WHY JOURNAL EDITORS SHOULD ENCOURAGE THE REPLICATION OF APPLIED ECONOMETRIC RESEARCH

Edward J. Kane*
Everett D. Reese Professor
Banking and Monetary Economics
Ohio State University

I. INTRODUCTION

Few economists would find fault with the sentiment embodied in the age-old prayer: "Lord, protect us from what we only think we know." But on the equally ancient hypothesis that God only helps those who help themselves, it is useful to inquire about the safeguards that, as a profession, we have erected against the accumulation and dissemination of false knowledge.

II. ASYMMETRIES IN EFFECTIVE REPLICAIVE DISCIPLINE

With purely theoretical forms of information, incentives to search for errors are substantial. Because this search focuses on the logical consistency of premises and conclusions, it may be undertaken with a minimal investment of application-specific resources. Successful efforts are likely to be rewarded in proportion to the novelty of the result, the perceived importance of the underlying research problem, and the standing of the economist or economists who previously espoused the error (cf. Stigler [10]). Searching for errors in empirical research is more costly and, when successful, notably less rewarding than searching for them in theoretical research. In what we could call the contemporary sociology of economics, uninvitably verifying someone else's empirical research is not a completely respectable use of one's time. Choosing such a task is widely regarded as prima facie evidence of intellectual mediocrity, revealing a lack of creativity and perhaps even a bullying spirit.

Resulting differences in expected reader interest in duplicating the steps of theoretical and empirical analyses lead the representative applied

*The author wishes to thank William Diewald for helpful comments on an earlier draft.
econometrician and journal editor to place far more weight on reporting statistical results in forms that are fully interpretable than on describing the research procedures followed in a fully reproducible manner. Neither individual researchers nor the vast majority of their readers want to spend much time on tedious procedural details.

Tedious though its requirements may be, reproducibility remains the touchstone of the scientific method. If an empirical finding is a fact, other researchers should be able to observe it, too. Successful and independent repetition of an econometric experiment increases professional confidence in the experiment's alleged results.

Incentives to cheat and to rest content with results that are less than thoroughly checked cannot be wished away. Whether fraudulent or merely providential, undiscovered error can confer distinct career benefits on its perpetrators. Recognizing this, the disciplined scientific mind must hold itself skeptical of every unverified assertion. Adopting a Bayesian framework, the scientific value of replication may be stated with particular clarity. Whether it is successful or unsuccessful, replication changes the evidence for (and therefore the probability of) some scientific hypothesis.

In contemporary economic journals, referees are expected to verify theoretical findings, but not empirical ones. This paradigmatic asymmetry leaves the foundations of empirically-based economic knowledge less adequately validated than the foundations of purely theoretical economic analysis. Sloppiness (and even dishonesty) may occur in the collection and processing of economic data that an editor would never tolerate in a strictly theoretical exercise.

To change this outcome, journals must subject applied econometricians to additional discipline, presumably by lowering replication costs and/or increasing professional rewards to econometric replication. As detailed in Dewald and Thursby [5], several prestigious economics journals have recently moved in this direction: The Journal of Political Economy, the Journal of the American Statistical Association, the Journal of Econometrics, and The Journal of Money, Credit, and Banking. These journals encourage authors to provide detailed descriptions of their data sources and computer programs and to submit their data sets for the journal's files (in the case of the JMCB) or at least to make them easily available to other researchers. Of course, mechanisms for enforcing this behavior are still at a rudimentary stage, especially when proprietary elements attach to an author's data set.
III. TYPES OF REPLICATION: ECONOMETRIC AUDITING VERSUS IMPROVISATIONAL REPLICATION

Replication includes but is not limited to slavish duplication of a predecessor study's specific procedures [8]. In an experimental science or in contexts where researchers make sampling and measurement decisions, these procedures include such controls as experimental conditions, the sampling framework, and specific techniques of measurement that are seldom exercised in econometric applications. For applied econometricians, sampling and measurement issues often reduce to procedures for making proxy-variable decisions and for assembling a well-edited data set from appropriate data tapes or published sources.

Replication which uses the same data sources, models, proxy variables, and statistical methods as its predecessor piece of research may be viewed as a procedure for conducting a careful econometric audit. The exercise's focal point is to gauge the accuracy of the results reported for a given econometric experiment. When a predecessor's results are completely affirmed, a well-written audit report is likely to deserve no more than a paragraph or two of journal space. Relative to the effort required to conduct an adequate audit, this is a poor payoff indeed. Even in instances where important errors (or fraud) are uncovered, the journal space an auditor may expect to command remains severely limited.

For this reason, econometric auditing might best be performed (as Dowdall and Thursby report that it is at Ohio State) by doctoral students as part of their econometric course work. Conducting econometric audits would benefit most students by routinizing important features of the often-painful transition from merely reading economics to actually doing economic research. At the same time, hands-on testing of nontextbook models should help students to develop experience in the artful side of econometrics. By signaling an interest (as this Journal has) in considering graduate student or instructor audit writeups for publication, journal editors would further increase the motivation for contemporary graduate students to cultivate expertise in data analysis.

At least through the early 1980s, econometric audits usually developed as preliminary steps in what we may call a process of improvisational replication. Anderson and Rauscher [3, pp. 811-812] provide an excellent example of such by-the-way auditing. The word improvisational is chosen to emphasize that the improvisational replicator seeks to depart in a carefully justified way from the rote reproduction of someone else's work. As in jazz
renditions of an old standard, the performer begins by restating some familiar thematic material and goes on to impart to it a series of original and illuminating personal twists. In econometric replication, the twists lie in using a different set of data, proxy variables, joint and alternative hypotheses, estimators, or postmortem sensitivity tests. By focusing on the effect of varying one or more essential features of an econometric experiment, this kind of replication seeks to gauge the generality of a reported result. Unlike econometric auditing, improvisational replication has an adaptively creative dimension that enhances the professional standing of its practitioners. The perceived importance of any particular replication, although it tends to be short-lived, varies directly with the inventiveness of the improvisation and the prior professional standing of the predecessor research.

IV. UNDISCOVERED ECONOMETRIC ERRORS ARE PROBABLY A WIDESPREAD PROBLEM

Replication may be likened to mining for error. A prospector would not want to sink a mine shaft without first convincing himself that surface evidence favored a strike. If applied econometricians never made important mistakes, would-be auditors would quickly find better uses for their time.

To our professional shame, theoretical analysis, supplemented by unsystematically accumulated empirical evidence, suggests that would-be econometric miners may expect to strike a rich lode. As we have seen, the theoretical case for rationally expecting widespread error is that the absence of external checks greatly lessens the expected value of professional penalties imposed on an applied econometrician who knowingly or unknowingly reports false results. Empirical evidence on the frequency of error is of two types:

1. Individual errors important enough to find their way into the economics literature or public press, and
2. Preliminary results from the econometric auditing experiment being carried out under National Science Foundation sponsorship on the data sets filed with the Journal of Money, Credit and Banking.

Two infamous examples of discovered errors are cited by Dewald and Thursby: a pension study by Martin Feldstein [6], whose error is reported in Arenson [3]; and a paper on the minimum wage by Finis Welch [11], whose error is reported in Siskind [9]. Other examples include: errors in M. A. Adelman's [1] analysis of trends in manufacturing concentration, discovered
by Gardiner Means [7, pp. 124-125] and verified by Ben Brough [4]; and
inadvertently misleading descriptions of statistical procedures used in Fed-
eral Reserve staff estimates of money-market models, discovered by Ande-
son and Rasche [2].

It remains to be seen whether such mistakes are isolated problems or
nuggets broken loose from a veritable motherlode of error. However, the
37.5 percent nonresponse rate and 76 percent error rate reported by Dowald
and Thursby for the JMCB's initial efforts to collect usable data sets from
its published authors suggest that journal editors everywhere would be wise
to encourage systematic replication of the empirical research they publish.
Until econometric auditing becomes a creditable professional activity,
applied econometrics will remain something less than a fully scientific
activity.

V. SUMMARY

Philosophers are disturbed by economists' insistence on using the word
methodology as if it were no more than a high-sounding synonym for meth-
ods. This note treats methodology properly construed—the philosophical
study of how we know.1 The argument draws and emphasizes a distinction
between improvisational replication and econometric auditing. Although our
profession values improvisational replication, it imposes great costs on
would-be replicators of either kind and tender little benefit to econometric
auditors. Given this benefit-cost ratio, the value of econometric auditing
must lie (as it does for graduate students and original investigators) mainly
in the doing. In that the QJBE's decision to encourage the submission of
replicate research improves net returns to would-be replicators, econom-
ists everywhere stand to benefit. By helping theoretical and empirical
research in economics to be evaluated more symmetrically, the implied
improvement in opportunities for publication promises to reduce the extent
to which propositions in economic theory remain truer (in a Bayesian sense)
than propositions in empirical economics.

---

1To a philosopher, knowledge consists of justifiedly true beliefs. Woody Allen once observed that he had learned two things in an introductory course in philosophy: first, that we know nothing at all, and second, how we know this.


5. William G. Dewald and Jerry G. Thursby, "The JMCB Data Storage and Evaluation Project" (Paper presented at the Midwest Economics Association Meetings, April, 1982).


