

Registration and Replication: A Comment

Richard G. Anderson

Research Division, Federal Reserve Bank of St. Louis, St. Louis, Missouri, and
Management School, University of Sheffield, Sheffield, UK
e-mail: randerson@stls.frb.org

Edited by R. Michael Alvarez

Social scientists long have debated how closely their research methods resemble the pure, classical ideal: sequentially (and absent pejorative “data snooping”) formulate a theory; develop empirically falsifiable hypotheses; collect data; and conduct the appropriate statistical tests. Meanwhile, in private quarters, they acknowledged that true research seldom proceeds in this fashion. Rather, in the event, the world is observed, data are collected, hypotheses formed, tests conducted, more data collected, and hypotheses revised (e.g., the classic Bernal 1974). Results are collected by the field’s scientists into a body of knowledge that defines “known science” and sets the accepted boundaries for future research. Kuhn (1970) labeled this a *paradigm*. Occasionally, he argued, results appear that lie outside the bounds of the extant paradigm—then, innovation occurs.

The articles included in this issue’s symposium discuss “registration” of empirical studies. The purpose is to reduce “publication bias,” that is, to prevent the scientific equivalent of schoolboy cheating: reporting tests of hypotheses that became evident only after the data were in hand. Registration has little power against the type of research fraud that is discovered from time to time. There are costs, however, when scientists operating within an accepted paradigm discourage researchers from exploring and reporting any/all relationships and correlations in a data set.

The articles acknowledge that registration likely is of little value for studies using published historical data. A story illustrates. Early in my academic career, an admired senior professor who always appeared to have National Science Foundation (NSF) funding confided to me: “Never submit a proposal for a project that you have not already completed.” His wisdom was genius: submitting a proposal for a project already completed permitted formulating appealing hypotheses within the prevailing paradigm, at least one of which was guaranteed to not be rejected. Registration has no power against such schemes; its value is limited to studies in which a large part of the research effort (and expense) is devoted to the collection of original data.

Replication is a well-known alternative, cherished by those who wish to fully understand published studies but despised by those who fear discovery of error, or worse, fraud. King (1995) eloquently summarizes the arguments in favor of replication as a quality enforcement mechanism, and King (2006) discusses replication as a springboard for future research.

In the first study of its kind in economics, funded by the NSF, Dewald, Thursby, and Anderson (1986) tested the replicability of a set of papers published by, or submitted to, the *Journal of Money, Credit, and Banking*. The study suggested three principal conclusions: (1) once an article was printed, the likelihood of obtaining from authors the materials necessary to replicate published results was low; (2) when replication materials were requested prior to publication, a much higher proportion of authors could supply adequate materials; and (3) the likelihood that a researcher could successfully replicate published results based on the authors’ accounts and descriptions was extremely low. (These results themselves were later replicated in Anderson and Dewald

Author’s note: The author wishes to thank the Leverhulme Trust, London, for financial support at the University of Sheffield during the preparation of this essay.

1994.) A number of critics (and perhaps most readers) dismissed our findings as irrelevant. After all, who has gained fame and fortune by reproducing the results of others?

Suppose, as discussed in this symposium, projects are preregistered. What is to be made of studies whose results differ from those proposed in the registration? Does failure to obtain at least one of the registered results signal a researcher's lack of ability to formulate a broad project within the current paradigm—and hence inadequate training or ability? Do attempts to publish results not included in the registration signal data mining? Kuhn (1970) suggests that discovery of results contrary to those anticipated by the current paradigm is the fundamental channel through which innovation occurs—if so, should the publication of results contrary to those in the registered design (presumably built to lie within the extant paradigm) be lauded? But, if such results are to be lauded, then what if any penalties are to be imposed on scientists whose results differ from their registered design? Absent penalties, does registration have any power to affect scientific research?

So what remains? I have argued elsewhere that the best system is peer-level pressure that presses authors to submit replication materials, including data and programs, to archival systems operated by third parties (Anderson 2006; Anderson et al. 2008). Further, Dewald, Thursby, and Anderson (1986) and King (1995) argue that the quality of published articles improves when authors prepare replication materials *even if the article's results are never replicated by others*. Unfortunately, three decades after Dewald, Thursby, and Anderson (1986) called for higher scientific standards, economics does not have a widely shared scheme for collecting, archiving, and distributing authors' data sets and programs. I do not see preregistration as a substitute.

References

- Anderson, Richard G. 2006. Replicability, real-time data, and the science of economic research: FRED, ALFRED, and VDC. *Federal Reserve Bank of St. Louis Review* 88(1):81–94.
- Anderson, Richard G., and William G. Dewald. 1994. Replication and scientific standards in applied economics a decade after the *Journal of Money, Credit, and Banking* project. *Federal Reserve Bank of St. Louis Review* 76(6):79–83.
- Anderson, Richard G., William H. Greene, B. D. McCullough, and H. D. Vinod. 2008. The role of data and program code archives in the future of economic research. *Journal of Economic Methodology* 15(1):99–119.
- Bernal, J. D. 1974. *Science in history*, 3rd ed. (1st ed., 1954). Cambridge, MA: MIT Press.
- Dewald, William G., Jerry G. Thursby, and Richard G. Anderson. 1986. Replication in empirical economics. *Journal of Money, Credit, and Banking project. American Economic Review* 76(4):587–603.
- King, Gary. 1995. Replication, replication. *PS: Political Science and Politics* 28(3):119–25.
- . 2006. Publication, publication. *PS: Political Science and Politics* 39(1):444–52.
- Kuhn, Thomas S. 1970. *The structure of scientific revolutions*, 2nd ed. (1st ed., 1962). Chicago: University of Chicago Press.