- Boruch, Robert F., and David S. Cordray. 1985. "Professional Codes and Guidelines in Data Sharing." In *Sharing Research Data*, Stephen E. Fienberg, Margaret E. Martin, and Miron L. Straf, eds., 199–223. Washington, DC: National Academy Press.
- Ceci, Stephen, and Elaine Walker. 1983. "Private Archives and Public Needs," *American Psychologist* 38 (April):414–23.
- Dewald, William G., Jerry G. Thursby, and Richard G. Anderson. 1986. "Replication in Empirical Economics: The *Journal of Money, Credit, and Banking* Project," *American Economic Review* 76, 4 (September):587–603.
- Diamond, A. M., Jr. 1988. "The Polywater Episode and the Appraisal of Theories." In *Scrutinizing Science*, A. Donovan et al., eds. New York: Kluwer Academic Publishers.
- Feigenbaum, Susan, and David M. Levy. 1993. "The Market for (Ir)reproducible Econometrics (with comments and response)," *Social Epistemology* 7, 3 (July–September):215–92.
- Fienberg, Stephen E., Margaret E. Martin, and Miron L. Straf, eds., *Sharing Research Data*. Washington, DC: National Academy Press.
- Griffin, Larry, and Charles C. Ragin. 1994. "Formal Methods of Qualitative Analysis," a special issue of *Sociological Methods and Research* 23, 1.
- Hamilton, David P. 1991. "Research Papers: Who's Uncited Now?" *Science*, 4 January, p. 25.
- King, Gary. 1991. "On Political Methodology," *Political Analysis* 2: pp. 1–30.
- King, Gary, John M. Bruce, and Michael Gilligan. 1993. "The Science of Political Science Graduate Admissions," *PS: Political Science & Politics* 4 (December):772–78.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Leimer, D. R., and Lesnoy, S. D. 1982. "Social Security and Private Saving," *Journal of Political Economy* 90:606–29.
- Neuliep, James W., ed. 1991. *Replication Research in the Social Sciences*. Newbury Park, CA: Sage Publications.
- Pendlebury, David. 1994. (Institute for Scientific Information.) Telephone conversation. 8 September.
- Siebert, Joan E., ed., 1991. *Sharing Social Science Data: Advantages and Challenges*. Newbury Park, CA: Sage Publications.

About the Author

Gary King is professor of government in the department of government at Harvard University. He is currently finishing a book entitled *A Solution to the Ecological Inference Problem* and, with Brad Palmquist, is completing (and will deposit with the ICPSR!) a large collection of merged precinct-level election and census data. King can be reached at Littauer Center, North Yard, Cambridge, Massachusetts 02138; e-mail: gking@harvard.edu; phone (617) 495-2027. The current version of this manuscript is available via gopher or anonymous FTP from haavelmo.harvard.edu.

Replication, Verification, Secondary Analysis, and Data Collection in Political Science

Paul S. Herrnson,¹ *University of Maryland, College Park*

Political scientists believe they ought to promote research as a social enterprise and develop a stronger sense of community within the discipline. They disagree on how to foster community and on the best way to promote research. The position taken in "Replication, Replication" (King 1995) is that these goals can be achieved by requiring researchers who collect data to surrender the products of their labor to others without compensation.

This data relinquishment or verification policy is mistakenly referred to as a replication policy, and would harm researchers, journals, the discipline, and the acquisition of knowledge about politics. A sense of community is bolstered more effectively by encouraging true replication of studies, which includes independent data collection, and by fostering arrangements of shared data that benefit those who collect data and those who seek to reanalyze them. These

approaches will result in more diverse, more interesting, and better scholarship in political science.

Defining Replication

"Replication, Replication" misstates the meaning of *replication* in the physical and life sciences, as well as political science. *Replication* is not the same as *reanalysis*, *verification*, or *secondary analysis*. The four terms have very different meanings.

A *reanalysis* studies the same problem as that investigated by the initial investigator; the same data base as that used by the initial investigator may or may not be used. If different, independently collected data are used to study the same problem, the reanalysis is called a *replication*. If the same data are used, the reanalysis is called a *verification*. In a *secondary analysis*, data collected to study one set of problems are used to study a different problem. Sec-

ondary analysis frequently, but not necessarily, depends on the use of multipurpose datasets. Data sharing is essential for all verifications and all secondary analyses; it may or may not be involved in replications (*Italics from the original. Committee on National Statistics, 1993, 9).*

Replication, verification, and secondary analysis are used for different purposes and require different kinds of activities (Neuliep 1991; Siebert 1991). Replication repeats an empirical study in its entirety, including independent data collection. It enables a researcher to comment on whether data used in an original study were collected properly or whether generalizations supported under one set of conditions are also supported under others. Replications increase the amount of information for an empirical research question and increase the level of confidence for a set of empirical generalizations (Rosenthal 1991, 2).