Replication in Economics: A Progress Report

M. Duvendack, R.W. Palmer-Jones and W.R. Reed

The Future of Scholarly Communication in Economics
ZBW, Hamburg
30-31 March 2015
Outline

• Introduction
• Brief history of data sharing and replication in economics
• Survey of replication policies
• Analysis of a collection of replication studies
• Concluding remarks: The future of replication research
Introduction

- Many empirical results in economics are not reproducible; and/or not generalizable to alternative empirical specifications, econometric procedures, extensions of the data, and other modifications to the original study.
- This concern not restricted to economics and social science but the ‘hard science’ too.
- Popular media now concerned with replication: The Economist, the New Yorker, the Atlantic, BBC Radio, the Los Angeles Times … but focus on academic fraud while academic community concerned with production of disproportionate rate of false positives.
- Replication not a panacea for the problems facing scientific verifiability but it can provide useful checks on the spread of incorrect results.
Brief history of data sharing and replication in economics

• From the early days of applied economics it has been acknowledged that sharing of data is desirable:
  • “In statistical and other numerical work presented in ECONOMETRICA the original raw data will, as a rule, be published, unless their volume is excessive. This is important in order to stimulate criticism, control, and further studies” (Frisch, 1933:3).

• Despite these early calls, limited practice of replication by economists

• From 1960s onwards the Journal of Human Resources (JHR) included replication as part of their analysis

• Mid-1970s the Journal of Political Economy (JPE) initiated a “Confirmations and Contradictions” section which existed from 1976 to 1999
By the 1980s few major economics journals had data sharing or replication policies in place, exception: *Journal of Money, Credit and Banking (JMCB)* which requested authors to submit data and code.

Subsequently an increasing number of journals (e.g. *AER, JAE, Econometrica*) adopted data sharing/archiving or replication policies, either requiring authors to (i) provide data and code upon request or (ii) deposit their data and code in journal-managed data archives upon submission of their article.

However, these requirements were by no means sufficient as in some cases these policies were not strictly enforced.
(Dis-)Incentives for replication

- Distorted researcher and publisher incentive systems partially explain lack of replication in economics
- Rewards in academia for innovation, or novelty; since replication seemingly involves repetition it is almost by definition not a rewarding activity
- Discussions of incentives include 3 actors:
  - **Replicators** concerned with the time to undertake replications and publication opportunities
  - **Journal editors** concerned with (per page) citations of replications - lower compared to original papers. Also, reputational effects when controversy between replicators and original authors ensues
  - **Original authors** concerned with costs of compiling data and code into usable forms – benefits small.
- Replication used in “disciplinary” ways: Some replications led to constructive interactions but many conducted more hostile disincentives
- Replication can undermine the authority that pertains to the expertise and the experts who practice econometrics
Survey of replication policies

• Surveyed 333 economics journals asking:
  (i) Does the journal regularly publish data and code for its empirical research articles?
  (ii) Does the journal’s website explicitly mention that it publishes data and code

• 27 of 333 journals regularly publish data and code for empirical research

• 10 explicitly mention that they publish replications:

• The lack of publishing outlets is perhaps the most serious obstacle to researchers interested in undertaking replication research
Analysis of replication studies I

- Analysis of 161 replication studies published between 1977-2014
- Since 2000 publications of replication studies increased in frequency

- JAE most frequent publisher accounting for 1/5, followed by JHR, AER, Econ Journal Watch, JDS and Experimental Economics – they account for 60% of all published replication studies.
Analysis of replication studies II

• **Summary?** Was the published article a full study, or did it only summarize the key results from a study?

• **Exact?** Did the replication study attempt to exactly reproduce the original findings?

• **Extension?** Did the replication study go beyond attempting to reproduce the original results by extending the analysis to different types of subjects, time periods, or test additional hypotheses?

• **Original Results?** Did the replication study report the findings of the original study in a way that facilitated comparison of results without having to access the original study?

• **Negative? Mixed? Positive?** Did the replication study confirm or disconfirm the original study, or were the results mixed?

• **Reply?** Did the journal publish a reply or response from the original authors?
### Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All (161)</td>
<td>0.050</td>
<td>0.640</td>
<td>0.516</td>
<td>0.584</td>
<td>0.665</td>
<td>0.124</td>
<td>0.211</td>
<td>0.205</td>
</tr>
<tr>
<td>JAE (31)</td>
<td>0.194</td>
<td>0.742</td>
<td>0.290</td>
<td>0.323</td>
<td>0.452</td>
<td>0.194</td>
<td>0.355</td>
<td>0.032</td>
</tr>
<tr>
<td>Experimental (11)</td>
<td>0.000</td>
<td>0.727</td>
<td>0.818</td>
<td>0.545</td>
<td>0.545</td>
<td>0.182</td>
<td>0.273</td>
<td>0.091</td>
</tr>
<tr>
<td>Non-JAE/Non-Experimental (119)</td>
<td>0.017</td>
<td>0.605</td>
<td>0.546</td>
<td>0.655</td>
<td>0.731</td>
<td>0.101</td>
<td>0.168</td>
<td>0.261</td>
</tr>
</tbody>
</table>

**NOTE:** Numbers in the table are averages of the respective 0-1 dummy variables. The numbers in parentheses in the Journals column indicates the number of replication studies in each journal category.

- **What can we learn from our analysis of replication studies?**
  - Main point: There appears to be a high rate of disconfirmation, approximately two out of every three studies disconfirmed the original findings.
Summary

- The importance of making data available to researchers in order to enable replication has long been noted.
- Publication of replication studies has been increasing slowly, however, there are few journals that publish replication studies → 6 journals account for approx 60% of all published replication studies.
- Only 10 journals have ever published more than 3 replication studies, JAE and JHR are exceptions.
- It is quite common to find that major results from empirical research in economics journals cannot be confirmed.
- Approx. four out of five replication studies fail to confirm one or more major findings from the original research.
Concluding thoughts I

- Many replication initiatives under way:
  - Reproducibility Initiative by the Center for Open Science
  - Berkeley Initiative for Transparency in the Social Sciences (BITSS)
  - Political Science replication initiative led by Gary King
  - “Replication in Economics” project at Goettingen University funded by Institute for New Economic Thinking

- Key forces that have driven the expansion of replication research in recent decades are:
  - the increasing availability of data and code,
  - technological innovations in the allocation of journal space,
  - societal factors that affects “tastes” for replication research – are likely to expand the use of replications in the future.
Concluding thoughts II

• Replication and publication bias:
  • Replication can potentially mitigate publication bias and uncover Type I errors

• Using replications more effectively:
  • Better Integration of replication and meta-regression
  • Meta-regression identifies study characteristics that “explain” why different studies reach different conclusions while replication studies then take these results and examine whether changing the empirical design of a study has the effect predicted by meta-analysis

• Replication is no panacea but a useful tool for assessing the reliability and validity of empirical results and we hope to further this cause.
Thank you!

For more details:
www.replicationnetwork.com

Contact

Dr Maren Duvendack
School of International Development
University of East Anglia
Norwich, United Kingdom
Email: m.duvendack@uea.ac.uk

Professor W. Robert Reed
Department of Economics and Finance
University of Canterbury
Christchurch, New Zealand
Email: bobreednz@yahoo.com