Data Sharing and Replication

Enabling Reproducible Research

Garret Christensen[1]

[1] UC Berkeley: Berkeley Initiative for Transparency in the Social Sciences
Berkeley Institute for Data Science

APHRC, Summer 2015

Outline

1 Introduction

2 Project Protocol, Reporting Standards

3 Data Sharing

4 Replication

5 Conclusion

Reproducibility & Transparency

- What are problems associated with reproducibility?
- What are solutions to these problems?
- What are practical tools to implement these solutions?

Introduction

Science advances by building on the work of others.

> *If I have seen further, it is by standing on the shoulders of giants*

–Sir Isaac Newton, 1676

Problems

What prevents us from building on others' work?

- Data not shared
- Analysis not shared
- Methods/protocol not shared

Solutions

What enables us to build on others' work?

- Data shared in trusted public repository
- Code/Analysis shared in trusted public repository
- Methods/protocol follow appropriate reporting standard
- Also: findings/scholarly publications available (open access)

Project Protocol, Reporting Standards

Make sure you report everything another researcher would need to replicate your research, including the exact methods.

What to report (following medicine):

- Find the appropriate reporting standard for your field and follow it.
- Enhancing the QUAlity and Transparency Of health Research (EQUATOR Network)
- The most widely-adopted standard: Consolidated Standards of Reporting Trials (CONSORT).
- Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT Statement).

Project Protocol, Reporting Standards

Make sure you report everything another researcher would need to replicate your research, including the exact methods.

What to report (following medicine):

- Find the appropriate reporting standard for your field and follow it.
- Enhancing the QUAlity and Transparency Of health Research (EQUATOR Network)
- The most widely-adopted standard: Consolidated Standards of Reporting Trials (CONSORT).
- Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT Statement).

Project Protocol, Reporting Standards

Where to report:
If not in the methods section of the article (of limited length), supplementary online appendix linked with article or in trusted digital repository.

Data Sharing

- To build on the work of others, data must be shared.
- Data sharing is associated with more citations (causality unclear). Piwowar et al. 2007

Data Sharing

History in Economics:

- Journal of Money Credit and Banking Project: Dewald, Thursby, Anderson *AER* 1986.
  - Low response rate to requests to share data.
  - Attempted to reproduce 9 papers, problems with all (some minor) even with help of original authors.

TABLE 1—RESPONSES TO REQUESTS FOR DATA FROM AUTHORS OF EMPIRICAL PAPERS[a]

	Published before Data Requested	Accepted before Data Requested	Under Review when Data Requested
Requests	62	27	65
Responses	42	26	49
Response Rate (Percent)	66	96	75
Mean Response Time (Days)	217	125	130
Not Submitted:			
Confidential Data	2	1[b]	0
Lost or Destroyed Data	14	2	1
Data Available, But Not Sent[c]	4	2	1
Nonrespondents	20	1	16
Total Not Submitted	40	6	18
Nonsubmission Rate (Percent)	66	22	28

[a] Includes all requests made through December 1984, and excludes authors whose papers were rejected.

[b] Two data sets were partially confidential.

[c] This category includes authors who (*i*) stated that their data were available from published sources, but did not send their data; and (*ii*) authors who claimed to have their data but were unwilling to sort through their papers to find the data.

Data Sharing

History in Economics:

- Journal of Money Credit and Banking Project: Dewald, Thursby, Anderson *AER* 1986.
  - Low response rate to requests to share data.
  - Attempted to reproduce 9 papers, problems with all (some minor) even with help of original authors.
- A Decade After JMCB: Anderson and Dewald, St Louis Fed 1994.
  - Repeated similar experiment
  - Similar bleak results
- Verifying the Solution from a Nonlinear Solver, McCullough and Vinod, *AER* 2003.
  - Different software programs get you different answers.
  - But finally change—*AER* institutes data sharing requirement. Policy

Data Sharing

History in Economics:

- Journal of Money Credit and Banking Project: Dewald, Thursby, Anderson *AER* 1986.
  - Low response rate to requests to share data.
  - Attempted to reproduce 9 papers, problems with all (some minor) even with help of original authors.
- A Decade After JMCB: Anderson and Dewald, St Louis Fed 1994.
  - Repeated similar experiment
  - Similar bleak results
- Verifying the Solution from a Nonlinear Solver, McCullough and Vinod, *AER* 2003.
  - Different software programs get you different answers.
  - But finally change—*AER* institutes data sharing requirement. Policy

Data Sharing

How are we doing as a discipline?

- *AER* internal review generally positive (Glandon 2010)
- Many, including McCullough, still skeptical of the ability to reproduce (Econ Journal Watch, 2007)
- Though *AER*, all AEA, and other top journals have a good policy, enforcement is limited, and shared data is often only the “analysis” data instead of raw data, and *QJE* has no policy whatsoever.
- A study by the Replication Network shows that fewer than 27 journals regularly publish data, only 10 explicitly state they publish replications. (Duvendack et al 2015)

Data Sharing

Why share your data in a trusted public repository?

- Find the appropriate repository: `http://www.re3data.org/`
- Repositories will last longer than your own website.
- Repositories are more easily searchable by other researchers.
- Repositories will store your data in a non-proprietary format that won't become obsolete.
- Repositories manage meta-data better.
- Repositories create digital citable identifiers (DOI).

Data Sharing

Examples of Trusted Repositories:

- Harvard's Dataverse
- Data Dryad
- figshare
- Open Science Framework
- Check the journal–they may use one of these
  - *REStat*'s Dataverse

APHRC Repository

APHRC has created the APHRC Microdata Portal

- 30 Studies and growing
- `http://aphrc.org/catalog/microdata/index.php/catalog`
- Managed by Cheikh Faye

Replication

- With data available, we can begin to replicate studies.
- We should be very careful about what we mean by “replication.”
- “The Meaning of Failed Replications” Michael Clemens, CGD Working Paper 399.

Table 1: A Proposed Definition to Distinguish Replication and Robustness Tests

	Sampling distribution for parameter estimates	Sufficient conditions for discrepancy	Types	Methods in follow-up study versus methods *reported* in original:			Examples
				Same specification	Same population	Same sample	
Replication	*Same*	*Random chance, error, or fraud*	Verification	*Yes*	*Yes*	*Yes*	*Fix faulty measurement, code, dataset*
			Reproduction	*Yes*	*Yes*	*No*	*Remedy sampling error, low power*
Robustness	*Different*	*Sampling distribution has changed*	Reanalysis	*No*	*Yes*	*Yes/No*	*Alter specification, recode variables*
			Extension	*Yes*	*No*	*No*	*Alter place or time; drop outliers*

The "same" specification, population, or sample means the same as *reported* in the original paper, not necessarily what was contained in the code and data used by the original paper. Thus for example if code used in the original paper contains an error such that it does not run exactly the regressions that the original paper said it does, new code that fixes the error is nevertheless using the "same" specifications (as described in the paper).

Replication

Why Replicate? Motivation and suggestions from Nicole Janz of Political Science Replication and Cambridge University

- For science in general:
  - Uncover misconduct and sloppy science
  - Confirm previous findings and generalizability
  - Point to misuse of statistical methods
- For you as researchers:
  - Learn statistics
  - Jump to research frontier
  - Publish
  - Make your own research routinely reproducible
  - Fun

Replication

Why Replicate? Motivation and suggestions from Nicole Janz of Political Science Replication and Cambridge University

- For science in general:
  - Uncover misconduct and sloppy science
  - Confirm previous findings and generalizability
  - Point to misuse of statistical methods
- For you as researchers:
  - Learn statistics
  - Jump to research frontier
  - Publish
  - Make your own research routinely reproducible
  - Fun

Replication

Why Replicate? Motivation and suggestions from Nicole Janz of Political Science Replication and Cambridge University

- For science in general:
  - Uncover misconduct and sloppy science
  - Confirm previous findings and generalizability
  - Point to misuse of statistical methods
- For you as researchers:
  - Learn statistics
  - Jump to research frontier
  - Publish
  - Make your own research routinely reproducible
  - Fun

Replication

Which study should you pick to replicate?

- Don’t select a study with methods that you don’t know or can’t learn within a reasonable time.
- Pick a recent study (<5 yo) from a good journal.
- Data (and code) should be publicly available.
- The journal that published the original study has published replications before.

Replication

Which journals publish replications?

- List from The Replication Network study, Duvendack et al.
- Sadly fairly limtied in economics (10).
- Selected journals from Janz (2015)

TABLE 2. Journals whose websites explicitly mention that they publish replications

1)	*Econ Journal Watch*
2)	*Economic Development and Cultural Change*
3)	*Economics of Education Review*
4)	*Empirical Economics*
5)	*Experimental Economics*
6)	*Explorations in Economic History*
7)	*International Journal of Forecasting*
8)	*Jahrbücher für Nationalökonomie und Statistik/Journal of Economics and Statistics*
9)	*Journal of Applied Econometrics*
10)	*Review of International Organizations*

Journals Open to Replication (selection)

Political Science

THE JOURNAL OF conflict RESOLUTION

R &P

Research & Politics ^

PRQ POLITICAL RESEARCH QUARTERLY

*

American Political Science Review

*

Psychology

Psychological SCIENCE

Perspectives on Psychological SCIENCE

Social Psychology

+

AMERICAN PSYCHOLOGIST

Economics

Empirical Economics

Journal of APPLIED ECONOMETRICS

Public Finance Review pfr

#

The American Economic Review

*original study was published in the same journal

+ home of the original 'Many Labs' project

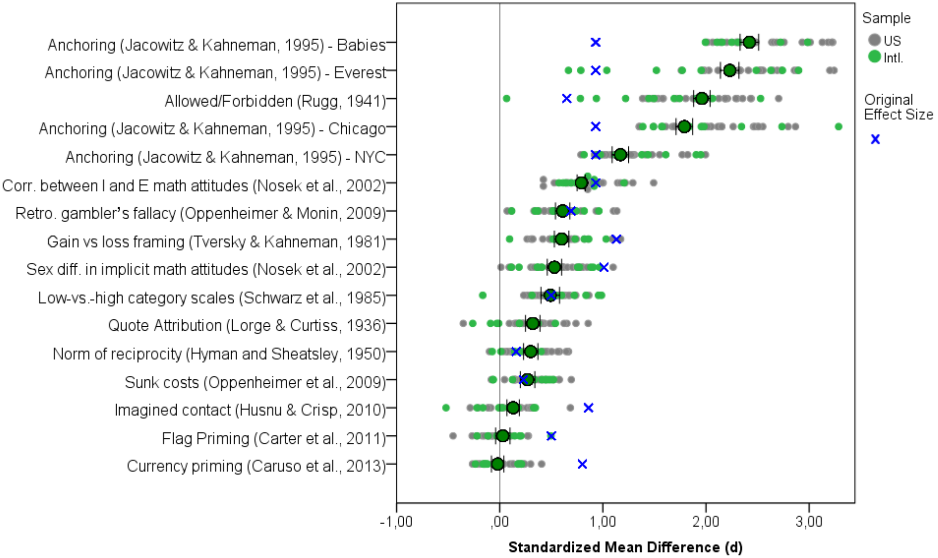
special issue dedicated to replications (March 2015)

^this journal invites replication studies

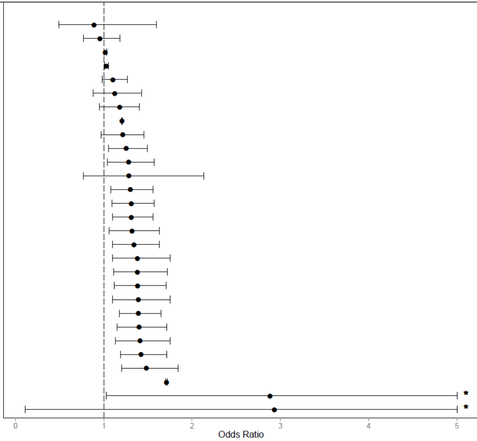
Replication

How exactly to replicate?

- Be systematic: write a pre-analysis plan.
- Don’t just go on a fishing expedition. We all know that if you dig hard enough, you can find a specification that makes results appear weaker. Don’t selectively report those specifications.
- Be courteous and professional.
- Take an entirely systematic approach:
  - Many Labs Project
  - Crowdsource your analysis


Anchoring (Jacowitz & Kahneman, 1995) - Babies
Anchoring (Jacowitz & Kahneman, 1995) - Everest
Allowed/Forbidden (Rugg, 1941)
Anchoring (Jacowitz & Kahneman, 1995) - Chicago
Anchoring (Jacowitz & Kahneman, 1995) - NYC
Corr. between I and E math attitudes (Nosek et al., 2002)
Retro. gambler's fallacy (Oppenheimer & Monin, 2009)
Gain vs loss framing (Tversky & Kahneman, 1981)
Sex diff. in implicit math attitudes (Nosek et al., 2002)
Low-vs.-high category scales (Schwarz et al., 1985)
Quote Attribution (Lorge & Curtiss, 1936)
Norm of reciprocity (Hyman and Sheatsley, 1950)
Sunk costs (Oppenheimer et al., 2009)
Imagined contact (Husnu & Crisp, 2010)
Flag Priming (Carter et al., 2011)
Currency priming (Caruso et al., 2013)
-1,00
,00
1,00
2,00
3,00
Standardized Mean Difference (d)
Sample
US
Intl.
Original Effect Size

Team	Analytic Approach	OR
12	Zero−inflated Poisson regression	0.89
17	Bayesian logistic regression	0.96
15	Hierarchical log−linear modeling	1.02
10	Multilevel regression and logistic regression	1.03
18	Hierarchical Bayes model	1.10
31	Logistic regression	1.12
1	Ordinary least squares with robust standard errors, logistic regression	1.18
4	Spearman correlation	1.21
14	Weighted least squares regression with referee fixed−effects and clustered standard errors	1.21
11	Multiple linear regression	1.25
30	Clustered robust binomial logistic regression	1.28
6	Linear Probability Model	1.28
26	Three−level hierarchical generalized linear modeling with Poisson sampling	1.30
3	Multilevel Binomial Logistic Regression using bayesian inference	1.31
23	Mixed model logistic regression	1.31
16	Hierarchical Poisson Regression	1.32
2	Linear probability model, logistic regression	1.34
5	Generalized linear mixed models	1.38
24	Multilevel logistic regression	1.38
28	Mixed effects logistic regression	1.38
32	Generalized linear models for binary data	1.39
8	Negative binomial regression with a log link analysis	1.39
20	Cross−classified multilevel negative binomial model	1.40
13	Poisson Multi−level modeling	1.41
25	Multilevel logistic binomial regression	1.42
9	Generalized linear mixed effects models with a logit link function	1.48
7	Dirichlet process Bayesian clustering	1.71
21	Tobit regression	2.88
27	Poisson regression	2.93



Conclusion

- Science builds on previous work
- To do that, work must be public
- Share your data and code publicly
- Replicate the work of others