Econ Journal Watch,
Volume 4, Number 3,
September 2007, pp 326-337.

ECONOMICS IN PRACTICE

Got Replicability?
The *Journal of Money, Credit and Banking* Archive

B.D. McCULLOUGH.[1]

ABSTRACT

During the period 1982-1984, all submissions to the *Journal of Money, Credit and Banking* (*JMCB*) were required to submit data and code to The journal's office. The results of this project, published in Dewald, Thursby and Anderson (1986), shocked the economics profession. The vast majority of the research could not be replicated. As a result, the *American Economic Review* (*AER*) adopted its "replication policy" whereby authors were required to provide their data and code to other researchers upon request (Ashenfelter et al 1986). The *AER* provided no formal enforcement procedure or penalty for failure to comply. Dewald, Thursby and Anderson (1986) recommended a mandatory archive instead of a mere "policy" lacking formal accountability. Many journals followed the *AER*'s lead and adopted "policies". When McCullough and Vinod (2003) showed that the *AER* policy was ineffective, then-editor Bernanke (2004) instituted a mandatory data+code archive. At that point, no one had ever analyzed any of the existing archives.

Kerry Anne McGeary, Teresa Harrison, and I analyzed several years of the *JMCB* data+code archive, 1996-2003. Of 266 articles published during that time, 193 were empirical and should have had data and code deposited in the archive. Of these, only 69 actually had anything in the archive; 11 of them had data only, and seven articles required software or other resources we did not have. Excluding these seven, the *JMCB* archive only enabled replication of 14 of 186 empirical articles. Two primary reasons that the archived data and code did not reproduce the published results (assuming that the results actually are replicable) is carelessness on the part of the authors and a failure of the editors to provide proper instruction to

1 Department of Decision Sciences, Drexel University, Philadelphia, PA 19104.
For comments and useful suggestions, thanks are due to Teresa Harrison, Kerry Anne McGeary, H. D. Vinod, and a referee.

authors on how to prepare replication files. The interested reader is invited to consult our paper (McCullough, McGeary, and Harrison, hereafter abbreviated MMH 2006), which gives several examples. The obvious conclusion is that the mandatory data+code archive of the *JMCB* did not work to ensure replicability. Based on our experience attempting to use the archives to replicate all those articles, we recommended several procedures to enhance the replicability of *JMCB* articles.

RECOMMENDATIONS WE MADE FOR AN EFFECTIVE ARCHIVE

Here I reiterate the recommendations we made for an effective archive, and make comments to explain or elaborate.

[A] The README file should list all the replication files with a brief description of each. It should clearly indicate which programs correspond to what results in the paper.

- H.D. Vinod and I attempted to replicate all the empirical articles in the June 1999 issue of the *American Economic Review* (McCullough and Vinod 2003). We noted that one of the authors (who was on the AER editorial board at the time), "after several months and numerous requests, finally supplied us with six diskettes containing over 400 files – and no README file. Reminiscent of the attorney who responds to a subpoena with truckloads of documents, we count this author as completely noncompliant" (McCullough and Vinod 2003, 887-888). Trying to reproduce the results without a good README file was hopeless. This problem also occurred when MMH attempted to replicate results in the *JMCB*.

[B] The README file should also contain a data dictionary that defines each variable and gives the provenance of all the data.

- Levitt (1997) published an article in which his primary contribution was the creation of an instrumental variable that would break the simultaneity between police and crime. This variable was constructed by noting the timing of mayoral and gubernatorial elections. When McCrary (2002) attempted to reconstruct this variable he was unable to do so. In his reply to McCrary's comment, Levitt (2002) admitted that he was unable to reconstruct this variable either. If Levitt had a data dictionary for his original article, the error would not have occurred. MMH attempted only to use the supplied data and code, not to verify that the data were correct. If they had attempted to verify the data, in most cases they would have been frustrated because most authors did not specifically describe where their data came from. It is one thing to say that data came from the *Survey of Current Business*. It is another thing to say which issue, table and page the data came from, since economic data may be revised several times.

[C] The README file should identify the version of the software used (by version number and/or release date) and similarly for the operating system on which the software runs.

- Different versions of the same software package sometimes produce different answers to the same problem, for example, when a bug is fixed. Run on different versions of an operating system, the same software can produce different answers. McCullough and Vinod (2004, 394) document a case in which the more recent version of the operating system had better math libraries. A very small number had incorrectly been evaluated as "zero" by the old operating system, producing a log-of-zero problem that did not appear with the more recent version of the operating system. With the old version the program came crashing to a halt because the log of zero does not exist.

[D] All data should be provided in ASCII format, and the version of the code submitted to the archive should call these same ASCII files.

- Data that are provided in machine-readable form often cannot be read by another package. Thus, someone trying to use a software package different than the one used by the original author might not be able to read the data. Further, the requirement that the code call the ASCII data ensures that the code will work with the data that are provided, and that the data provided really do reproduce the published results (below I tell of this problem as encountered in trying to replicate the *JMCB* article by Lastrapes in the December 2006 issue of the *JMCB*), as discussed below.

[E] Authors should provide the original data from which the final dataset is derived and all instructions/code necessary to turn the original data into the dataset analyzed.

- Often data from different sources are combined to produce a usable dataset, or data are transformed before use. Without the exact record of how that was done, it may be impossible to replicate the dataset. A good example of this is the Caroline Hoxby/Jesse Rothstein debate over schooling. Rothstein (2004, 8) wrote, "A major difficulty in replicating Hoxby's sample is the matching of NELS schools to MSAs, as the NELS offers several indirect indications of schools' locations but no direct MSA code. Hoxby reports an 8th grade sample size of 10,790 students from 211 MSAs, but does not report her geocoding algorithm. I am unable to replicate her exact sample size." Naturally, if Rothstein couldn't get the correct sample, he couldn't reproduce Hoxby's published results.

[F] The author should provide code such that the data and code, when placed in the same subdirectory, will execute. Also, the output from doing so should be provided. The author should check to make sure that this runs correctly and produces the results in his paper.

- Xiangrong Liu and I encountered that problem when trying to replicate the Meiers and Mueller article in the December 2006 issue of the *JMCB* (McCullough and Liu 2007), as discussed below.

[G] The archive should list each paper regardless of whether the paper has been exempted from the rules. Sometimes a paper has been exempted from the data requirement. In such cases, the archive should say that the paper has been exempted from the data requirement, and the code should still be required.

- That problem marks the Brevoort and Hannan (2006) article in the December 2006 issue of the *JMCB* .

[H] The journal should issue conditional acceptance letters, with a formal acceptance letter being sent only after the data+code have been archived.

- The *Journal of Applied Econometrics* has 99% compliance, because the editors do not send an acceptance letter until the archive manager, James MacKinnon, has informed them that he has received the required material. MMH (2007) compared the proportion of empirical articles that have archive entries for several journals. Some of their results are summarized in Table 1.

Table 1: Lifetime Compliance

Journal	Empirical articles	Entries	Compliance %
J. App. Econometrics	292	290	99
Fed. St. Louis Review	219	162	74
JMCB	193	66	34
J. Bus. Econ. Statistics	342	121	35

Number of empirical articles (that should have archive entries), actual number of archive entries, percentage of empirical articles that have archive entries.

Most journals with archives do not bother to make sure that anything is archived. The missing 1% of *JAE* articles were due to special issues that were exempted from the usual process. Note, however, that the *JAE* requires only data, not code, so there is still room for significant improvement.

[I] Managing the archive should be an editorial function.

- Let us learn from the success of the *Journal of Applied Econometrics*, where managing the archive is an editorial function. I think it fairly easy to find a fastidious junior professor who would pay necessary attention to the archive. A major problem with the old *JMCB* archive was that managing the archive was a secretarial function.

[J] The pages of the journal should give space to replication attempts, at a minimum by publishing a one-page summary with supporting materials placed in the archive.

- In the history of the *JMCB* archive surveyed by MMH, most articles could not be replicated. During that period, the editors published not a single note on the replicability of any article. Authors knew that the lack of replicability of their research would not be exposed by the journal. The editors should welcome replication attempts, and publish responsible ones. (In the meantime, send your critiques to *Econ Journal Watch*!)

Again, the *Journal of Applied Econometrics* sets a good example. The *JAE* has a Replication Section, edited by Baldev Raj, and will correct published errors—see, for example, Kleiber and Zeileis's (2005) correction of Bai and Perron's (2003) influential article and software on structural breaks. The *JAE* is not afraid to admit that researchers might make mistakes or (in the case of Bai and Perron) that software has bugs.

JMCB Response to McCullough, McGeary, & Harrison

The MMH article was formally accepted in 2003, but for whatever reason it was not published until June 2006, which is when the editors followed with a brief editorial about new procedures (*JMCB* Editors 2006). They announced in the June 2006 issue:

> Since January 2005, the *JMCB* research assistants have been replicating some of the results from the *JMCB* Archive for the purpose of helping the editors to evaluate various methods to improve the policy. So far, we have adopted the following new procedure:
>
> 1. When a paper is submitted, the *JMCB* office writes to the corresponding author the following: Publication of any paper regarded as empirical is conditional on compliance with the data archiving policy below. If your paper is empirical, please start to prepare data, programs, and a README file now. If providing data is infeasible, please write to the *JMCB* office to ask for an exemption in writing as soon as possible.
>
> 2. If the exemption is requested, then the Editor will notify the *JMCB* office whether or not the data requirement is waved for the particular paper, and the *JMCB* office will let the author know by the time of the first turn-around time. If an exemption is requested and refused, the Editor will let the author know as soon as possible.
>
> 3. When a paper is accepted, the *JMCB* office will remind the corre-

> sponding author that the publication is conditional on our receiving data, programs, and a README file unless that the paper is given an exemption by the accepting Editor. (The *JMCB* Editors 2006, 1108)

The three new procedures are in the spirit of the MMH recommendations. The editors refrained from the other recommended procedures, stating that they "do not agree with all the arguments made by [MMH]." The editors did not say which arguments they disagreed with, or the grounds for disagreement.

Since the editorial appeared in June 2006 and states, "we have adopted the following new procedures," it seems reasonable to suppose that articles appearing four issues later, in December 2006, would be subject to the new regime. Thus, enough time has passed to see whether the changes implemented by the editors have been effective at ensuring that *JMCB* publishes replicable research. The present paper examines the December 2006 issue of the journal and concludes that new policy is not yet working.

Table 2: Articles in the December 2006 issue of *JMCB*

		Empirical?	Archive?
1.	"A New Analysis of the Determinants of the Real Dollar-Sterling Exchange Rate: 1871-1994" by Ivan Paya and David A. Peel	✓	0
2.	"Commercial Lending and Distance: Evidence from Community Reinvestment Act Data" by Kenneth P Brevoort and Timothy H. Hannan	✓	0
3.	"Is There a Cost Channel of Monetary Policy Transmission? An Investigation into the Pricing Benaviour of 2,000 Firms" by Eugenio Gaiotti and Alessandro Secchi	✓	0
4.	"When Did the FOMC Begin Targeting the Federal Funds Rate? What the Verbatim Transcripts Tell Us" by Daniel L. Thornton	✓	0
5.	"Dollarization Traps" by John Duffy, Maxim Nikitin, and R. Todd Smith	0	0
6.	"Fleshing Out the Monetary Transmission Mechanism: Output Composition and the Role of Financial Frictions" by Andre Meier and Gernot J. Mueller	✓	✓
7.	"Technical Trading-Rule Profitability, Data Snooping, and Reality Check: Evidence from the Foreign Exchange Market" by Min Qi and Yangru Wu	✓	0
8.	"Inflation and the Distribution of Relative Prices: The Role of Productivity and Money Supply Shocks" by William Lastrapes	✓	✓
9.	"Linking Individual and Aggregate Price Changes" by Attila Ratfa	✓	0
10.	"Cities and Countries" by Andrew K. Rose	✓	0
11.	"Inflation Inertia and the Optimal Hybrid Inflation/Price-Level Target" by Oisten Roisland	0	0

The December 2006 issue has 11 papers, which are listed in Table 2. As can be seen, 9 of the 11 papers were decidedly empirical and yet, the archive had entries only for two, the paper by Meier and Mueller and the paper by Lastrapes.

Empirical Articles with No Data/Code in the Archive

With respect to the seven December 2006 *JMCB* empirical articles for which there was nothing in the archive,[2] it is of interest to ascertain why there are no archive entries and whether data and code actually exist. To do this, an email was sent to all authors of each article, inquiring whether the data and code were submitted to the *JMCB* and asking for the data and code. Our interest is only in the extent to which the journal ensures replicability of published results, not whether the articles themselves are replicable. So no attempt is made to replicate the articles in this section.

Ivan Paya and David A. Peel

No response was received from either author.

Kenneth P. Brevoort and Timothy H. Hannan

Brevoort responded that their article used confidential data—something not indicated in the archive. Further, the journal only exempted the data from the archive requirement, not the code. If some researcher wishes to apply the same method to a different data set, should he have to re-invent the wheel? Even if the data are not available, ought not the code be available for inspection? MMH made two recommendations that are relevant here. First, if confidential data are used, this should be noted in the archive. Second, if confidential data are used, the code nonetheless should be made available.

Eugenio Gaiotti and Alessandro Secchi

No response was received from either author (aside from an "out of the office" auto-reply from Giaotti's email account).

Daniel L. Thornton

Thornton's article made minimal use of the data, computing only summary statistics. Thornton replied that he no longer had the original data, but did quickly put together some similar data and supplied them. The original data on which the results are based are now lost and unrecoverable. If the journal had requested his data at the time of publication, this article would be replicable. Ideally, he would have supplied code and information about the software used (including version)

2 In fact, for those seven articles, there continued to be nothing in the archive right up to August 2007 when the present article was finalized.

and the operating system, since, as noted, both can affect computed results.

Min Qi and Yangru Wu

Qi responded that he supplied the data and code to the journal in August 2005, and that he had no idea why it has not been posted to the archive. In that same email he sent the data and code. Perhaps his email never made it to the journal, but it is clear that somebody at the journal should be, but is not, checking to see whether data and code have been submitted before sending articles to the publisher.

Antilla Ratfai

No response was received.

Andrew K. Rose

Rose responded that the data and code are posted at his personal website, and he did not recall whether he submitted the data and code to the journal. I checked his website (on 22 August 2007) and did not find the code available. I did find the output from running his code. The output contained the original commands, but to produce usable code a replicator would have to do much cutting and pasting.

EMPIRICAL ARTICLES WITH DATA AND CODE IN THE ARCHIVE

There were two empirical articles with data and code in the archive. A tenet of Gary King's "Replication Standard" is that the replication should be feasible without contacting the original author or anyone else (King 1995, 444). The two papers were ultimately replicable, but only with the assistance of the original authors.

The authors of the first of the two empirical articles with data and code, Andre Meier and Gernot J. Mueller, submitted machine-readable MATLAB code that could be read only by MATLAB; similarly, their data files were readable only by MATLAB. They did have a README file in ASCII, but it was missing much of the information recommended by MMH. If the purpose of the archive is only to reproduce published results, then there is no need for the code to be readable by humans. The primary purpose of the archive is to support the extension of research, and the archived data and code should facilitate the porting of the code from one package to another. Machine-readable data/code cannot do this. MMH recommended that data and code be human-readable.

When Xiangrong Liu and I ran Meier and Mueller's data and code, the output produced results that clearly differed from the published results (McCullough and Liu 2007). The authors, when contacted, provided necessary details for modifying the code so that it would reproduce some of the published results. That

information should have been in the README file. Even then, only two of three tables could be reproduced, and it was necessary for the authors to provide yet further instruction before the third table could be reproduced. What should have been a simple matter of loading data and code into a common directory and executing the code took several hours and required multiple communications with the authors spread out over several days.

The second empirical article with data and code in the archive was authored by William Lastrapes. In the course of replicating the Lastrapes article, Bret Meyers and I discovered that Professor Lastrapes had omitted a variable from his dataset, and also included the incorrect version of a necessary subroutine (McCullough and Meyers 2007). Professor Lastrapes, when contacted, quickly provided the missing variable and the correct version of the subroutine. the journal should have provided him with instructions on how to provide replication files that satisfy King's Replication Standard. For example, Lastrapes' code loaded the data from several Excel files. the journal told him to provide ASCII data, so he just exported the data to unlabelled ASCII (no variable names, just numbers) and did not change his code. Since a variable was missing, there was no possible way that we could correctly identify the names of the variables in the ASCII data files. Furthermore, even if the missing variable had been provided, only someone with a great deal of expertise in RATS could have matched the variable names in the code to the columns in the data files.

Again, MMH recommended that the journal instruct authors to prepare their data and ASCII code so that it will all run from one subdirectory. This would have enabled Lastrapes to catch his errors before submitting his data and code.

Further, there were version-dependent errors produced by the code that could have been avoided if the version of RATS run by Lastrapes had been known (this information was obtained by contacting Lastrapes). MMH recommended that authors be require to place this information in the archive.

Many researchers are under the impression that if they just provide some version of their data and code, then any other researcher can reconstruct the results. That was shown to be untrue by, among others, Dewald et al (1986), McCullough and Vinod (2003), and MMH. Even original authors are often unable to reproduce their published results. It is clear that authors need some written instruction on how to prepare replication files, and that the journal should provide it.

Conclusions

Some say that economics needs math for rigor, clarity, and accountability. Surely that view contains some truth. Yet critics have often argued that math can also produce opacity and a lack of accountability. As noted at the start of this paper, investigations into the replicability of econometric results have dispelled

the vision of the applied econometrician as saint of replicable research. Investigations I have undertaken have convinced me that accountability will not be achieved by fuzzy hopes and gestures. In addition to the failures in documenting and providing data and code, there is the simple matter of authors' responding to inquiries about their work. As indicated here, a significant portion of authors simply "blow off" research inquiries.

MMH criticized the *JMCB* archive project, identifying two major flaws: (1) most authors of empirical articles did not contribute data and code; and, (2) when data and code were contributed, the data and code did not reproduce the published results (MMH, 1005). In response, The editors instituted a new set of procedures.

A spot check of the December 2006 issue of the journal reveals that the new procedures have not ensured that the data and code in the archive can reproduce the published results. Of eleven articles in that issue, nine were empirical and only two had data/code in the archive. Furthermore, the data/code in the archive failed to reproduce the published results.

Perhaps the December 2006 issue of the journal was a fluke. Perhaps all the other issues since the publication of The editors (2006) have data and code archived for each empirical article, and the data and code reproduces the published results of each article. The interested reader is invited to check the *JMCB* archive (link.)

Such a reader would find as of this writing (September 1, 2007) that the archive stops with December 2006, even though 2007 issues have been published for February, March/April, June and August. It is difficult to imagine that it would take more than six months to archive the data and code from the February issue, for example.

All available evidence indicates that replicable economic research is the exception and not the rule, and this appears to be the case at the *JMCB* even though it nominally has a mandatory data+code archive. The editors of the *JMCB* are to be commended for wanting to make The journal more scientific and for making gestures toward ensuring that results published in their journal are replicable. Achieving this, however, will require more than mere gestures. I believe that the procedures needed were recommended by MHH. They will require serious editorial commitment.

REFERENCES

Anderson, Richard, William H. Greene, B. D. McCullough and H. D. Vinod. 2008. The Role of Data/Code Archives in the Future of Economic Research. *Journal of Economic Methodology* (to appear).

Ashenfelter, Orley, Robert H. Haveman, John G. Riley, and John T. Tay-

lor. 1986. Editorial Statement. *American Economic Review* **76**(4): v.

Bai, Jushan and Pierre Perron. 2003. Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics* 18(1): 1-22.

Dewald, William G., Jerry Thursby and Richard G. Anderson. 1986. Replication in Empirical Economics: The *Journal of Money, Credit and Banking* Project. *American Economic Review* 76: 587-603.

JMCB Editors. 2006. *Journal of Money, Credit, and Banking* **38**(4): 1108.

King, Gary. 1995. Replication, Replication. *Political Science & Politics* (September): 444-452.

Kleiber, Christian and Achim Zeileis. 2005. Validating multiple structural change models-a case study. *Journal of Applied Econometrics* 20(5): 685-690.

Levitt, Steven D. 1997. Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime. *American Economic Review* 87(3): 1230-1250.

Levitt, Steven D. 2002. Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Reply. *American Economic Review* 92(4): 1244-1250.

McCrary, Justin. 2002 Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment. *American Economic Review* 92(4): 1236-1243.

McCullough, B. D. and Bret Meyers. 2007. Replicating Lastrapes' 2006 *JMCB* Article. *Indian Journal of Economics and Business* (to appear).

McCullough, B. D. and Xiangrong Liu. 2007. Replicating Meier and Mueller's 2006 *JMCB* Article. *Indian Journal of Economics and Business* (to appear).

McCullough B. D., Kerry Anne McGeary and Teresa Harrison. 2006. Lessons from the *JMCB* Archive. *Journal of Money, Credit and Banking* **38**(4): 1093-1107.

McCullough B. D., Kerry Anne McGeary and Teresa Harrison. 2007. Do Economics Journal Archives Promote Replicable Research? *SSRN.com.*

McCullough, B. D. and H. D. Vinod. 2003. Verifying the Solution from a Nonlinear Solver: A Case Study. *American Economic Review* **93**(3): 873-892.

McCullough, B. D. and H. D. Vinod. 2004. Verifying the Solution from a Nonlinear Solver: Reply. *American Economic Review* 94(1): 391-396.

Rothstein, Jesse. 2004, Does Competition Among Public Schools Benefit Students and Taxpayers? Comment. Manuscript. Princeton University

ABOUT THE AUTHOR:

B. D. McCullough received a B.A. in economics from Georgetown University in 1981 and a Ph.D in economics from the University of Texas at Austin in 1989. After five years as an assistant professor of economics at Fordham University, he joined the Federal Communications Commission as a senior economist where he began his work on the accuracy of statistical software and the replicability of published research. He has published replicability results and commentary in *American Economic Review*, *Journal of Economic Methodology*, *Journal of Money, Credit and Banking*, *Indian Journal of Economics and Business, Journal of Economic Perspectives,* and *Journal of Economic and Social Measurement*. Since 2001 he has been on the faculty of the Decisions Sciences Department at Drexel University, being promoted to Professor in 2006. His email is bdmccullough@drexel.edu.

Go to September 2007 Table of Contents with links to articles