

Richard G. Anderson and William G. Dewald

Richard G. Anderson is a research officer at the Federal Reserve Bank of St. Louis. William G. Dewald is director of research at the Federal Reserve Bank of St. Louis.

Replication and Scientific Standards in Applied Economics a Decade After the Journal of Money, Credit and Banking Project

SINCE EARLY 1993, the Research Department of the Federal Reserve Bank of St. Louis has made the data and programs for articles published in the Bank's *Review* available to the public on its electronic bulletin board.¹ During the first year, files from articles in the *Review* were downloaded from the bulletin board more than 200 times. More recently, about 30 files have been downloaded each month.

The Research Department of the Bank develops the program and data files on our bulletin board during a replication of each article prior to publication. A research analyst first checks the author's data against original sources. Because databases may have been updated or revised after the research began, this can require searching for the original published data. In a few cases, data errors have been corrected, fortunately with only minor impact on the author's results. Next,

an annotated version of the computer program is prepared and all statistical results recalculated. Finally, bibliographic and other references are checked by the analyst against original source documents. We believe this practice both assures the accuracy of the empirical results and allows the interested reader to delve into the details of the author's research.

THE ROLE OF DATA IN ECONOMIC EXPERIMENTS

Although empirical knowledge in both the physical and social sciences arises from repeated experiments, the role of data differs. In the physical sciences, scientists control a relatively small number of variables such as temperature, atmospheric pressure, diet or family characteristics. Since some variables are neither observed nor controlled, no two repetitions of an experiment will be

¹ The bulletin board is advertised as the Federal Reserve Economic Database, or FRED. FRED's phone number is (314) 621-1824. (The Federal Reserve System does not have a server on the Internet.) Dewald, Thursby and

Anderson (1986) summarized the *Journal of Money, Credit and Banking* project mentioned in the title.

identical. Response surface analysis and the newer field of research synthesis provide tools for analyzing the dependence of experimental results on the settings of the conditioning variables.²

In economics, however, unlike the physical sciences, researchers can only condition on the observed values of the environmental variables, not control them. Consider a simple model of an economic experiment:

1. Form hypotheses.
2. Collect data.
3. Develop theoretical and econometric framework.
4. Estimate.
5. Test hypotheses, draw conclusions.

The values of the conditioning variables are collected in step 2. Published articles typically describe steps 1, 3 and 5, but are most often silent on step 2. In principle, a researcher armed with the values of the conditioning variables and the computer code for step 4 should be able to *exactly* reproduce an economic experiment.³ Unlike the physical sciences, the experiment is deterministic, given the data.

Appraising the robustness of the results of an economic experiment requires knowing the values of the conditioning variables used by the researcher. Obtaining the data may sometimes be difficult. Datasets and programs may be mislaid or lost during the interval between completion of the research and publication of an article. Further, requests to authors for data may raise suspicions that the reader hopes (or expects) to find errors in the authors' research. An individual researcher has strong incentives not to share data and programs. If the materials are shared and results confirmed, the confirmation provides little (if any) reward to the researcher beyond the original publication of his findings. If results are found faulty, however, the researcher faces the likelihood of some professional embarrassment.

The trepidation of authors aside, scientific progress depends on challenging received wisdom. In applied economics, these challenges fall into three categories: replication of published results using the previous authors' data and programs; applying new statistical methods or techniques to authors' datasets; and application of existing statistical methods (including those used by previous authors) to new datasets.⁴ That most applied economic research falls within the third category is not surprising, since the first two depend on access to previous authors' datasets. Only with the authors' data may the reader repeat, or replicate, all five steps of the scientific experiment. Selecting a new set of values for the conditioning variables from published sources may yield results close to those obtained by the author, or results that are quite different. Unfortunately, it is difficult to predict the sensitivity of authors' results to variations in the values of the conditioning variables. For an example of the wide range of results that may arise when mixing different sources and vintages of data, see the computer simulation experiment reported in Dewald, Thursby and Anderson (1986).

THE *JMCB* PROJECT AFTER 10 YEARS

The *Journal of Money, Credit and Banking* project, conducted from 1982-84 at the editorial offices of the *JMCB* at The Ohio State University in Columbus, was the first attempt by a professional journal to make authors' programs and data available to its readers.⁵ During the project, the *JMCB* asked authors to submit data and programs to the journal's office. Conceptually, we regarded the research reported in each article as the outcome of an experiment. A complete understanding of the experiment required the researchers' data and computer programs, as well as the published summary descriptions and conclusions. For a subset of these submissions, we attempted to repeat steps 2 and 4 by collecting data from the sources cited by the authors and re-running the authors' computer programs.

² See Cooper and Hedges (1994).

³ A complication not dealt with here are errors and inconsistencies in econometric computer programs. In the *Journal of Money, Credit and Banking* project (described further below), we requested that authors provide the version, release date and serial number of the computer program used for their estimation. See Lovell and Selover (1994) for examples of the variation in econometric packages.

⁴ Various classification schemes and nomenclatures have been discussed by Kane (1984), Mittelstaedt and Zorn (1984), Hubbard and Vetter (1991), Lindsay and Ehrenberg (1993), and Fuess (1994).

⁵ The primary research team was William Dewald, Jerry Thursby, Richard Anderson and Hashem Dezbaksh. The project's findings are summarized in Dewald, Thursby and Anderson (1986). The project was supported in part by the National Science Foundation.

Table 1
Datasets for the *JMCB* Project

Datasets Requested by Month and Year through May 1989							
	1983	84	85	86	87	88	89
Jan.	0	0	12	3	9	5	8
Feb.	0	2	1	1	8	6	4
March	0	1	9	5	4	1	1
April	14	2	2	6	0	1	1
May	1	0	14	1	1	5	2
June	1	0	2	6	5	2	
July	0	4	5	3	2	2	
Aug.	3	2	4	3	0	2	
Sept.	9	4	4	2	1	1	
Oct.	11	6	7	38	2	4	
Nov.	1	2	6	4	3	4	
Dec.	0	1		6	0	3	

Datasets Available by Issue as of May 1989									
	1980	81	82	83	84	85	86	87	88
Feb.	2	3	4	6	4	5	3	2	2
May	1	3	2	5	5	5	3	5	6
Aug.	1	4	3	6	8	2	6	2	6
Nov.	4	0	3	7	7	5	7	1	3
Total	8	10	12	24	24	17	19	10	17

SOURCE: *JMCB* editorial office staff. No later data are available from the *JMCB* staff.

We found during the *JMCB* project that many authors could not furnish data and programs following publication of their articles. We initially requested data from the authors of 62 articles published during 1980-82, prior to the beginning of the project on July 1, 1982. About one-third of the authors did not respond to either a first or second request for data. Among the responding authors, one-half either could not locate their data or chose not to submit them. Most of these authors said that they could have done so if the materials had been requested when the manuscript was first submitted to the *JMCB*. Data and programs were often apparently mislaid during the relatively long delay between completion of the research and publication of its findings.

We next requested data from the authors of papers that either had only recently been accepted for publication or were under editorial review. More than three-fourths furnished their data. We

concluded that it is important for journals to request data from authors immediately following completion of the research, and for the journal to retain the data to avoid its loss during the interval between completion of the research and publication of the paper.

In the second part of the *JMCB* project, we studied whether the materials submitted by authors were in fact sufficient to allow another researcher to repeat their experiment.

For many articles, repeating step 2—searching for the authors' data in their stated sources—was impossible. Descriptions of sources were either too vague to allow us to locate the data and/or the data were not included in the cited sources. Although 54 datasets were submitted to the *JMCB* during the project, we judged only eight as satisfactory and 14 as valueless in helping us understand the authors' published work. Others were deficient in at least one important respect. For a few articles, we discovered data errors during comparison to published sources. In the most severe case, we found that an author's conclusions were reversed (prior to publication of the article) when an error was corrected. Where the data were adequate, we usually obtained numerical results from authors' programs very close to those reported by the authors.

Beyond encouraging readers to explore authors' methodology and the robustness of published results, we believed that requesting data and programs from authors would encourage them to exercise added care during their research. We also expected that other journals would adopt similar requirements to increase the value of their articles to readers. Although the *JMCB* project stimulated discussion of the role of replication in economics, no other journal adopted a policy of requesting data from authors during the 1980s. Some journals adopted editorial statements that authors should stand ready to provide data and programs to other researchers. Such statements, in and of themselves, may not solve the two major problems identified during the project: Data often are mislaid prior to publication of the article, and the author may be suspicious of the motives of a researcher requesting his data and programs.

A decade after the *JMCB* project, the replication of previous studies as a part of new research seems an infrequent occurrence. During the last decade, no papers or notes in the *JMCB* have focused primarily on replication, and only about two papers per year have included a direct compari-

son of the authors' results to those in previous studies, whether published in the *JMCB* or elsewhere.⁶ The collection of data by professional journals also remains rare. The *JMCB* discontinued requesting data from authors in 1993. To our knowledge, today only two journals—the *Journal of Applied Econometrics* and the *Journal of Business and Economic Statistics*—routinely request data from authors, and neither requests their programming.

From January 1983 through mid-1989, the *JMCB* received nearly 150 submitted datasets and about 300 requests, as shown in Table 1.⁷ Except for a surge in requests following the publication of Dewald, Thursby and Anderson (1986), on balance only a few datasets were requested each month even though the number of available datasets increased significantly during the decade.

The higher request rate during the past two years for data from the St. Louis bulletin board may suggest that the modest costs of requesting data from the *JMCB* still exceeded the marginal value to an individual researcher of replicating a previous study. To obtain data for a *JMCB* article, a reader had to call the editorial office to ask the price of the data, submit payment by mail, wait for the data to be reproduced and mailed, and perhaps re-enter the data into a computer. By contrast, the St. Louis bulletin board is free, delivery is immediate, and data are machine-readable.

THE ROLE OF THE NATIONAL SCIENCE FOUNDATION

The economics program of the National Science Foundation (NSF) has sought to build a heightened awareness of the value of data collection, archival and distribution among economists during the last decade. Following publication of Dewald, Thursby and Anderson (1986), the NSF established an archive for the storage and distribution of authors' data at the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan. Initially, some anticipated that the NSF's effort would extend the *JMCB*'s practice of requesting and distributing authors' data to

a much larger number of journals. The editors of 22 journals, however, declined invitations from NSF's economics program to request that their authors place data in the ICPSR archive.⁸

The National Science Foundation has also adopted guidelines to reduce the cost of replications. The guidelines require that authors place any data used and/or developed in conjunction with an NSF-funded project in a public archive not later than six months following the end of the grant period. Applications for additional NSF funds must contain a statement of how the author has complied with this requirement. The ICPSR accepts data from any author who has received NSF funds.⁹

National Science Foundation initiatives have also assisted users of copyrighted and confidential data. Some data obtained by researchers from commercial vendors are copyrighted and may not be further distributed by the researcher without the vendor's permission. One such vendor, the Center for Research in Security Prices (CRISP), has agreed to maintain researchers' datasets as part of its own database and make them available to all licensed users of its data. For confidential data, the Bureau of the Census and the NSF are exploring opening regional offices that would allow researchers access to confidential data, including datasets used in previously published studies. A pilot office is operating in Boston.

REPLICATION OF THE *JMCB* PROJECT IN ST. LOUIS

Our experience at the *JMCB* during 1982-84 was itself only one trial of an experiment. Would another sample of authors also have difficulty providing data following publication of their articles? Or were our original findings anomalous, leading us to greatly exaggerate the problem, as some critics have suggested?

During 1992-93, we repeated the *JMCB* experiment at St. Louis, in part, by requesting data and programs from the authors of papers presented at the Bank's annual economic policy conference in October 1992. We did not tell authors prior

⁶ Replication also has been relatively rare even in journals that encourage submission of such papers. See Fuess (1994).

⁷ Recall that in 1982 we began requesting data for articles that had been published as early as 1980.

⁸ Some authors have since proposed models of how such collective disinterest among professional journals might arise and be sustained. See Feigenbaum and Levy (1994, 1993).

⁹ Materials should be submitted to User Support, ICPSR, P.O. Box 1248, Ann Arbor, MI 48106. Data may be retrieved from ICPSR via an Internet server; see Goffe (summer 1994; March 1994).

to the conference that we would be requesting their programs and data. Their responses were very similar to those we had observed at the *JMCB* a decade earlier. Authors who had completed their studies just prior to the conference promptly submitted their data and programs, while those who had largely completed their research as much as several years earlier found it costly to organize and submit data. The following year, we informed participants for our 1993 conference at the time their paper was invited that we expected data to be submitted with the manuscript. Authors generally found it imposed little burden to submit data and programs with their manuscripts *so long as they were aware of the requirement in advance*, although the Bank's staff had to make some follow-up calls to clarify documentation.

THE FUTURE ROLE OF REPLICATION

How should economists, as scientists, interpret their profession's apparent, collective lack of interest in replicating prior studies? Does the failure of journals to request and distribute data and programs reflect a lack of scientific discipline in economics, as some have suggested? Do the costs to a reader of obtaining data from authors exceed the benefit?

The *JMCB* project demonstrated that professional journals could reduce the costs of replication and improve the quality of applied economic research by collecting data and programs from authors. For the first time, the reader of an empirical article could obtain data and programs anonymously (with respect to the author) from the journal that published the article. Further, collection of the authors' data and programs seemed to encourage additional care by authors to avoid inadvertent errors during the conduct of the research.

Since January 1994, the *Journal of Applied Econometrics* has accepted papers for publication conditional on authors furnishing an acceptable dataset. The publication reviews datasets for completeness and then places them on an Internet server at Queens University.¹⁰ The *Journal of Business and Economic Statistics* requests, but does not require, that authors submit data for

distribution via an Internet server at Duke University. Tauchen (1993) argues, as we did in 1986, that readers of journals should be interested in authors' data and programs, and a journal's prestige (and circulation) should increase when such data and programs are made readily available to readers. We believe that published empirical research generally would benefit if the practices of the *Journal of Applied Econometrics*, *Journal of Business and Economic Statistics*, and the *St. Louis Review* were more widely adopted.

REFERENCES

- Cooper, Harris, and Larry V. Hedges, eds. *The Handbook of Research Synthesis*. New York, 1994.
- Dewald, William G., Jerry G. Thursby, and Richard G. Anderson. "Replication in Empirical Economics: The *Journal of Money, Credit and Banking* Project," *American Economic Review* (September 1986).
- Feigenbaum, Susan, and David M. Levy. "The Self-Enforcement Mechanism in Science," mimeo, paper presented at the American Economic Association Meetings, Boston, Massachusetts, January 1994.
- _____, and _____. "The Market for (Ir)Reproducible Econometrics," *Social Epistemology* (no. 3, 1993), pp. 243-44.
- Fuess, Scott M. "On Replication in Business and Economic Research: The *QJBE* Case," paper presented at the Eastern Economic Association meetings, Boston Massachusetts, March 1994 (mimeo, Department of Economics, University of Nebraska, October 1993).
- Goffe, William L. "Computer Network Resources for Economists," *Journal of Economic Perspectives* (summer 1994).
- _____. "Resources for Economists on the Internet," electronic document maintained on the Internet (March 1994 version). Contact bgoffe@whale.st.usm.edu.
- Hubbard, Raymond, and Daniel E. Vetter. "Replications in the Finance Literature: An Empirical Study," *Quarterly Journal of Business and Economics* (autumn 1991).
- Kane, Edward J. "Why Journal Editors Should Encourage the Replication of Applied Econometric Research," *Quarterly Journal of Business and Economics* (winter 1984).
- Lindsay, Murray R., and A.S.C. Ehrenberg. "The Design of Replicated Studies," *The American Statistician* (August 1993), pp. 217-28.
- Lovell, Michael, and David Selover. "Econometric Software Accidents," *The Economic Journal* (May, 1994), pp. 713-25.
- MacKinnon, James. "Guidelines for Users: Journal of Applied Econometrics FTP Archive Site," *Journal of Applied Econometrics* (April-June 1994), pp. 229-30.
- Mittelstaedt, Robert A., and Thomas S. Zorn. "Econometric Replication: Lessons from the Experimental Sciences," *Quarterly Journal of Business and Economics* (winter 1984).
- Tauchen, George. "Remarks on My Term at JBES," *Journal of Business & Economic Statistics* (October 1993).

¹⁰ For discussion of the *Journal of Applied Econometrics* server, see MacKinnon (1994).