

Research Preregistration in Political Science: The Case, Counterarguments, and a Response to Critiques

James E. Monogan III, *University of Georgia*

ABSTRACT

This article describes the current debate on the practice of preregistration in political science—that is, publicly releasing a research design before observing outcome data. The case in favor of preregistration maintains that it can restrain four potential causes of publication bias, clearly distinguish deductive and inductive studies, add transparency regarding a researcher's motivation, and liberate researchers who may be pressured to find specific results. Concerns about preregistration maintain that it is less suitable for the study of historical data, could reduce data exploration, may not allow for contextual problems that emerge in field research, and may increase the difficulty of finding true positive results. This article makes the case that these concerns can be addressed in preregistered studies, and it offers advice to those who would like to pursue study registration in their own work.

A conversation is emerging in political science about the merits of study registration and how well the concept fits with research in the discipline. Registering a study means that before observing outcome data, researchers craft and publicly release their plan for data analysis that they believe will offer the most honest means of testing a hypothesis. Proponents argue that study registration can restrain publication bias and distinguish deductive from inductive research. The concept is not entirely new: clinical trials in biomedical research often are preregistered, and several regulators including the US Food and Drug Administration require it (albeit compliance often lags behind mandates; see Prayle, Hurley, and Smyth 2012).

The idea of preregistration emerged as part of the trend toward increased transparency in political science. More journals are requiring writers to share replication data and more researchers are volunteering supplemental information. The American Political Science Association is weighing Guidelines for Data Access and Research Transparency (DA-RT) in the quantitative and qualitative traditions as a means to promote openness, about which *PS: Political Science and Politics* published a symposium

(Lupia and Elman 2014). Leaders in the discipline have developed transparency-focused innovations including resources for free data sharing (King 2007), an emphasis on replication as an important activity (King 1995), a case to keep lab books public (Lupia 2008), and initiatives for digitally preserving at-risk data (Gutmann et al. 2009). These efforts add clarity to research and allow others to complete replication and meta-analysis projects. This article presents the case for preregistration as a next step in transparency, as well as the current controversy about the tradeoffs of registration.

A REMEDY FOR PUBLICATION BIAS

Several scholars argue that study registration can be useful in social science (Asendorpf et al. 2013; Casey, Glennerster, and Miguel 2012; Chambers 2013; Humphreys, de la Sierra, and van der Windt 2013; King et al. 2007, 2009; Monogan 2013). Chiefly, preregistration can restrain publication bias, which is the tendency for positive results to be disproportionately prone to publication relative to null findings. There is evidence of publication bias in political science articles (Gerber and Malhotra 2008; Gerber et al. 2010). Four possible causes of this pattern include a journal's rejection of null findings, an author's self-selecting to submit only those studies with significant results, an author's expansion of samples after failing significance tests, and an author's search

James E. Monogan III is assistant professor of political science at the University of Georgia. He can be reached at monogan@uga.edu.

for specifications that generate significant results (Gerber and Malhotra 2008, 314).

Monogan (2013, 23–4) argues that preregistration can restrain all four causes of publication bias. First, registration would make research design more central to the review process, thereby reducing the importance of significance tests in publication decisions. Some proposals call for a publication decision based strictly on the research design, removing significance tests from consideration (Chambers 2013). Even if the decision is made after observing the results, however, registration would highlight the research design. If several researchers test the same hypothesis differently when the null hypothesis is true, a Type I error occasionally will emerge, which then may be the only published result (Gill 1999). When a sound theoretical idea is tested rigorously yet yields a negative finding, publishing the study may prevent duplicate efforts until a false positive misleads the discipline.

Second, registration would alleviate the problem of null findings that authors never submit (Rosenthal 1979) because it would provide a record of the registered study, even if an article was not written. This record would convey to others that pursuing a research question may not be fruitful. Preregistration also could ameliorate the “file-drawer problem” (Rosenthal 1979) by changing how null findings are perceived. That is, scholars who conduct rigorous research may be willing to submit manuscripts if they expect them to be evaluated for thoroughness rather than significance levels.

Third, concerning the expansion of the sample size to achieve significance, preregistration would signal in advance the appropriate sample size for a given research question. Adding data is the least problematic source of publication bias because larger samples reduce the scope for fishing (i.e., model manipulation for a desired result) (Humphreys, de la Sierra, and van der Windt 2013, 6–7). However, the analysis can be monitored until a positive result emerges before stopping the data collection. Simmons, Nelson, and Simonsohn (2011, 1362) therefore argue that the rule for terminating data collection should be decided in advance. Preregistered information (e.g., the target sample size) indicates to readers that the result was the consequence of informed planning.

Fourth, preregistration can prevent fishing because the investigator must specify the model in advance. As Humphreys, de la Sierra, and van der Windt (2013) argue, even nonbinding registration can communicate to readers whether a study adhered exactly to the preregistered design, deviated on grounds defended by the researchers, or was not preregistered and hence could be exploratory work. By sorting out the best specification using theory and past work in advance, a researcher can commit to the results of a well-reasoned model. Simmons, Nelson, and Simonsohn (2011, 1359) define “researcher degrees of freedom” as the investigator’s discretion in choice of dependent variables, covariates, and sample size. With only a few degrees of freedom, absurd results can be manufactured. The findings from preregistered research should be trustworthy because the investigator eliminated those researcher degrees of freedom.

BENEFITS OF TRANSPARENCY BEYOND PUBLICATION BIAS

In addition to these four ways that preregistration can diminish publication bias, it can help the discipline in three more ways. First, registration distinguishes deductive and inductive research. The act of study registration lends itself more naturally to deductive research in which a theory is formulated and then hypothesis

tests are developed. However, nothing would prevent researchers from stating up front that they will learn from data and explaining how they will do this. Inductive studies are valuable; however, if an author learns from the data and crafts the article to appear as if it tests a theory, then the discipline is being misled about the nature of the study. The added transparency can clarify for the discipline how a study should be evaluated. If a researcher wants readers to be certain about whether a study was deductive or inductive, providing proof by preregistering can reduce erroneous perceptions.

Second, without transparency, researchers’ motivations can be misjudged. Some findings may prompt readers to accuse an author of motivated reasoning, which may be unintentional. As Feynman (1999, 212) advised scientists, “You must not fool yourself—and you are the easiest person to fool.” For instance, a reader who suspects that a political viewpoint led an author to measure variables and specify the model in such a way to produce a desired result may not accept the finding—even if it was obtained honestly. Preregistration allows researchers to declare measurement and specification decisions without reference to outcomes, thereby signaling to readers that fishing is impossible. From a disciplinary perspective, when investigators register their designs, the ability to fish is eliminated—even if any motivated reasoning is unintended.

Third, Casey, Glennerster, and Miguel (2012, 1758–9) argue that preregistration can be useful in policy studies because it ties the researchers’ hands when they “may face professional incentives to affirm the priors of their academic discipline or the agenda of donors and policy makers.” They show how useful preregistration can be in practice by creating a “pre-analysis plan” for an investigation of a randomly assigned governance program in Sierra Leone. The analysis revealed only short-run treatment effects, which is in contrast to a prevailing notion that governance programs can have sustainable effects. With deviations from the pre-analysis plan, the researchers could generate misleading results with either positive or negative treatment effects (Casey, Glennerster, and Miguel 2012, 1804–5). Study registration can liberate researchers who simply want trustworthy results, even in the face of governmental, financial, or academic pressure for a specific finding.

COUNTERARGUMENTS ON PREREGISTRATION

In contrast to this case in favor of study registration, several arguments call for skepticism in making preregistration a new norm in political science. First, Anderson (2013) emphasizes that registration is most useful for studies that collect original data; however, in the analysis of historical data, preregistering cannot send as clear a signal. Anderson also contends that discouraging reports of all observed empirical relationships can be detrimental for scientific development (Kuhn 1962). Finally, Anderson suggests replication as an alternative to registration; that is, beyond the quality enforcement that occurs from replication projects, the quality of published articles improves when replication materials are furnished even without attempts at replication (King 1995). Thus, preregistration cannot be a substitute for sharing replication information.

Second, Gelman (2013) generally supports preregistration by proposing that this step would reduce the number of printed results that turn out to be false. He does have concerns, however: it would be problematic if preregistration led to robotic data

analysis in which the simple evaluation of hypotheses came at the expense of broader data exploration. Furthermore, studies should present visualizations of data and embrace the uncertainty of estimation.

unpublished studies are sharing data. Thus, preregistration symbiotically supports the sharing of replication information.

Second, regarding the argument that preregistration may stifle reports from broader data exploration, Gelman (2013) notes

Another concern that Laitin raises is that many of the most important developments in knowledge came from inductive findings, and many studies evolve in a cycle between theory testing and learning from data.

Third, Laitin (2013) argues that additional moves in transparency are worthwhile in political science, but certain issues should be considered before adopting preregistration as the next step. One concern is that study registration works well for clinical research but that the incentives are stronger in that field than in political science. Clinical researchers often work in labs that are funded by companies that want to market proposed treatments, and the costs of Type I errors are steeper in clinical research. Also, context matters for field-based political studies in ways that do not emerge in the clinical setting. Another concern that Laitin raises is that many of the most important developments in knowledge came from inductive findings, and many studies evolve in a cycle between theory testing and learning from data. Furthermore, Laitin espouses that the review process and replication projects can serve as a better means for ensuring reasonable results. Finally, Laitin is concerned that preregistration may be adopted so zealously that nonregistered studies will be perceived as inferior.

A fourth argument against preregistration is that finding true positive results can be difficult, and preregistration may increase the difficulty. Many models call for diagnostic assessments after estimation, and evidence that assumptions have been violated calls for remedial measures. Without these corrections, a model's findings can be misleading. For example, suppose an investigator preregistered a study with a plan to fit a regression model in which all predictors were held in linear form. If later diagnostics indicated that a nonlinear functional form was necessary, then the simple linear results would be misleading. Indeed, inaccurate functional forms can produce false-negative findings. Therefore, a full prescription of the data analysis prior to observing the outcome may run the risk of ineffective modeling.

RESPONSES TO THE CRITIQUES

This section responds to the counterarguments and renews the case that preregistration is appropriate for political science. The first concern is the notion that the required provision of replication information is a better path to transparency. Study registration should not replace the sharing of replication data but rather enhance it. Our discipline's commitment to replicability is critical to open knowledge. There may not be as much acceptance in the sharing of replication information as is ideal (Anderson 2013, 39), but an increase in journal requirements implies that moves are being made in the right direction. As long as journals require public sharing of replication data, study registration can further transparency because it also requires an author to illustrate more about the research process. Furthermore, releasing pre-outcome data as a part of study registration increases sharing because even

that it does not necessarily preclude such activity. For instance, data visualization still is possible when completing the work of a preregistered study (Gelman 2013, 40). Anderson (2013) and Gelman (2013) rightly observe that researchers who conduct preregistered studies should be wary of completing their analysis wearing blinders. In any publication regime that includes preregistration (voluntary or required), reporting auxiliary findings from data should be encouraged—provided that the central hypothesis is evaluated using the registered design. Because much learning occurs through a cycle of deductive and inductive inference (Laitin 2013, 44), describing additional empirical results as observations from data—rather than hypothesis tests—can provide useful data-oriented insights.

Regarding the contention that political studies fundamentally differ from medical research, Laitin (2013, 42–3) raises several valid points. The structure of medical research, in which treatment manufacturers fund labs, places clinical researchers out on a limb that null findings can sever. At the same time, Type I errors in clinical research allow treatments that are ineffective or not worth the added patient risk to become pervasive. Hence, investigators must tie their hands before evaluating a treatment in clinical studies. Although incentives are stacked more in medical research, political scientists are expected to generate findings that advance knowledge; null findings rarely are regarded as novel—even if they should be in some cases.

Laitin's argument about the distinction from biomedical research focuses not only on the greater need in clinical trials for irreversible early commitments but also on the lower costs of preregistration in clinical research. Unexpected implementation difficulties are unlikely to occur in laboratories. By contrast, field experiments and observational data often require attention to social contexts and the role of external events on political behavior. Requiring political scientists to anticipate all contingencies could be unreasonable. Any registration regime must consider this point: there may be legitimate reasons why the initial plan was altered in a registered study—for example, an unanticipated impediment to gathering data as originally intended or a local crisis that changed whether individuals would respond to a treatment. Among the proposals for political registries, one idea is for registry staff to rate a study's compliance with its design. Cases in which there are easily justified deviations may be rated differently than those for which the investigator successfully anticipated any implementation difficulties and necessary workarounds. According to Laitin's point, it would be essential for studies with well-reasoned justifications for deviation to be regarded as highly as those with no deviation whatsoever. Although this point highlights the need for a particular caveat within a preregistration

program, the consideration of this provision should allow political research to follow its necessary course while still expanding transparency.

Regarding the argument that preregistration does not fit some studies as well as others, Anderson (2013) reiterates that preregistration is more informative for some than for others. Specifically, with historical data, a scholar may have glimpsed the data before the study registration. Laitin (2013) is concerned that overzealous support for preregistration might lead to the perception that nonregistered studies are inferior. For these reasons, it is critical in any registration regime that scholars have the option to briefly explain why preregistering their study would not be effective. In historical analysis, the self-evident reason why a scholar might not preregister is that there is no way to guarantee that a preliminary analysis was not conducted. In an inductive study, scholars could register the process of learning from data. Alternately, they could state that they are not holding a hypothesis to scrutiny, so reporting findings at the end of the process is valid without registration. No policy should threaten the diversity of the discipline's studies. Making preregistration a feature of political research, however, would identify studies that conduct deductive tests.

Finally, regarding the argument that finding true positive results can be difficult, it is worth reiterating that registration is not a permit to work wearing blinders. Authors should consider details from exploring data, running diagnostics, and responding to reviews. Previous arguments maintained that changes from the preregistered plan should be acceptable if the findings of the original design are reported with justification for the changes (Monogan 2013, 24–5). If journal editors prefer to print revised results, then placing the original estimates and the justification on the registry page allows readers to see the entire process from design to final result.

The issues could be guest-edited if journal editors preferred not to develop a two-track system.

For journals that implement preregistration procedures, the online appendix accompanying this article lists several proto-registries that editors can rely on as hosts for public posting of research designs. If single-blind review is permissible, authors can easily include registration information from a proto-registry in the manuscript. To include registration information under double-blind review, some third-party registries allow investigators' names to remain temporarily anonymous as a solution (see the online appendix). In fact, "The 2011 Debt Ceiling Controversy and the 2012 US House Elections," also in this issue of *PS: Political Science and Politics*, was subject to double-blind review; as such, blinded preregistration materials were shared with reviewers. (Editors had access to the non-anonymous materials.) At present, editors and reviewers must assess adherence to the original research design; however, there are proposals to create a sustainable general registry with staff who could verify the degree of compliance. Creating a more comprehensive registry would be a major step forward for transparency in social research. Yet, even incremental steps by journals could assess whether the broader discipline would buy into a preregistration regime by giving the opportunity for ground-up acceptance of a new paradigm.

To that end, the real debate will emerge when more political scientists use firsthand experience to evaluate study registration in practice. For those considering self-registration of their research, it is worth considering that preregistration is more beneficial to readers under certain conditions than in others. A few cases in which registration would be less useful include theory-building projects, whether positive or normative. They would be difficult to register because authors could not guarantee that they had not already worked on developing the argument.

Editors could give authors the option of submitting a research design before the outcome variable is observed, which would allow pre-acceptance of an article before seeing the results—provided the research design is followed precisely.

DIRECTIONS FORWARD

This article presents the current debate on preregistration in political science and makes the case for registering research designs to restrain publication bias and to distinguish deductive from inductive studies. As this debate expands, several intermediate steps can be taken. At the journal level, editors could allow a new publication track that is similar to a policy implemented at the journal *Cortex* (Chambers 2013). Editors could give authors the option of submitting a research design before the outcome variable is observed, which would allow pre-acceptance of an article before seeing the results—provided the research design is followed precisely. Even without adopting a different track, several journals now acknowledge open-research practices such as preregistration by placing badges on publications; editors could adopt this practice. Another journal option would be to publish special issues on topics that can be studied easily with preregistered designs, such as elections, policy studies, and experimental research. For the special issue, only those studies with a preregistered design in a journal-approved forum would be considered.

Studies using big data often will not gain substantially from registration, particularly if they use existing information (e.g., scraped text) or if the endeavor is to learn inductively from such data (e.g., to characterize the content of the text). In any study using existing information, whether past surveys, time series of economic data, or historical records, authors unfortunately cannot provide proof that they did not glance at the information beforehand; therefore, an author could register a study that uses existing data. However, doing so does not provide as much information to the reader as a researcher who posts a time-stamped design before entering the field for a survey on designated dates. Moreover, scholars conducting inductive studies certainly could register a plan for how they want to learn from data. However, when exploring the data, flexibility and creativity may lead to new and unexpected findings in a way that using only a prescribed plan may not produce.

By contrast, in certain types of studies, there is a clear opportunity to offer more transparency to readers by preregistering. In deductive studies that test one or a few hypotheses using original

data, readers can see that authors have tied their hands by registering before data collection. Therefore, when researchers conduct an experiment, field a survey, or plan to study an upcoming election, they can definitively identify for readers that the study is deductive by releasing all elements of the design before collecting the data. Policy studies, studies of election returns, and lab-based studies of reactions to psychological or economic treatments are all substantive areas in which preregistration has high value. In fact, the article titled “The 2011 Debt Ceiling Controversy and the 2012 US House Elections” in this issue demonstrates how registration can be implemented in an election study.

Particularly in these cases, preregistering a study in practice is helpful for applied researchers to understand the process and develop an informed opinion. The registries in the appendix to this article provide tools that are available for authors to use, and several studies provide examples of self-registration in practice (Casey, Glennerster, and Miguel 2012; Humphreys, de la Sierra, and van der Windt 2013; King et al. 2007; Monogan 2013). In conference discussions, many editors have expressed receptiveness to the self-initiative from authors for including preregistered research in journals. A greater number of printed preregistered studies will provide our discipline with a broader view of the tradeoffs in this step in transparency.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <http://dx.doi.org/S1049096515000189>.

ACKNOWLEDGMENTS

For helpful assistance, I thank Phillip J. Ardoyn, Jamie L. Carson, Keith L. Dougherty, Kevin M. Esterling, N. Macartan Humphreys, Anthony J. Madonna, Patrick McNeal, Ryan T. Moore, Brian A. Nosek, Keith T. Poole, and several anonymous reviewers. A previous version of this research was presented at the “Effects of the 2012 Presidential Election Conference,” organized by Keith T. Poole and Jamie L. Carson, on November 30, 2012, in Athens, Georgia, and at the “St. Louis Area Methods Meeting” on April 19, 2013, in Iowa City, Iowa. ■

REFERENCES

- Anderson, Richard G. 2013. “Registration and Replication: A Comment.” *Political Analysis* 21 (1): 38–9.
- Asendorpf, Jens B., Mark Conner, Filip de Fruyt, Jan de Houwer, Jaap J. A. Denissen, Klaus Fiedler, Susann Fiedler, et al. 2013. “Recommendations for Increasing Replicability in Psychology.” *European Journal of Personality* 27 (2): 108–19.
- Casey, Katherine, Rachel Glennerster, and Edward Miguel. 2012. “Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan.” *Quarterly Journal of Economics* 127 (4): 1755–812.
- Chambers, Christopher D. 2013. “Registered Reports: A New Publishing Initiative at *Cortex*.” *Cortex* 49 (3): 609–10.
- Feynman, Richard P. 1999. *The Pleasure of Finding Things Out*. New York: Basic Books.
- Gelman, Andrew. 2013. “Preregistration of Studies and Mock Reports.” *Political Analysis* 21 (1): 40–1.
- Gerber, Alan S., and Neil Malhotra. 2008. “Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals.” *Quarterly Journal of Political Science* 3 (3): 313–26.
- Gerber, Alan S., Neil Malhotra, Conor M. Dowling, and David Doherty. 2010. “Publication Bias in Two Political Behavior Literatures.” *American Politics Research* 38 (4): 591–613.
- Gill, Jeff. 1999. “The Insignificance of Null Hypothesis Significance Testing.” *Political Research Quarterly* 52 (3): 647–74.
- Gutmann, Myron P., Mark Abrahamson, Margaret O. Adams, Micah Altman, Caroline Arms, Kenneth Bollen, Michael Carlson, et al. 2009. “From Preserving the Past to Preserving the Future: The Data-PASS Project and the Challenges of Preserving Digital Social Science Data.” *Library Trends* 57 (3): 315–37.
- Humphreys, Macartan, Raul Sanchez de la Sierra, and Peter van der Windt. 2013. “Fishing, Commitment, and Communication: A Proposal for Comprehensive Nonbinding Research Registration.” *Political Analysis* 21 (1): 1–20.
- King, Gary. 1995. “Replication, Replication.” *PS: Political Science and Politics* 28 (3): 444–52.
- . 2007. “An Introduction to the Dataverse Network as an Infrastructure for Data Sharing.” *Sociological Methods and Research* 36 (2): 173–99.
- King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T. Moore, Clayton Nall, Nirmala Ravishankar, et al. 2009. “Public Policy for the Poor? A Randomized Assessment of the Mexican Universal Health Insurance Programme.” *Lancet* 373 (9673): 1447–54.
- King, Gary, Emmanuela Gakidou, Nirmala Ravishankar, Ryan T. Moore, Jason Lakin, Manett Vargas, Martha Maria Tellez-Rojo, et al. 2007. “A ‘Politically Robust’ Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program.” *Journal of Policy Analysis and Management* 26 (3): 479–506.
- Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Laitin, David D. 2013. “Fisheries Management.” *Political Analysis* 21 (1): 42–7.
- Lupia, Arthur. 2008. “Procedural Transparency and the Credibility of Election Surveys.” *Electoral Studies* 27 (4): 732–9.
- Lupia, Arthur, and Colin Elman (eds.). 2014. “Symposium: Openness in Political Science: Data Access and Research Transparency.” *PS: Political Science and Politics* 47 (1): 19–83.
- Monogan III, James E. 2013. “A Case for Registering Studies of Political Outcomes: An Application in the 2010 House Elections.” *Political Analysis* 21 (1): 21–37.
- Prayle, Andrew P., Matthew N. Hurley, and Alan R. Smyth. 2012. “Compliance with Mandatory Reporting of Clinical Trial Results on ClinicalTrials.gov: Cross-Sectional Study.” *British Medical Journal* 344: d7373.
- Rosenthal, Robert. 1979. “The ‘File-Drawer Problem’ and Tolerance for Null Results.” *Psychological Bulletin* 86 (3): 638–41.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. “False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant.” *Psychological Science* 22 (11): 1359–66.