

Replication Section

Introduction

B.D. McCullough

Department of Decision Sciences Drexel University Philadelphia, PA, USA

Replication, widely acknowledged as the cornerstone of science, brings many benefits to the discipline that embraces it. First among them, successful replication ensures that a particular contribution to the cumulated body of knowledge is correct. Second, unsuccessful replication purges incorrect results from the cumulated body of knowledge. Other benefits are enumerated in the article by King [7].

The economic “science” has denied itself these benefits and perhaps even jeopardized its standing as a science, due its longstanding refusal to embrace replication. This is particularly noteworthy because all available evidence strongly suggests that the economic literature is filled with incorrect results. Dewald et al. ([3], hereafter “DTA”) were able to replicate only 2 of 54 articles, while McCullough et al. [9] were able to replicate only 14 of 62. It appears that the cumulated body of economic knowledge may be filled with incorrect results. Of course, nobody knows which results are correct and which are incorrect. The only prospect for purging these incorrect results is for the principle of replication to be embraced actively by all parts of the economics profession: journals, researchers, professors and students.

The philosopher C.S. Peirce [12] identified the “unlimited community” of investigators as the sole guarantor of the integrity of the scientific method. In order for the scientific method to succeed at uncovering truth, the members of the community must not be selfish – to the contrary, they must identify their own interests with those of the unlimited community. In this way, they can correct each other’s errors and build effectively on each other’s work, and in so doing ultimately arrive at the truth.

Regrettably, many researchers prefer to put their own interests ahead of those of the scientific community, and do not make their research available for inspection by others. While this is “rational” behavior by a utility-maximizing economist (see, e.g., Mirowski and Sklivas [11]), one would hope that institutional constructs would provide incentives to counteract this behavior. Yet too many journals refuse to ensure that the “research” they publish can be checked for accuracy by independent third parties.

Some journals have paid lip service to the scientific method. For example, in response to DTA, the *American Economic Review* adopted a “replication policy” [1] whereby authors would agree to make available their data and code for purposes of

replication – despite the fact that DTA recommended a mandatory archive. Other journals followed suit. McCullough and Vinod [10] showed that these replication policies are largely ignored by authors. In response, the *AER* adopted a mandatory data-code archive [2]. Yet an archive alone is insufficient to ensure that authors participate in the unlimited community. The *Journal of Money, Credit and Banking* attempted had a similar requirement, but neglected an important part of the incentive scheme: they published no replication attempts. Authors, knowing that their errors (or even refusal to supply data and code) would not be sanctioned by the journal, all but ignored this policy. See McCullough et al. [9] for a discussion.

Some researchers actively participate in the search for truth, even at the expense of having their own mistakes uncovered. Martin Feldstein, Harvard Professor, former chief of the President's Council of Economic Advisers is one such. Writing about those who uncovered his programming error, he noted that [5]

They [Leimer and Lesnoy] set an admirable example of the tradition of replication on which all scientific work ultimately rests. As economic research increasingly involves large and complex computer programs to analyze microeconomics datasets or simulate models that cannot be solved analytically, replication studies like that of Leimer and Lesnoy should become increasingly important.

Steve Levitt, University of Chicago Professor, editor of the *Journal of Political Economy* and winner of the Bates Clark Medal has twice been so exposed. A publication in the *AER* was corrected by a graduate student [8] and another in the *Quarterly Journal of Economics* was corrected by pair of Boston Fed economists [6].

In setting such a fine example for others, Feldstein and Levitt understand that, in exposing their own research to correction by others, they are not “risking their reputations.” Rather, their reputations as scientists are enhanced by showing how dedicated they are to the scientific method. Their actions stand in sharp contrast to those who do not embrace replication: those who ensure that their work cannot be inspected by others do not advance knowledge but, rather, retard it.

Lately, several top journals have joined the *AER* in facilitating replication by adopting a mandatory data/code archive, including *Econometrica*, *Review of Economic Studies* and *Journal of Political Economy*. These top journals might not be able to publish all the replication work that their archives inspire. Where might such authors publish their work? Here, in the replication section of the *Journal of Economic and Social Measurement*.

Other journals that stay outside the scientific method desperately need their research to be subjected to replication attempts. Yet, where might such intrepid researchers publish their results? Certainly not in the original journal (which doubtless subscribes to the fiction that *every* article in the journal is 100% correct) but here in the replication section of the *Journal of Economic and Social Measurement*.

Researchers who have analyzed the problem of replication in economics have advocated the use of graduate students as a means of replicating published work in economics journals (DTA; Feigenbaum and Levy [4]). In anticipation of launching

this section, last year I sent e-mails to a dozen professors who teach econometrics and who require their students to write papers, often replication papers. I informed the professors of this section, and told them that they could recommend papers for publication in this section. To date not a single professor has recommended a single paper. So it appears that this section, rather than being regular, will be only occasional.

Here in the inaugural replication section, I am extremely pleased to have an unsuccessful replication by Breusch and Gray, of an article written by Chapman et al. The piece by Breusch and Gray details many of the problems with producing replicable research. It is a fine piece of detective work, very lucidly written, and could not have been done without the full cooperation of Chapman et al. who, together with Feldstein and Levitt, are members in good standing of the unlimited community. The usual response to an unsuccessful replication, e.g., Feldstein and Levitt, is a discussion of the original article and its relation to the replication attempt. Such responses are generally of interest only to readers of the original article. I am particularly pleased to have the response of Chapman and Gray, who have gone a step further: they have produced a thoughtful and considered reflection on the role that their original article and the replication attempt together play in the larger literature and, in so doing, have produced a piece that is important in its own right. Their comment draws some important lessons for the literature, and points toward new areas of research. Professors Breusch and Gray, for reasons that will be obvious to the reader, do not exercise their right of reply.

References

- [1] O. Ashenfelter, R.H. Haveman, J.G. Riley and J.T. Taylor, Editorial Statement, *American Economic Review* **76**(4) (1986), p. v.
- [2] B.S. Bernanke, Editorial Statement, *American Economic Review* **94**(1) (2004), 404.
- [3] W.G. Dewald, J. Thursby and R.G. Anderson, Replication in Empirical Economics: The *Journal of Money, Credit and Banking* Project, *American Economic Review* **76** (1986), 587–603.
- [4] S. Feigenbaum and D. Levy, The Market for (Ir)Reproducible Econometrics, *Social Epistemology* **7**(3) (1993), 215–232.
- [5] M.S. Feldstein, Social Security and Private Saving: Reply, *Journal of Political Economy* (1982), 630–642.
- [6] C. Foote and C. Goetz, Testing Economic Hypotheses with State-Level Data: A comment on Donohue and Levitt (2001), Federal Reserve Bank of Boston working paper, 2005.
- [7] G. King, Replication, Replication, *Political Science & Politics* **Sep 95** (1995), 444–452.
- [8] J. McCrary, Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment, *American Economic Review* **92**(4) (2002), 1236–1243.
- [9] B.D. McCullough, K.A. McGeary and T.D. Harrison, Lessons from the *JMCB* Archive, *Journal of Money, Credit and Banking* (2006), (to appear).
- [10] B.D. McCullough and H.D. Vinod, Verifying the Solution from a Nonlinear Solver: A Case Study (with H.D. Vinod), *American Economic Review* **93**(3) (2003), 873–892.
- [11] P. Mirowski and S. Sklivas, Why Econometricians Don't Replicate (Although They Do Reproduce), *Review of Political Economy* **3**(2) (1991), 146–163.
- [12] C.S. Peirce, Doctrine of Chances, *Popular Science Monthly* **12** (March 1878), 604–615.