American Economic Review: Papers & Proceedings 2017, 107(5): 56–59
https://doi.org/10.1257/aer.p20171033

Replication, Meta-Analysis, and Research Synthesis in Economics†

By Richard G. Anderson and Areerat Kichkha*

I. Replication

Replicability is essential to assure the scientific integrity of empirical economic research: beginning with an author's dataset and applying the mathematical operations specified by the author, a researcher should obtain the author's numerical results. These actions—labeled "econometric auditing" by Kane (1984) and "pure replication" by Hamermesh (2007)—are issues with which economics, as a putative science, continues to struggle.[1] Absent replicability, the "quality" of a published study must be judged by nonscientific criteria (the prominence of the author, the skill of the writing, a clever statistical innovation, personal connections, etc.) which, while valuable, complicate assessing the study's contribution to science—the essential act of research synthesis. Modern discussion of the issue largely dates from Dewald, Thursby, and Anderson's (1986) summary of their project at the *Journal of Money, Credit and Banking*. Other authors in this session have surveyed the mixed state of replication in economics today.

*Anderson: Lindenwood University, 209 S. Kingshighway, St. Charles, Missouri 63301 (e-mail: rganderson@alum.mit.edu); Kichkha: Lindenwood University, 209 S. Kingshighway, St. Charles, Missouri 63301 (e-mail: akichkha@lindenwood.edu).

† Go to https://doi.org/10.1257/aer.p20171033 to visit the article page for additional materials and author disclosure statement(s).

[1] Mistakes by a putative replicator that might cause assertions that a study is not replicable are not evidence that a study is not replicable. But, echoing Dewald et al. (1986), Anderson et al. (2008), and McCullough, McGeary, and Harrison (2006, 2008), we insist that the programming code is essential for replication; see Monroe (2009) for a more general argument in all fields of science.

II. Meta-Analysis

Replication is but one statistical tool that labors in service of the concept of "research synthesis" (Cooper and Hedges 2009a, b; Cooper 2017). Meta-data analysis is another. Care is necessary when applying meta-analysis in economics because, unlike studies in psychology or the health sciences, empirical studies in most of economics are nonstochastic and analyze publicly available data, although dataset selection may differ study-to-study. Such studies correspond to Kane's (1984) "improvisational replication" and Hamermesh's (2007) "scientific replication." This definition is familiar in the physical and biological sciences, where "data" arise from experiments that are inherently stochastic: "replication" is "repetition" and "replication of the experiment" creates a second (or third, or fourth) set of data. But even within these disciplines, *conditional on analyzing the same set of data and using the same tools*, subsequent researchers should obtain the same statistical results as a previous researcher.

Advocates of meta-analysis suggest that traditional literature surveys cannot contribute (very much) to research synthesis (e.g., Stanley and Jarrell [1989] 2005, p. 6):

> *Literature surveys are instrumental in summarizing the contending economic theories and in framing the remaining issues at stake. Nonetheless, there remains a great deal of subjectivity in literature surveys. The reviewer often impressionistically chooses which studies to include in the review, what weights to attach to the results of these studies, how to interpret the results, and which factors are responsible for the differences among these results.*

But, disguised somewhat by choice of language, these *same choices* are made in meta-analysis. Meta-analysis is not an "a-scientific technique"—decisions are required regarding which studies to include, what weights to attach to the studies, and how to interpret the results. Meta-analysis is but a statistical technique in service of the goal of research synthesis; it is not the synthesis itself. The rules and language of statistics apply.

A topic frequently addressed by meta-analysts has been the employment effects of minimum wages; see for example Congressional Budget Office (2014). Economists' skepticism of meta-analysis in this literature is well-expressed by Neumark (2015, p. 2):

> *Some studies use "meta-analysis," averaging across a set of studies to draw conclusions. For example, Doucouliagos and Stanley (2009) [British Journal of Industrial Relations] report an average elasticity across studies of −0.19, consistent with earlier conclusions, but argue that the true effect is closer to zero; they suggest that the biases of authors and journal editors make it more likely that studies with negative estimates will be published. However, without strong assumptions it is impossible to rule out an alternative interpretation—that peer review and publication lead to more evidence of negative estimates because the true effect is negative. In addition, meta-analyses do not assign more weight to the most compelling evidence. Indeed, they often downweight less precise estimates, even though the lower precision may be attributable to more compelling research strategies that ask more of the data. In short, meta-analysis is no substitute for critical evaluation of alternative studies.*

Publication bias is a well-honed theme in meta-analysis. Searches for publication bias (almost) always are conducted outside a meta-analysis, confined to subsets of the published literature: publication bias does not exist in a well-executed meta-analysis because such studies, by construction, include all published and unpublished primary studies. Extant statistical tests for publication bias consist (largely) of scatter plots that compare a standardized effect size to either the study's sample size or the effect's standard error. The power of such tests arises solely from the assertion that, if the true effect is nonzero, then the standard error should decrease as sample size increases. Unfortunately, if the true effect in the population does not exist (equals zero) the tests have no power, that is, the standard error may fail to decrease because the effect, in fact, does not exist (Stanley 2005; Sutton 2009).[2]

In economics, the primary focus has been on regression coefficients. Stanley and Jarrell ([1989] 2005) proposed the model[3]

$$\hat{\beta}_j - \beta = \sum_{k=1}^{K} \alpha_k Z_{jk} + \varepsilon_j \quad j = 1,2,\ldots,L,$$

where $\hat{\beta}_j$ is the estimate of β reported in study j, and the Z_{jk} are covariates coded by the meta-analyst. As illustrations they included:

(i) Dummy variables which reflect whether potentially relevant independent variables have been omitted from (or included in) the primary study.

(ii) Specification variables that account for differences in functional forms, types of regression, and data definitions or sources.

(iii) Sample size.

(iv) Selected characteristics of the authors of the primary literature.

(v) Measures of research or data quality.

Overall, it is difficult to propose a more subjective set of covariates: Except perhaps for sample size, the setting of these covariates inherently requires subjective judgments of the researcher. But, before even these covariates can be coded, the meta-analyst must specify the precise hypotheses to be tested: this decision delimits the range of studies to be considered. It is important that this range be broad: ample examples exist

[2] Sutton (2009, p. 448), cautions: "Publication bias is a difficult problem to deal with because the mechanisms causing the bias are usually unknown, and the merit of any method to address it depends on how close the assumptions the method makes are to the truth."

[3] Roberts (2005, p. 1), in a volume co-edited with T. D. Stanley, writes that "Stanley and Jarrell (1989) has stood the test of time and remains a landmark in the exposition of MRA [meta-regression analysis]." Hence, the focus herein on the statistical procedures recommended in that article.

where study validity was compromised by a narrow hypothesis and inadequate literature search (e.g., Cordray and Morphy 2009). But broadening the scope increases the diversity of included primary studies, complicating choice and coding of the meta-regression's explanatory variables. Because these choices are the meta-analyst's equivalent of a model specification search, issues arise regarding specification choice and robustness, belaying the apparent linear simplicity of the commonly used fixed-effects ANOVA model. Particularly troublesome are dummy variables to measure the inclusion/exclusion of "potentially relevant" explanatory variables and to measure research or data "quality": How is the meta-analyst to judge quality absent replicating the primary study? Eventually, implicitly or explicitly, a quality index q_i emerges, either binary $q_i = [0,1]$ such that study i is discarded when $q_i = 0$ and included without question when $q_i = 1$, or continuous $q_i = [R^1]$.

Formally, assume that a regression coefficient and its [conditional] variance, $\{\beta_i, \nu_i\}$ are observed in k independent primary studies, yielding $\{\hat{\beta}_1, \ldots, \hat{\beta}_k, \hat{\nu}_1, \ldots, \hat{\nu}_k\}$.[4] Assuming a fixed-effects model, assert that $\beta_1 = \cdots = \beta_k = \beta$, such that β is our synthesis. (It is unlikely that $\nu_1 = \cdots = \nu_k = \nu$.) Letting $\{w_1, \ldots, w_k\}$ denote a set of weights, under the null an unbiased estimator of β is $\hat{\beta}_{\bullet} = \sum_{i=1}^{k} q_i w_i \hat{\beta}_i / \sum_{i=1}^{k} q_i w_i$. The variance of $\hat{\beta}_{\bullet}$ is minimized by setting $w_i = \nu_i^{-1}$ if known or, if unknown, to $w_i = \hat{\nu}_i^{-1}$, where regression covariates may be included to improve the estimates $\{\hat{\nu}_i\}$.

III. Cautions[5]

#1.—Meta-analysis delivers correlation, not causation. The latter only can be addressed in primary studies, not in meta-analysis.

#2—Because of #1, significance of variables in meta-analysis does not signal important directions for future research.

#3.—If primary researchers begin to follow the signals in #2, meta-analytic explanatory variables will become endogenous, challenging the statistical consistency of meta-estimates.

#4.—Meta-analytic studies should be replicated. Choices with respect to the set of studies included, variable definitions, coding, etc. should be audited and replicated/reproduced, as in other empirical economic studies.

In Brief.—Replication, meta-analysis, and traditional literature surveys all serve the goal of research synthesis. "Pure replication" has a clear practical goal: reproduce the empirical findings of an analysis beginning with the same dataset and using the same methods. The others contain substantial judgment.

[4] Bandwagon effects, common in most fields, suggest that studies may not be independent.

[5] A number of meta-analysts have suggested the items mentioned here. Due to space limitations, I do not cite them individually.

REFERENCES

Anderson, Richard G., 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* 15 (1): 99–119.

Congressional Budget Office. 2014. *The Effects of a Minimum-Wage Increase on Employment and Family Income.* Washington, DC: Congressional Budget Office.

Cooper, Harris. 2009. "Hypothesis and Problems in Research Synthesis." In *The Handbook of Research and Synthesis and Meta-Analysis, Second Edition*, edited by Harris Cooper, Larry V. Hedges, and Jeffrey C. Valentine, 19–35. New York: Russell Sage Foundation.

Cooper, Harris. 2017. *Research Synthesis and Meta-Analysis, Fifth edition*. New York: SAGE Publications.

Cooper, Harris, and Larry V. Hedges. 2009a. "Potentials and Limitations." In *The Handbook of Research and Synthesis and Meta-Analysis, Second Edition*, edited by Harris Cooper, Larry V. Hedges, and Jeffrey C. Valentine, 561–72. New York: Russell Sage Foundation.

Cooper, Harris, and Larry V. Hedges. 2009b. "Research Synthesis as a Scientific Process." In *The Handbook of Research and Synthesis and Meta-Analysis, Second Edition*, edited by Harris Cooper, Larry V. Hedges, and Jeffrey C. Valentine, 3–16. New York: Russell Sage Foundation.

Cordray, David S., and Paul Morphy. 2009. "Research Synthesis and Public Policy." In *The Handbook of Research and Synthesis and Meta-Analysis, Second Edition*, edited by Harris Cooper, Larry V. Hedges, and Jeffrey C. Valentine, 473–93. New York: Russell Sage Foundation.

Dewald, William G., Jerry G. Thursby, and Richard G. Anderson. 1986. "Replication in Empirical Economics: The Journal of Money, Credit and Banking Project." *American Economic Review* 76 (4): 587–603.

Hamermesh, Daniel S. 2007. "Viewpoint: Replication in Economics." *Canadian Journal of Economics* 40 (3): 715–33.

Kane, Edward J. 1984. "Why Journal Editors Should Encourage the Replication of Applied Econometric Research." *Quarterly Journal of Business and Economics* 23 (1): 3–8.

McCullough, B. D., Kerry Anne McGeary, and Teresa D. Harrison. 2006. "Lessons from the JMCB Archive." *Journal of Money, Credit, and Banking* 38 (4): 1093–107.

McCullough, B. D., Kerry Anne McGeary, and Teresa D. Harrison. 2008. "Do Economics Journal Archives Promote Replicable Research?" *Canadian Journal of Economics* 41 (4): 1406–20.

Monroe, Don. 2009. "When Data is Not Enough." *Communications of the ACM* 58 (12): 12–14.

Neumark, David. 2015. "The Effects of Minimum Wages on Employment." *Federal Reserve Bank of San Francisco Economic Letter* 2015–37.

Roberts, Colin J. 2005. "Issues in Meta-Regression Analysis: An Overview." In *Meta-Regression Analysis: Issues of Publication Bias in Economics*, edited by Colin J. Roberts, and T. D. Stanley, 1–4. Hoboken: Wiley-Blackwell Publishing.

Roberts, Colin J., and T. D. Stanley. 2005. *Meta-Regression Analysis: Issues of Publication Bias in Economics*. Hoboken: Wiley-Blackwell Publishing.

Stanley, T. D. 2005. "Beyond Publication Bias." In *Meta-Regression Analysis: Issues of Publication Bias in Economics*, edited by Colin J. Roberts, and T. D. Stanley, 15–51. Hoboken: Wiley-Blackwell Publishing.

Stanley, T. D., and Stephen B. Jarrell. 1989. "Meta-Regression Analysis: A Quantitative Method of Literature Surveys." *Journal of Economic Surveys* 3 (2): 161–70, 2005.

Sutton, Alex J. 2009. "Publication Bias." In *The Handbook of Research and Synthesis and Meta-Analysis, Second Edition*, edited by Harris Cooper, Larry V. Hedges, and Jeffrey C. Valentine, 435–52. New York: Russell Sage Foundation.

This article has been cited by:

1. Hiroyoshi Iwata, Yoshikazu Miwa, Yumi Ikehara, Toshifumi Yodoshi, Shinichiro Ueda. 2022. The clinical value of hepatojugular reflux on congestive heart failure: A meta-analysis. *Journal of General and Family Medicine* **23**:6, 393-400. [Crossref]
2. Clay Collins, Brad R. Humphreys. 2022. Contest Outcome Uncertainty and Fan Decisions: A Meta-Analysis. *Journal of Sports Economics* **23**:6, 789-807. [Crossref]
3. Joern Block, Christopher Hansen, Holger Steinmetz. 2022. Are Family Firms Doing More Innovation Output With Less Innovation Input? A Replication and Extension. *Entrepreneurship Theory and Practice* **29**, 104225872210842. [Crossref]
4. Christopher Hansen, Holger Steinmetz, Jörn Block. 2022. How to conduct a meta-analysis in eight steps: a practical guide. *Management Review Quarterly* **72**:1, 1-19. [Crossref]
5. S Anukriti, Sonia Bhalotra, Eddy H F Tam. 2022. On the Quantity and Quality of Girls: Fertility, Parental Investments and Mortality. *The Economic Journal* **132**:641, 1-36. [Crossref]
6. Arjen van Witteloostuijn, Marcus Dejardin, Jeffrey M. Pollack. 2021. Replication in entrepreneurship. *Journal of Business Venturing Insights* **16**, e00271. [Crossref]
7. Fernando-Ignacio Sánchez-Martínez, Jorge-Eduardo Martínez-Pérez, José-María Abellán-Perpiñán, José-Luis Pinto-Prades. 2021. The value of statistical life in the context of road safety: new evidence on the contingent valuation/standard gamble chained approach. *Journal of Risk and Uncertainty* **63**:2, 203-228. [Crossref]
8. Émilie Caldeira, Ali Compaore, Alou Adessé Dama, Mario Mansour, Grégoire Rota-Graziosi. 2020. Effort fiscal en Afrique subsaharienne : les résultats d'une nouvelle base de données. *Revue d'économie du développement* **Vol. 27**:4, 5-51. [Crossref]
9. Jerome Geyer-Klingeberg, Markus Hang, Andreas Rathgeber. 2020. Meta-analysis in finance research: Opportunities, challenges, and contemporary applications. *International Review of Financial Analysis* **71**, 101524. [Crossref]
10. Kenneth Button. 2020. Studying the empirical implications of the liberalization of airport markets. *Competition and Regulation in Network Industries* **21**:3, 223-243. [Crossref]
11. Zareh Asatryan, Annika Havlik, Friedrich Heinemann, Justus Nover. 2020. Biases in fiscal multiplier estimates. *European Journal of Political Economy* **63**, 101861. [Crossref]
12. Silvester Van Koten. 2020. Forward premia in electricity markets: A replication study. *Energy Economics* **89**, 104812. [Crossref]
13. Kenneth Button. Is Regional Science Just Economics with a "dij" Added to All Equations? Some Thoughts of an Economist 23-42. [Crossref]
14. Eszter Czibor, David Jimenez-Gomez, John A. List. 2019. The Dozen Things Experimental Economists Should Do (More of). *Southern Economic Journal* **86**:2, 371-432. [Crossref]
15. Steven Cook, Jack Fosten. 2019. Replicating rockets and feathers. *Energy Economics* **82**, 139-151. [Crossref]
16. Kenneth Button. 2019. The value and challenges of using meta-analysis in transportation economics. *Transport Reviews* **39**:3, 293-308. [Crossref]
17. Frank Mueller-Langer, Benedikt Fecher, Dietmar Harhoff, Gert G. Wagner. 2019. Replication studies in economics—How many and which papers are chosen for replication, and why?. *Research Policy* **48**:1, 62-83. [Crossref]
18. Eszter Czibor, David Jimenez-Gomez, John A. List. 2019. The Dozen Things Experimental Economists Should Do (More Of). *SSRN Electronic Journal* **23**. . [Crossref]

19. Jerome Geyer-Klingeberg, Markus Hang, Andreas Rathgeber. 2019. Meta-Analysis in Financial Economics: Opportunities, Challenges, and Contemporary Applications. *SSRN Electronic Journal* . [Crossref]
20. Armenak Antinyan, Zareh Asatryan. 2019. Nudging for Tax Compliance: A Meta-Analysis. *SSRN Electronic Journal* . [Crossref]
21. Gerald Eric Daniels, Venoo Kakar. 2018. Normalized CES supply-side system approach: how to replicate Klump, McAdam, and Willman (2007). *Economics* **12**:1. . [Crossref]
22. Anna Josephson, Jeffrey D. Michler. 2018. Viewpoint: Beasts of the field? Ethics in agricultural and applied economics. *Food Policy* **79**, 1-11. [Crossref]
23. Randolph Luca Bruno, Maria Cipollina. 2018. A meta-analysis of the indirect impact of foreign direct investment in old and new EU member states: Understanding productivity spillovers. *The World Economy* **41**:5, 1342-1377. [Crossref]