



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Credibility Crisis and some Solutions: Registration and Pre-Analysis Plans

Garret Christensen^{1,2}
Fernando Hoces de la Guardia¹

¹UC Berkeley:
Berkeley Initiative for Transparency in the Social Sciences
²Berkeley Institute for Data Sciences

American Institutes for Research, May 2018
Slides available online at <https://goo.gl/RtfxqX>



Outline

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

1 Introduction

2 Ethical Research

3 Publication Bias — Registrations

4 P-Hacking — Pre-Analysis Plans

5 Replications — Guidelines & Tools

6 Conclusion



Outline

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

1 Introduction

2 Ethical Research

3 Publication Bias — Registrations

4 P-Hacking — Pre-Analysis Plans

5 Replications — Guidelines & Tools

6 Conclusion



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES



ICPSR



Outline

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

1 Introduction

2 Ethical Research

3 Publication Bias — Registrations

4 P-Hacking — Pre-Analysis Plans

5 Replications — Guidelines & Tools

6 Conclusion



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Ethical Research

- Transparency is part of being an ethical researcher.
- Scientific values espoused by Robert Merton (Merton 1942):
 - Universalism: anyone can make a claim regardless of status.
 - Communism: open sharing of knowledge.
 - Disinterestedness: truth as motivation, not financial gains (COI).
 - Organized skepticism: peer review, replication.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Ethical Research

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

- Fraud exists (Simonsohn 2013), but mostly we should admit that we're human, subject to bias and motivated reasoning, transparency can help with this (Nosek, Spies, Motyl 2012).
- Those of us who run experiments or use data with personal identifying information should take IRBs seriously as part of transparency (Ch. 11–13 Morton & Williams 2010, Desposato 2014).



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Ethical Research

Why we worry:

- (Anderson, Martinson, De Vries 2007)
- (John, Loewenstein, Prelec 2011)

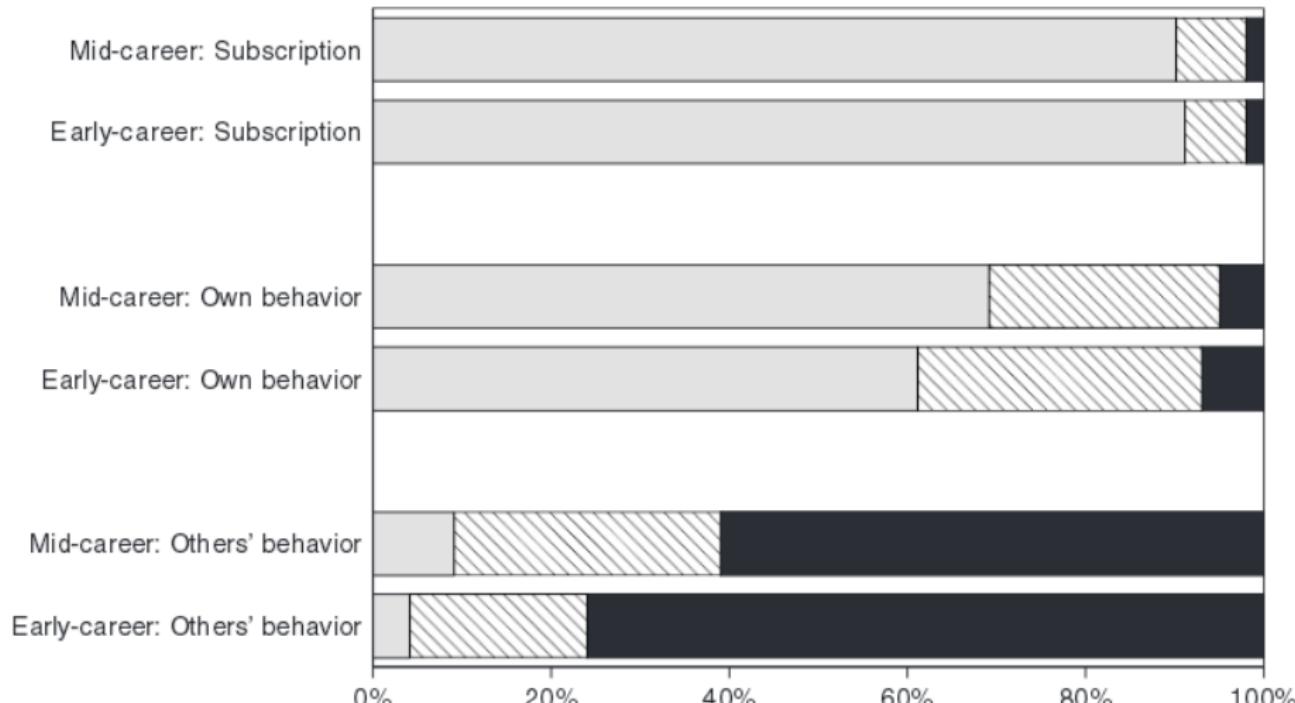


FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

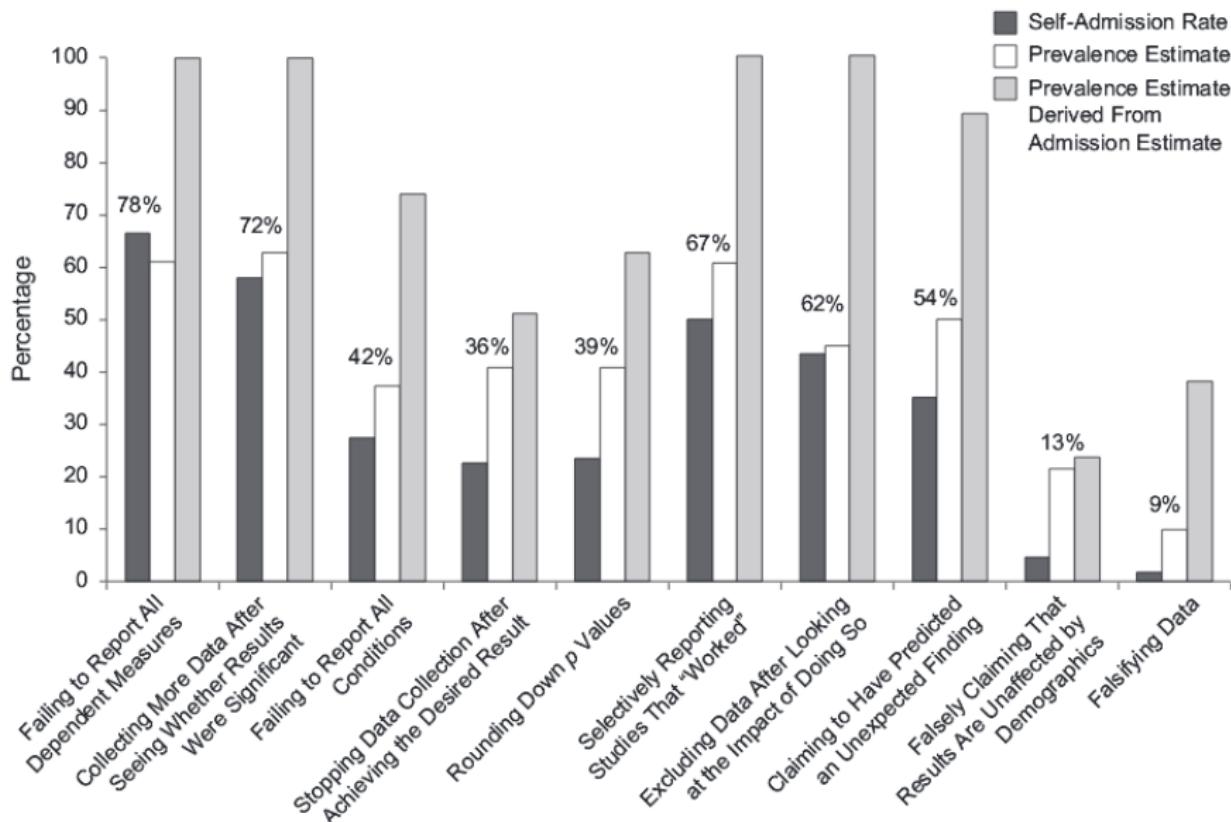


Fig. 1. Results of the Bayesian-truth-serum condition in the main study. For each of the 10 items, the graph shows the self-admission rate, prevalence estimate, prevalence estimate derived from the admission estimate (i.e., self-admission rate/admission estimate), and geometric mean of these three percentages (numbers above the bars). See Table I for the complete text of the items.



Outline

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

1 Introduction

2 Ethical Research

3 Publication Bias — Registrations

4 P-Hacking — Pre-Analysis Plans

5 Replications — Guidelines & Tools

6 Conclusion



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Publication Bias

Existence of the problem:

- Effect sizes diminish with sample size (Gerber, Green, Nickerson 2001)
- There is a higher fraction of rejected hypothesis tests in social compared to hard sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- Data on the complete set of experiments run shows strong results are 40pp more likely to be published, and 60pp more likely to be written up. The file drawer problem is large. (Franco, Malhotra, Simonovits 2014)

All Fields

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

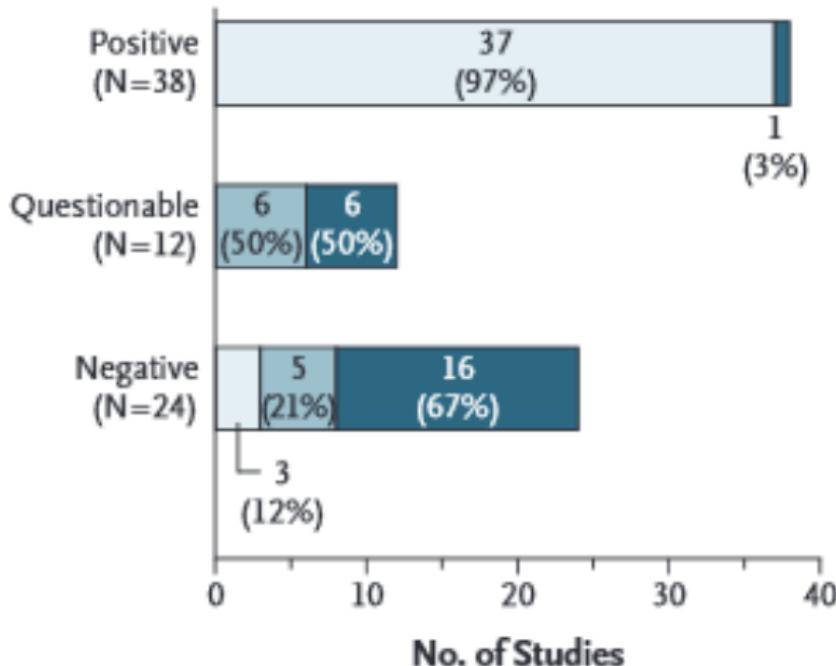
Conclusion

- Medicine: (Turner et al. 2008)
- Social Sciences: (Franco, Malhotra, Simonovits 2014)
- Economics: (Brodeur et al. 2016)
- Sociology: (Gerber and Malhotra 2008)
- Political Science: (Gerber and Malhotra 2008)

- Published, agrees with FDA decision
- Published, conflicts with FDA decision
- Not published

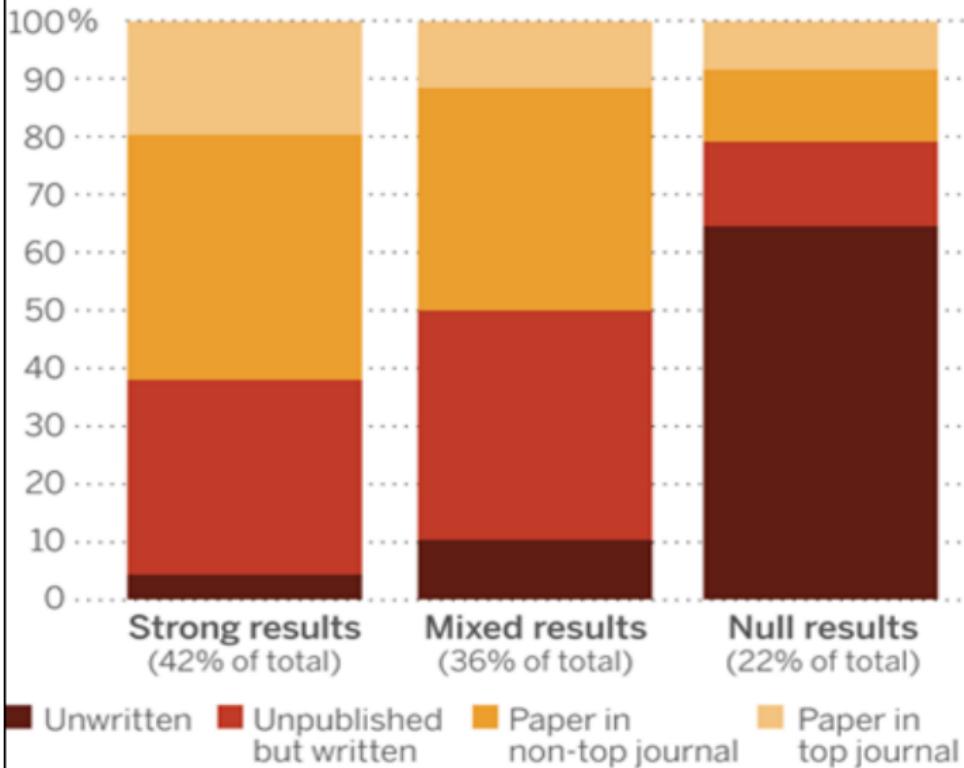
A Studies (N=74)

FDA Decision



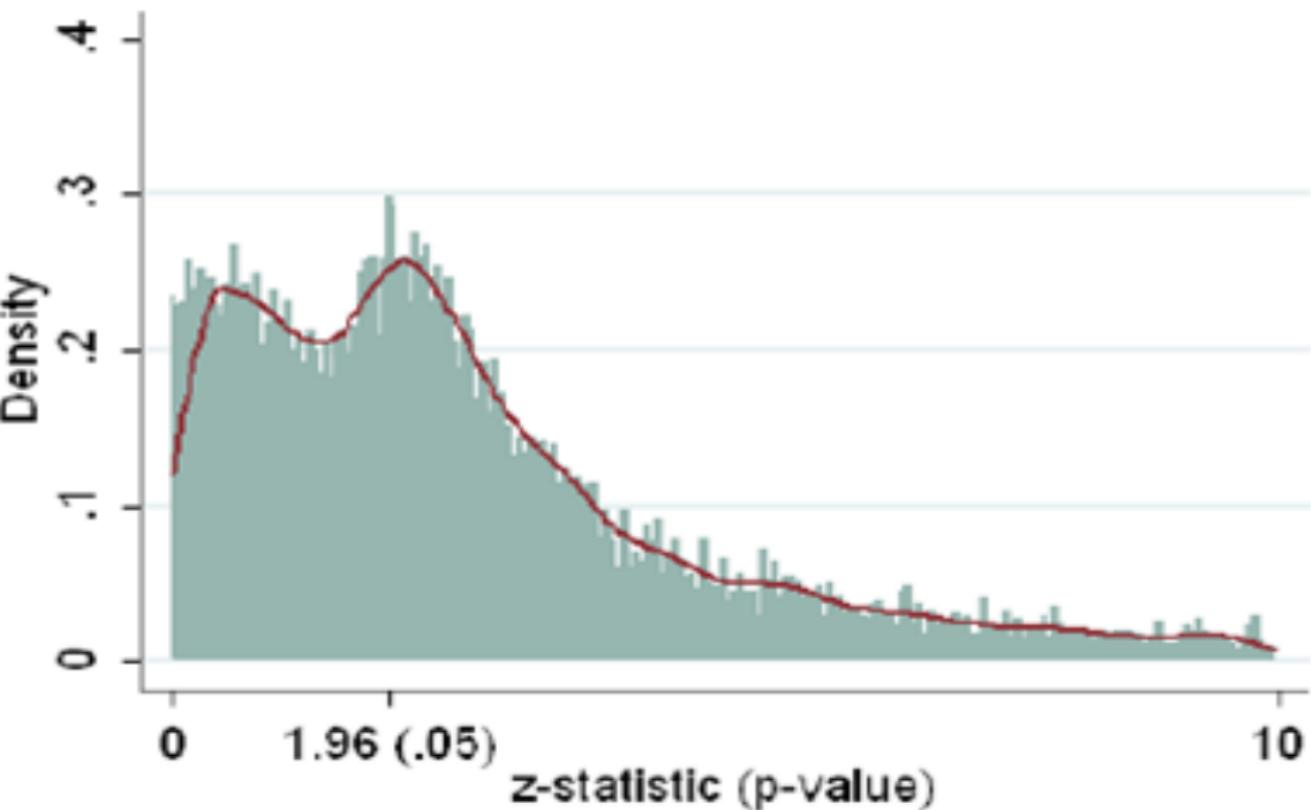
Most null results are never written up

The fate of 221 social science experiments

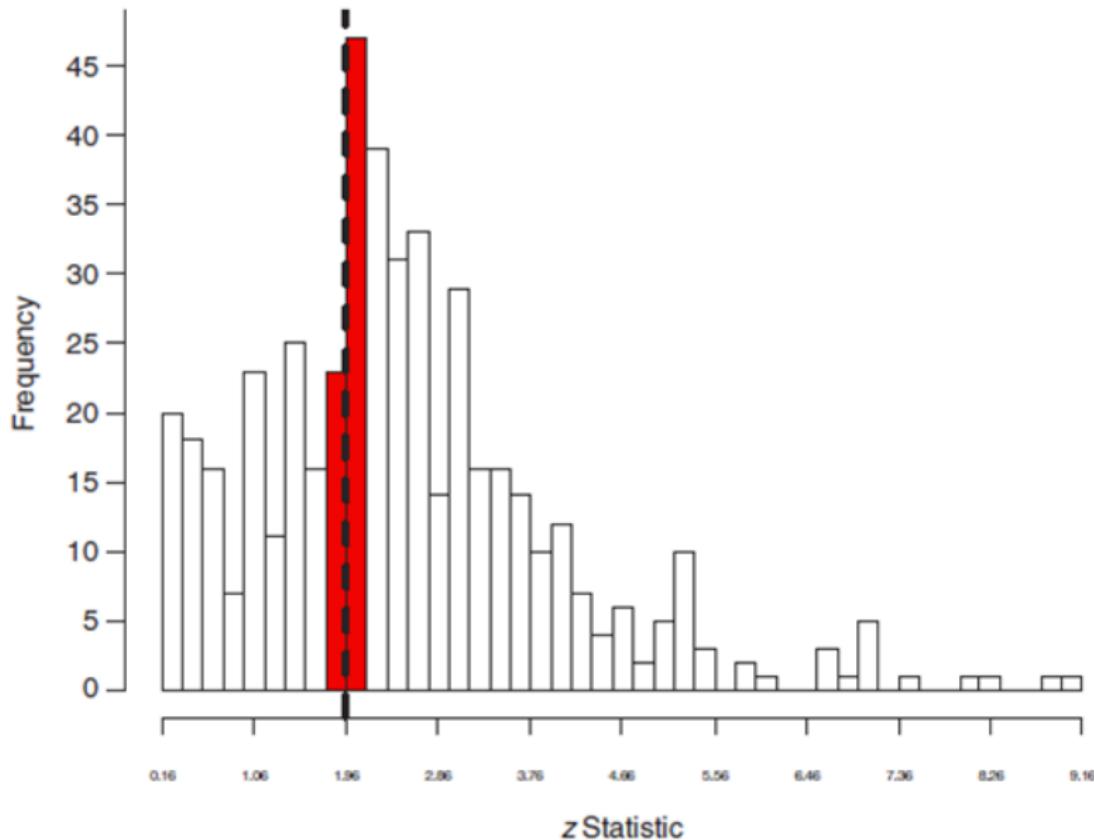


Source: A. Franco et al., *Science* (28 August)

(b) Unrounded distribution of z-statistics.



Histogram of z Statistics From the *American Sociological Review*, the *American Journal of Sociology*, and *The Sociological Quarterly* (Two-Tailed)



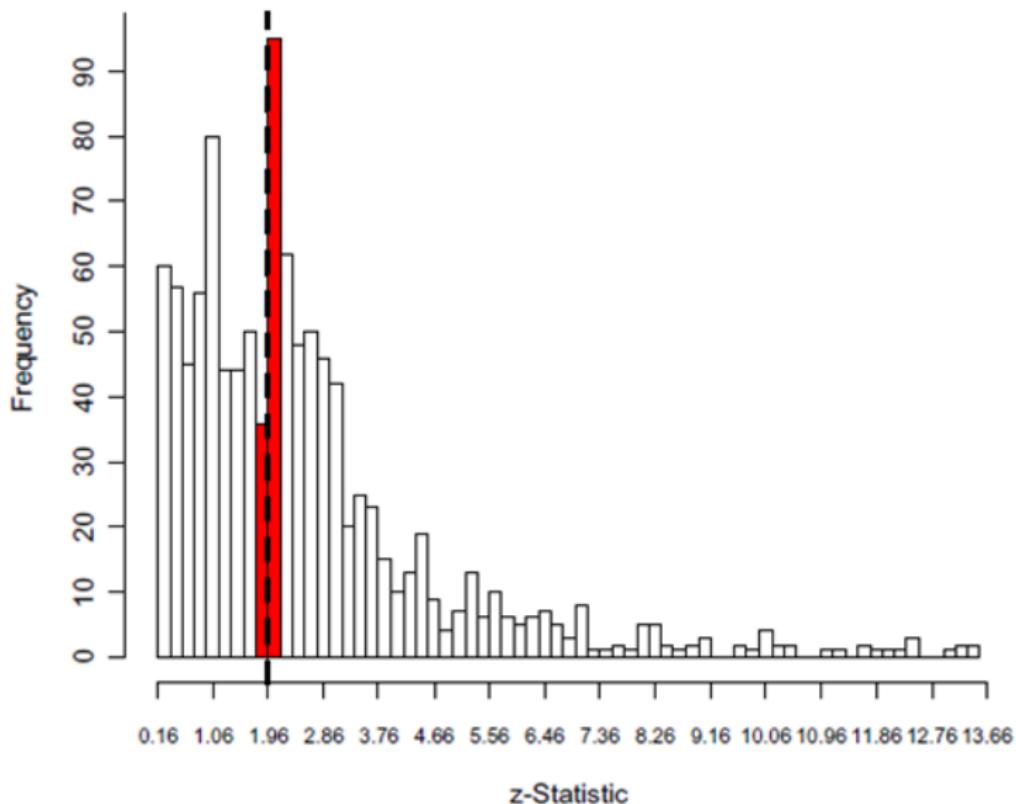


Figure 1(a). Histogram of z -statistics, *APSR* & *AJPS* (Two-Tailed). Width of bars (0.20) approximately represents 10% caliper. Dotted line represents critical z -statistic (1.96) associated with $p = 0.05$ significance level for one-tailed tests.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility Crisis and some Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Publication Bias

If we only write up/publish significant results, and we have no record of all the insignificant results, we have no way to tell if our ‘significant’ results are real, or if they’re the 5% we should expect due to noise.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration

Registration as Solution to Publication Bias:

- Publicly stating all research you will do, what hypotheses you will test, prospectively.
- Near universal adoption in medical RCTs. Top journals (ICMJE) won't publish if it's not registered.
<http://clinicaltrials.gov>
- Even better if registry requires outcomes from after study. Currently limited, but NIH is moving on this.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

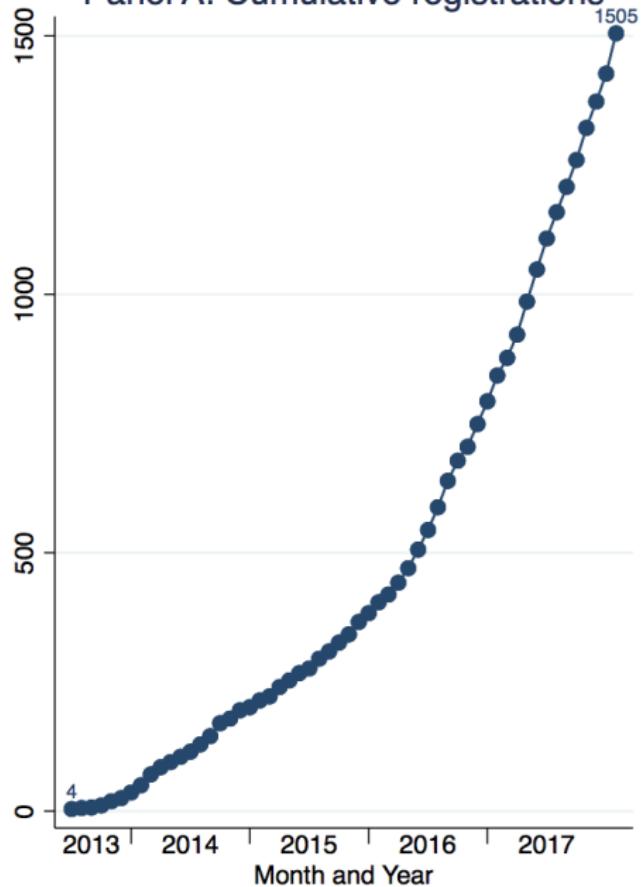
Conclusion

Registration

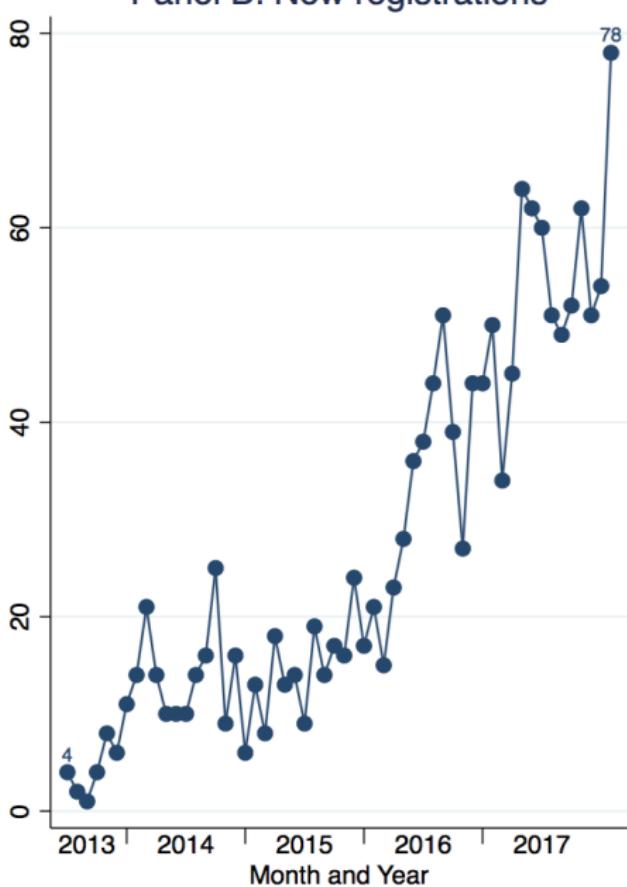
Newer to social sciences, but:

- AEA registry, currently only for RCTs.
<http://socialscienceregistry.org>
- EGAP registry
<http://egap.org/design-registration>
- 3ie registry, for developing country evaluations.
<http://ridie.3ieimpact.org>
- Open Science Framework
<http://osf.io>
 - Open format
 - Will soon sync with above
- Simple: <http://aspredicted.org>

Panel A: Cumulative registrations



Panel B: New registrations





BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Hands On Activity: #1 Little Experiment

Explanation to participants:

- Read and complete the sheet. Do not look at others sheets.
- Go to the website below and complete with your answers.

<https://goo.gl/aj8W61>



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Explanation to researchers

- You participated in (highly simplified) version of **The Ultimatum Game**
- The goal of the UG is to measure attitudes about fairness and/or expectations about (econ) rational behavior.
- Our little experiment was trying to measure if the responses to the UG can be anchored by a completely irrelevant number: **The ID number at the beginning of your sheet!**



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Explanation to researchers

- Treatment was receiving an ID number between 960 and 999.
- Control receive an ID number between 10 and 49.
- Outcome: Offer made in the UG



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Hands On Activity: #2 Registration.

Based on our little experiment:

- Explore AEA Registry at
www.socialscienceregistry.org
- Create a draft of using Open Science Framework at
osf.io:
 - Open format
 - AsPredicted (will work with this one)



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

Using Aspredicted format:

- **Research question:** Does exposure to a large number increases the offer made in the ultimatum game?
- **Dependent variable:** Amount offered in the ultimatum game.
- **Treatment:** Participant will be randomly assigned a large number ([960, 999]) or a small number ([10, 49]) to be read and remember, before reading the ultimatum game question.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

Using Aspredicted format:

- **Research question:** Does exposure to a large number increases the offer made in the ultimatum game?
- **Dependent variable:** Amount offered in the ultimatum game.
- **Treatment:** Participant will be randomly assigned a large number ([960, 999]) or a small number ([10, 49]) to be read and remember, before reading the ultimatum game question.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

Using Aspredicted format:

- **Research question:** Does exposure to a large number increases the offer made in the ultimatum game?
- **Dependent variable:** Amount offered in the ultimatum game.
- **Treatment:** Participant will be randomly assigned a large number ([960, 999]) or a small number ([10, 49]) to be read and remember, before reading the ultimatum game question.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

Using Aspredicted format:

- **Research question:** Does exposure to a large number increases the offer made in the ultimatum game?
- **Dependent variable:** Amount offered in the ultimatum game.
- **Treatment:** Participant will be randomly assigned a large number ([960, 999]) or a small number ([10, 49]) to be read and remember, before reading the ultimatum game question.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

Using Aspredicted format:

- **Research question:** Does exposure to a large number increases the offer made in the ultimatum game?
- **Dependent variable:** Amount offered in the ultimatum game.
- **Treatment:** Participant will be randomly assigned a large number ([960, 999]) or a small number ([10, 49]) to be read and remember, before reading the ultimatum game question.



Registration of our Little Experiment

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Using Aspredicted format:

- **Research question:** Does exposure to a large number increases the offer made in the ultimatum game?
- **Dependent variable:** Amount offered in the ultimatum game.
- **Treatment:** Participant will be randomly assigned a large number ([960, 999]) or a small number ([10, 49]) to be read and remember, before reading the ultimatum game question.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

Using Aspredicted format:

- **Research question:** Does exposure to a large number increases the offer made in the ultimatum game?
- **Dependent variable:** Amount offered in the ultimatum game.
- **Treatment:** Participant will be randomly assigned a large number ([960, 999]) or a small number ([10, 49]) to be read and remember, before reading the ultimatum game question.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility Crisis and some Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

- **Analyses:** OLS regression of amount offer as dependant variable and treatment as regressor.
- **Outliers and Exclusions:** Will exclude participants with missing information in any field. Amounts beyond plausible values (eg offers above max dollar value) will be top-coded.
- **Sample size:** We will define our sample by the number of participants in the workshop.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility Crisis and some Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

- **Analyses:** OLS regression of amount offer as dependant variable and treatment as regressor.
- **Outliers and Exclusions:** Will exclude participants with missing information in any field. Amounts beyond plausible values (eg offers above max dollar value) will be top-coded.
- **Sample size:** We will define our sample by the number of participants in the workshop.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Registration of our Little Experiment

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

- **Analyses:** OLS regression of amount offer as dependant variable and treatment as regressor.
- **Outliers and Exclusions:** Will exclude participants with missing information in any field. Amounts beyond plausible values (eg offers above max dollar value) will be top-coded.
- **Sample size:** We will define our sample by the number of participants in the workshop.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

- **Analyses:** OLS regression of amount offer as dependant variable and treatment as regressor.
- **Outliers and Exclusions:** Will exclude participants with missing information in any field. Amounts beyond plausible values (eg offers above max dollar value) will be top-coded.
- **Sample size:** We will define our sample by the number of participants in the workshop.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

- **Analyses:** OLS regression of amount offer as dependant variable and treatment as regressor.
- **Outliers and Exclusions:** Will exclude participants with missing information in any field. Amounts beyond plausible values (eg offers above max dollar value) will be top-coded.
- **Sample size:** We will define our sample by the number of participants in the workshop.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

- **Analyses:** OLS regression of amount offer as dependant variable and treatment as regressor.
- **Outliers and Exclusions:** Will exclude participants with missing information in any field. Amounts beyond plausible values (eg offers above max dollar value) will be top-coded.
- **Sample size:** We will define our sample by the number of participants in the workshop.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Registration of our Little Experiment

- **Analyses:** OLS regression of amount offer as dependant variable and treatment as regressor.
- **Outliers and Exclusions:** Will exclude participants with missing information in any field. Amounts beyond plausible values (eg offers above max dollar value) will be top-coded.
- **Sample size:** We will define our sample by the number of participants in the workshop.

Meta-Analysis

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Synthesize results systematically

Organizations:

- Cochrane Collaboration (Medicine)
- Campbell Collaboration (Policy)
- What Works Clearinghouse (US Gov't, Education)
- CLEAR (US Gov't, Labor)
- MAER-NET (Economics)



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Meta-Analysis

Tools:

- Funnel plots of sample size vs. effect size or precision (Card & Krueger 1995)
- Funnel Asymmetry Test (Stanley & Doucouliagos 2012)
- P-curve (Simonsohn et al. 2014)
 - ▶ Online App
 - One for all P-checker
 - ▶ Shiny App

Relevant? Null Results Increase After Requiring Registrations (Kaplan & Irvin 2015)

Credibility Crisis and some Solutions

Introduction

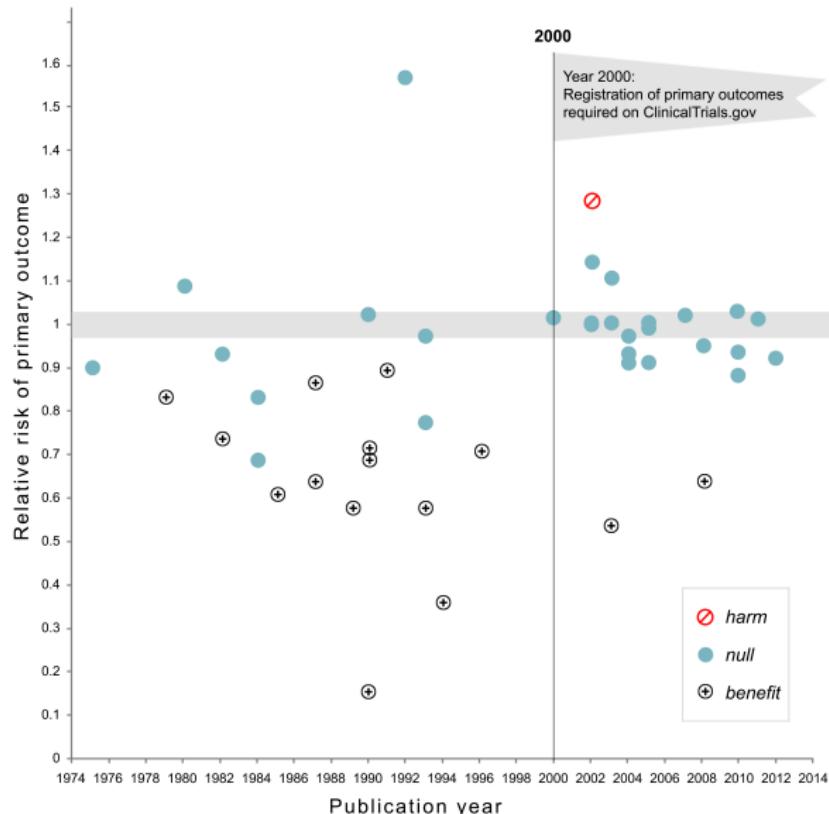
Ethical Research

Publication Bias — Registrations

P-Hacking — Pre-Analysis Plans

Replications — Guidelines & Tools

Conclusion





BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility Crisis and some Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Outline

1 Introduction

2 Ethical Research

3 Publication Bias — Registrations

4 P-Hacking — Pre-Analysis Plans

5 Replications — Guidelines & Tools

6 Conclusion

P-Hacking

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Define the problem:

- Definition: flexibility in data analysis allows portrayal of *anything* as below an arbitrary p-value threshold; significance loses its meaning.
- Also called fishing, researcher degrees of freedom, or data-mining.
- Not something only evil people do. It's subconscious, or simply built into statistics (Gelman, Loken 2013).

The Threat of P-Hacking (Casey et al. 2012)

| Outcome variable | (1) Mean for controls | (2) Treatment effect |
|---|-----------------------------|----------------------------|
| Panel A: GoBifo “weakened” institutions | | |
| Attended meeting to decide what to do with the tarp | 0.81 | -0.04 ⁺ |
| Everybody had equal say in deciding how to use the tarp | 0.51 | -0.11 ⁺ |
| Community used the tarp (verified by physical assessment) | 0.90 | -0.08 ⁺ |
| Community can show research team the tarp | 0.84 | -0.12* |
| Respondent would like to be a member of the VDC | 0.36 | -0.04* |
| Respondent voted in the local government election (2008) | 0.85 | -0.04* |
| Panel B: GoBifo “strengthened” institutions | | |
| Community teachers have been trained | 0.47 | 0.12 ⁺ |
| Respondent is a member of a women’s group | 0.24 | 0.06** |
| Someone took minutes at the most recent community meeting | 0.30 | 0.14* |
| Building materials stored in a public place when not in use | 0.13 | 0.25* |
| Chiefdom official did not have the most influence over tarp use | 0.54 | 0.06* |
| Respondent agrees with “Responsible young people can be good leaders” and not “Only older people are mature enough to be leaders” | 0.76 | 0.04* |
| Correctly able to name the year of the next general elections | 0.19 | 0.04* |



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility Crisis and some Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

The Extension of P-Hacking (Athey 2018)

What We Say v. What We Do (Econometrics)

What We Say

- Causal inference and counterfactuals
- God gave us the model
- We report estimated causal effects and appropriate standard errors
- Plus a few additional specifications for robustness

What we do

- Run OLS or IV regressions
 - Try a lot of functional forms
 - Report standard errors as if we ran only one model
 - Have research assistants run hundreds of regressions and pick a few “representative” ones
- Use complex structural models
 - Make a lot of assumptions without a great way to test them



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility Crisis and some Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

The Practice of P-Hacking (BuzzFeed 2018!)

From: [Brian Wansink](#)
To: [David Just](#)
Cc: [Collin Payne](#); [Sandra Cuellar](#)
Subject: Can Branding Improve School Lunches?
Date: Saturday, January 7, 2012 7:17:42 AM
Attachments: [Elmo Icon-AJPH - 1-7-12.doc](#)
[ATT00001.htm](#)

Hi David,

Here's the Elmo study we are going to spin off and submit.
I think we start with the AJPH as a Brief (80 word abstract and 800 word paper),
and go from there. I'll give Sandra a list of the journals and the priority order we
should consider. Let's consider these two first:

Brief -- American Journal of Public Health

Research Letter – Archives of Pediatric and Adolescent Medicine

One sticking point is that although the stickers increase apple selection by 71%, for
some reason this is a p value of .06. It seems to me it should be lower. Do you
want to take a look at it and see what you think. If you can get the data, and it
needs some tweaking, it would be good to get that one value below .05.

Best,

Brian



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

P-Hacking DIY

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

- “Science isn’t Broken” —538 journalism piece with interactive demo [Link](#)
- Train your p-hacking skills R/Shiny App. [Link](#)
- An Exact Fishy Test [Link](#)



Pre-Analysis Plan

Explain the solution:

- From 3ie: “A pre-analysis plan is a **detailed description of the analysis** to be conducted that is written **in advance of seeing the data** on impacts of the program being evaluated. It may specify **hypotheses to be tested, variable construction, equations to be estimated, controls to be used**, and other aspects of the analysis. A key function of the pre-analysis plan is to increase transparency in the research. By setting out the details in advance of what will be done and before knowing the results, the plan guards against data mining and specification searching. Researchers are encouraged to develop and upload such a plan with their study registration, but it is not required for registration.”

Suggested Items (M)

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

World Bank Development Impact Blog

- 1 Description of the sample to be used in the study
- 2 Key data sources
- 3 Hypotheses to be tested throughout the causal chain
- 4 Specify how variables will be constructed
- 5 Specify the treatment effect equation to be estimated
- 6 What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
- 7 Procedures to be used for addressing survey attrition
- 8 How will the study deal with outcomes with limited variation?
- 9 If you are going to be testing a model, include the model
- 10 Remember to archive it

Examples

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

- J-PAL Hypothesis Registry (11), see <http://www.povertyactionlab.org/Hypothesis-Registry>
6 published papers:
 - Sierra Leone CDD, Oregon Medicare, Turkey Job Training, El Salvador TOMS, two in Indonesia (Olken et al.)
 - Psychology: Hawkins, Fitzgerald, Nosek—Conception Risk and Prejudice

Wide range of when exactly to write and how detailed to make the plan. At the extreme level of detail you would have your entire code already written before you got any data.

Why Do We Need PAPs? The Social Planner View (Haushofer, 2017)

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Benefits:

1. Clear ex-ante what the researcher planned.
2. Reduces false positives.
3. Reduces the file drawer problem.

Costs:

1. Time cost. Not that relevant.
2. Stifles exploratory work. Same, see below.
3. Pre-specifying the wrong analyses (ex ante or ex post).
This is potentially serious.

Reducing costs:

1. Time cost: make the PAP your methods section later.
2. Exploratory work: data mine all you want. Just label it properly.
3. Pre-specifying the wrong analyses: Be honest about your thought process and hope for sensible readers/referees.

Why Do We Need PAPs? The Social Planner View (Haushofer, 2017)

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Benefits:

1. Clear ex-ante what the researcher planned.
2. Reduces false positives.
3. Reduces the file drawer problem.

Costs:

1. Time cost. Not that relevant.
2. Stifles exploratory work. Same, see below.
3. Pre-specifying the wrong analyses (ex ante or ex post).
This is potentially serious.

Reducing costs:

1. Time cost: make the PAP your methods section later.
2. Exploratory work: data mine all you want. Just label it properly.
3. Pre-specifying the wrong analyses: Be honest about your thought process and hope for sensible readers/referees.

Why Do We Need PAPs? The Social Planner View (Haushofer, 2017)

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Benefits:

1. Clear ex-ante what the researcher planned.
2. Reduces false positives.
3. Reduces the file drawer problem.

Costs:

1. Time cost. Not that relevant.
2. Stifles exploratory work. Same, see below.
3. Pre-specifying the wrong analyses (ex ante or ex post).
This is potentially serious.

Reducing costs:

1. Time cost: make the PAP your methods section later.
2. Exploratory work: data mine all you want. Just label it properly.
3. Pre-specifying the wrong analyses: Be honest about your thought process and hope for sensible readers/referees.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Why write one? Individual View

Benefits:

- 1 Strengthens the quality of your research.
- 2 Signals dedication to honesty and rigor.
- 3 Now you can get in-principle acceptance to a good PAP (JDE).



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility Crisis and some Solutions

Introduction

Ethical Research

Publication Bias — Registrations

P-Hacking — Pre-Analysis Plans

Replications — Guidelines & Tools

Conclusion

Register Reports at the JDE.

[Home](#) > [Journals](#) > [Journal of Development Economics](#)



ISSN: 0304-3878

[Submit Your Paper](#)

[View Articles](#)

[Guide for Authors](#)

[Abstracting/ Indexing](#)

[Track Your Paper](#)

[Order Journal](#)

[Journal Metrics](#)

CiteScore: 2.96

More about CiteScore

Impact Factor: 1.889

5-Year Impact Factor: 3.305

Source Normalized Impact per Paper
(SNIP): 2.732

Journal of Development Economics

Supports Open Access

Editor in Chief: A. Foster

[View Editorial Board](#)

The *Journal of Development Economics* publishes original research papers relating to all aspects of economic development - from immediate policy concerns to structural problems of underdevelopment. The emphasis is on quantitative or analytical work, which is novel and relevant. The Journal does not publish...

[Read more](#)

[Most Downloaded](#) [Recent Articles](#) [Most Cited](#) [Open Access Articles](#)

Fertilizing growth: Agricultural inputs and their effects in economic development John W. McArthur | Gordon C. McCord

Do refugee camps help or hurt hosts? The case of Kakuma, Kenya Jennifer Alix-Garcia | Sarah Walker | ...

Informal finance: A theory of moneylenders Andreas Madestam

[View All Most Downloaded Articles](#)

Announcements

NEW: Registered Reports

As part of a pilot project, the JDE now offers authors the opportunity to have their prospective empirical projects reviewed and approved for publication before the results are known (referred to as 'Registered Reports'). This pre-results review track may be particularly suitable for authors working on research projects for which they have not

Guidelines and Checklist



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

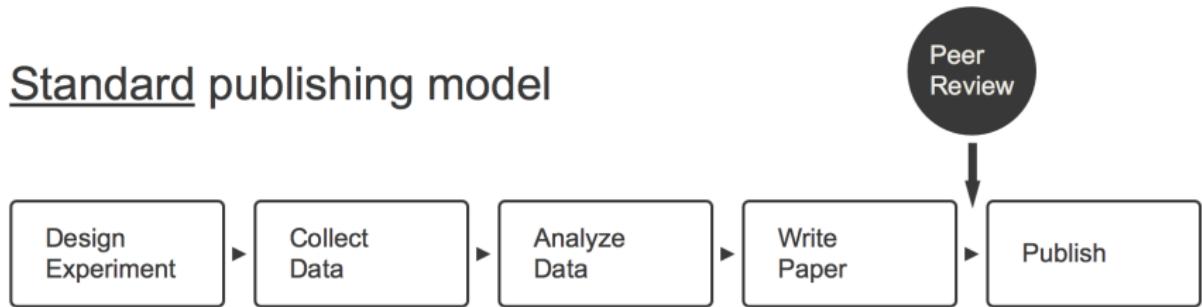
Design-Based Publication

AKA Registered Reports, moves peer review before data gathering, results, and analysis.

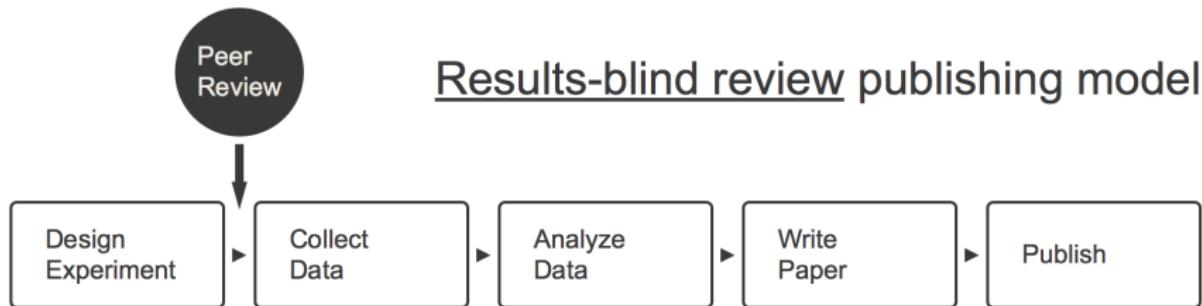
- 1 Design a project
- 2 Submit
- 3 Reviewed based on importance of question and quality of design
- 4 Get in-principle acceptance
- 5 Follow through, and nulls get published

20 Journals, 5 more with Special Issues [▶ Link](#)

Standard publishing model



Results-blind review publishing model





THE PREREGISTRATION CHALLENGE

Learn more at cos.io/prereg



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

PAP—Observational Studies

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

- Debated in public health/epidemiology and economics (Burlig 2018).
- Difficult, but not impossible, to verifiably pre-specify.
- Example: Government data releases
- Example: Minimum Wage (Neumark 2001)
- Example: US, Italian Elections



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Similarities between Registration and PAP?

- Time stamped document, public (or to be published in a specified date), that describes a prospective study.
- Both go in same registries:
 - Medicine: clinicaltrial.gov
 - Social Science (RCTs): socialscienceregistry.org (AEA)
 - Social Science (Observational in dev. countries): RIDIE (3ie)
 - All disciplines and methods: osf.io



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Difference between Registration and PAP?

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

- Key difference is the amount of detail/effort.
- Registration: very easy, goal is to track publication bias.
- PAP: much more detail required. Similar to grant application/work plan.
- It is more a matter of degree.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Final Considerations for Registrations & PAPs

Time dimension:

- Both PAPs and Registrations should be submitted to a public registry *before* looking at the *entire dataset*.
- A broadly defined registration should not change much so the earlier the better.
- A precise PAP need as much information as possible. Ok to look at data, as long as can prove lack of access either treatment or outcome variable.

Deviations:

- It is completely fine to deviate from the original PAP. Just label it properly.
- Really good example of how to handle deviations: Green's SOPs (Article, SOP)



Outline

- 1 Introduction
- 2 Ethical Research
- 3 Publication Bias — Registrations
- 4 P-Hacking — Pre-Analysis Plans
- 5 Replications — Guidelines & Tools
- 6 Conclusion

Replication

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

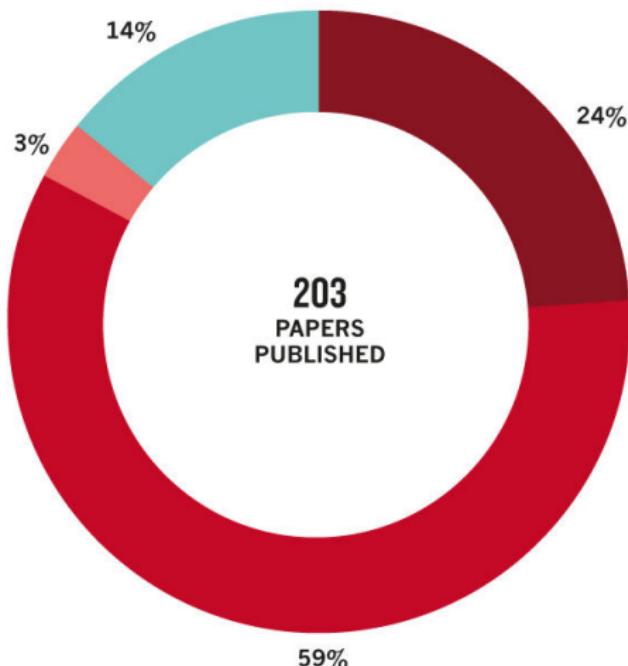
- 1 Lack of replicability in code:**
 - 1986: JMCB Project
 - 2018: Gertler, Galiani, Romero
- 2 Also a lack of replicability in experiments: Camerer et al 2016, Many Labs, Reproducibility Project: Psychology**
- 3 TOP Guidelines**
- 4 Organizing Workflow**
- 5 Code & Data Sharing**
- 6 Reporting Guidelines**

REPLICATION RARELY POSSIBLE

An analysis of 203 economics papers found that fewer than one in seven supplied the materials needed for replication.

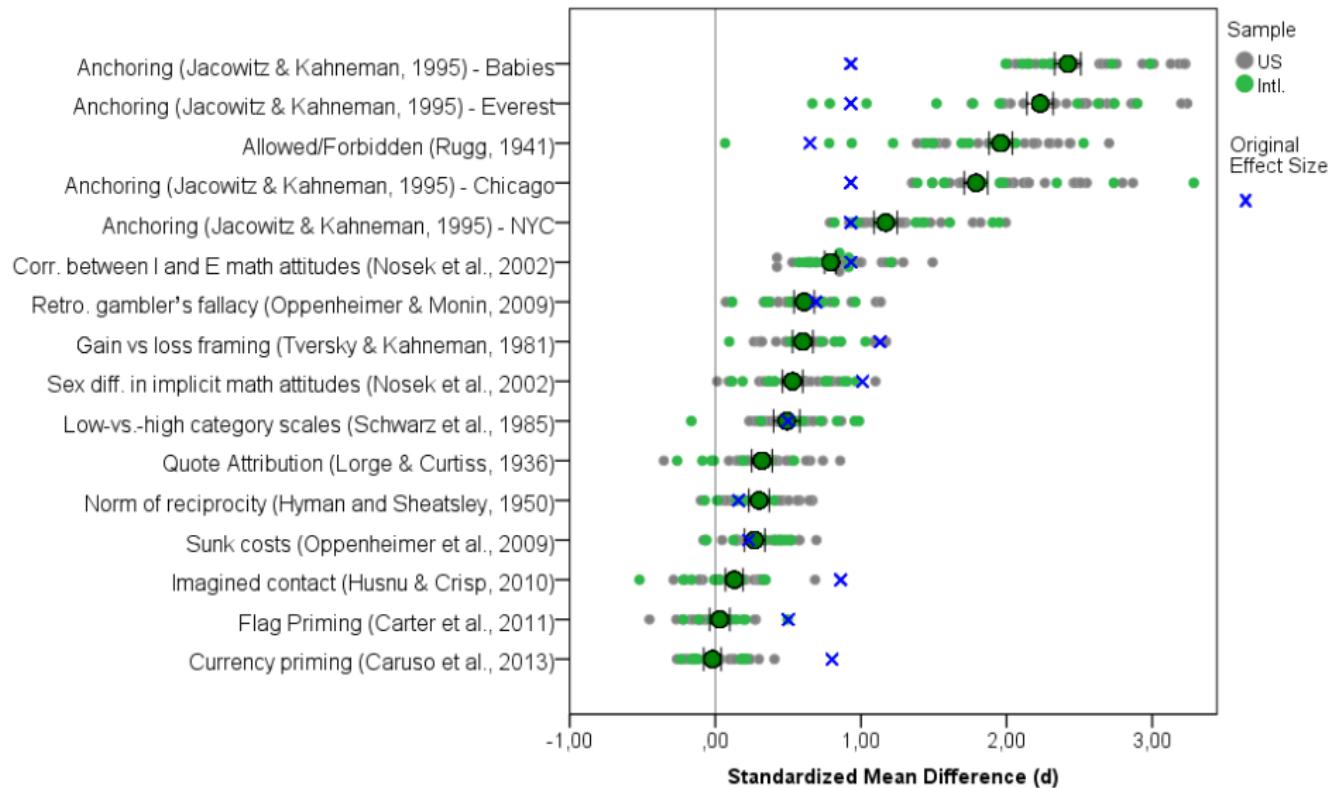
ELEMENTS PROVIDED*:

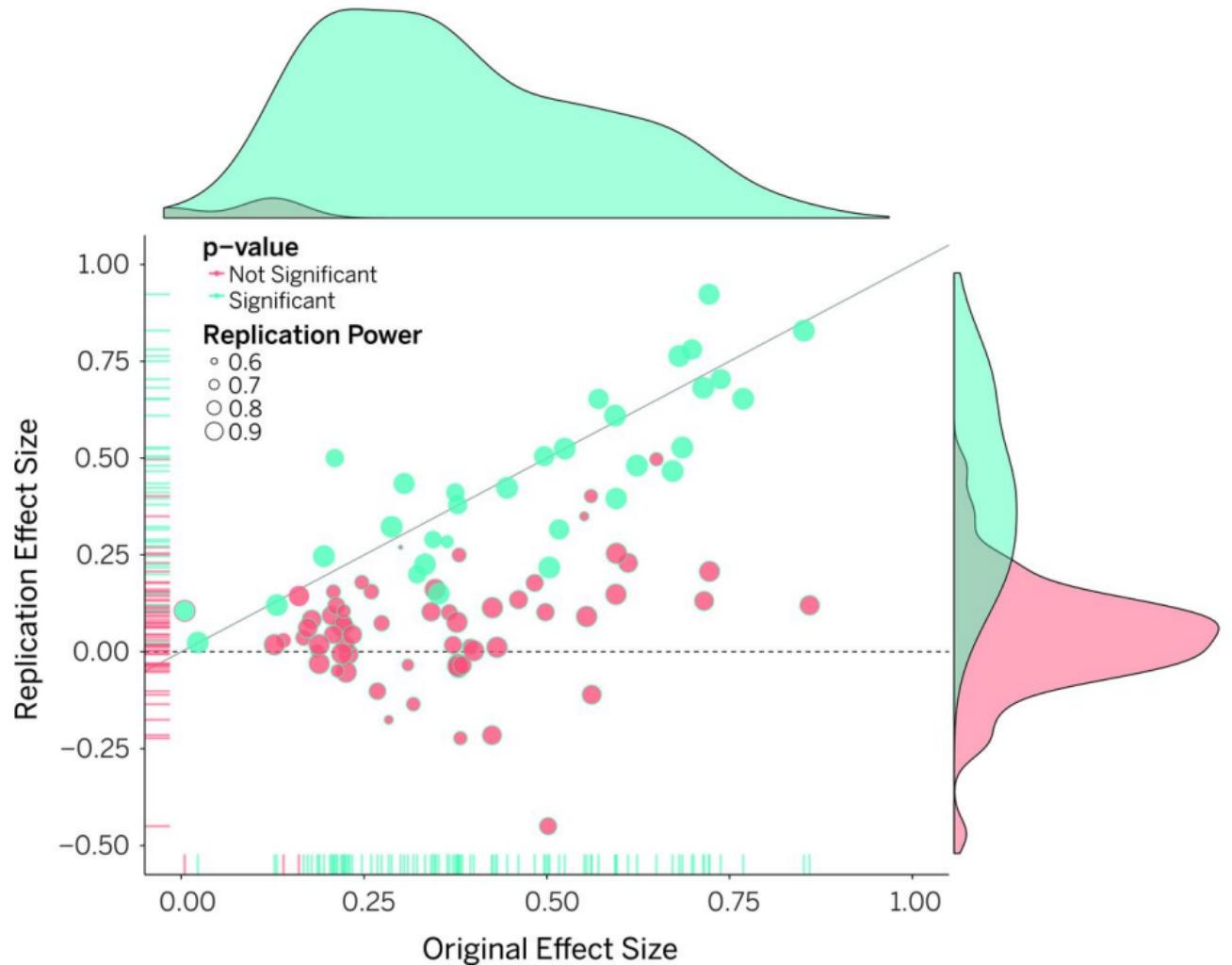
- None ■ One or more missing
- All, code doesn't run ■ All, code runs



*The elements assessed were raw data, raw code, estimation data and estimation code.

©nature







BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Project Protocol, Reporting Standards

Make sure you report everything another researcher would need to replicate your research.

- Find the appropriate reporting standard for your field and follow it: <http://www.equator-network.org>
- Report the nuts and bolts of the project implementation in a detailed protocol:
<http://www.spirit-statement.org>
- Transparency and Openness Promotion (TOP) Guidelines: <http://cos.io/top>

SHARE[POLICY FORUM](#) | [SCIENTIFIC STANDARDS](#)

Promoting an open research culture



B. A. Nosek*, G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. Levy Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, T. Yarkoni

+ Author Affiliations

*Corresponding author. E-mail: nosek@virginia.edu



Workflow

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

“Reproducibility is just collaboration with people you don’t know, including yourself next week”
—Philip Stark, UC Berkeley Statistics

Workflow

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

- Practical coding and organizational suggestions
 - Making any changes to a file that has been posted/shared means it gets a new name.
 - Use version commands to ensure others get same results.
 - Long (2008) *The Workflow of Data Analysis Using Stata*
- Literate programming (extensive commenting, making the aim of code reading by a human)
- Version Control
- Dynamic Documents

Version Control

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

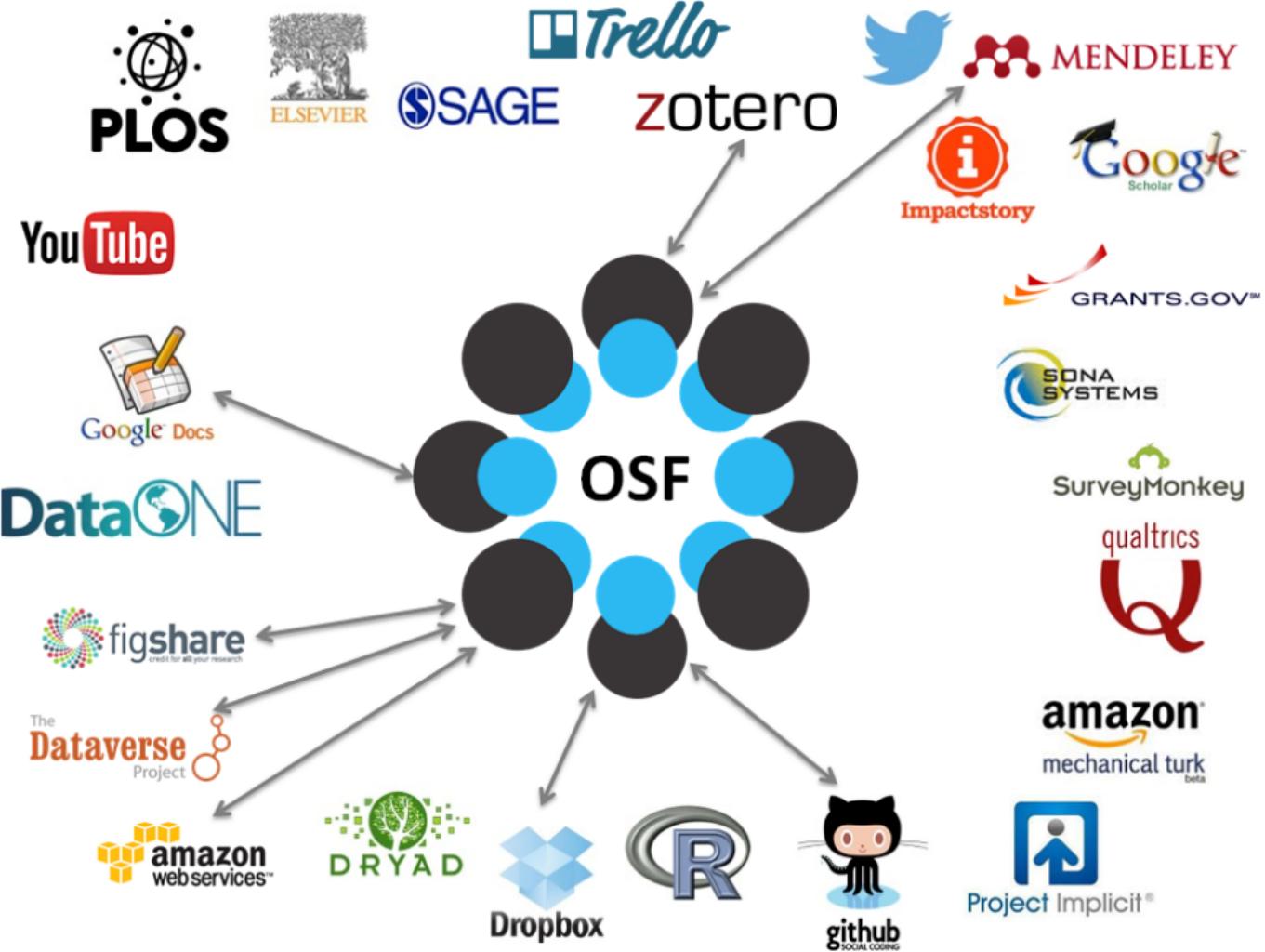
Conclusion

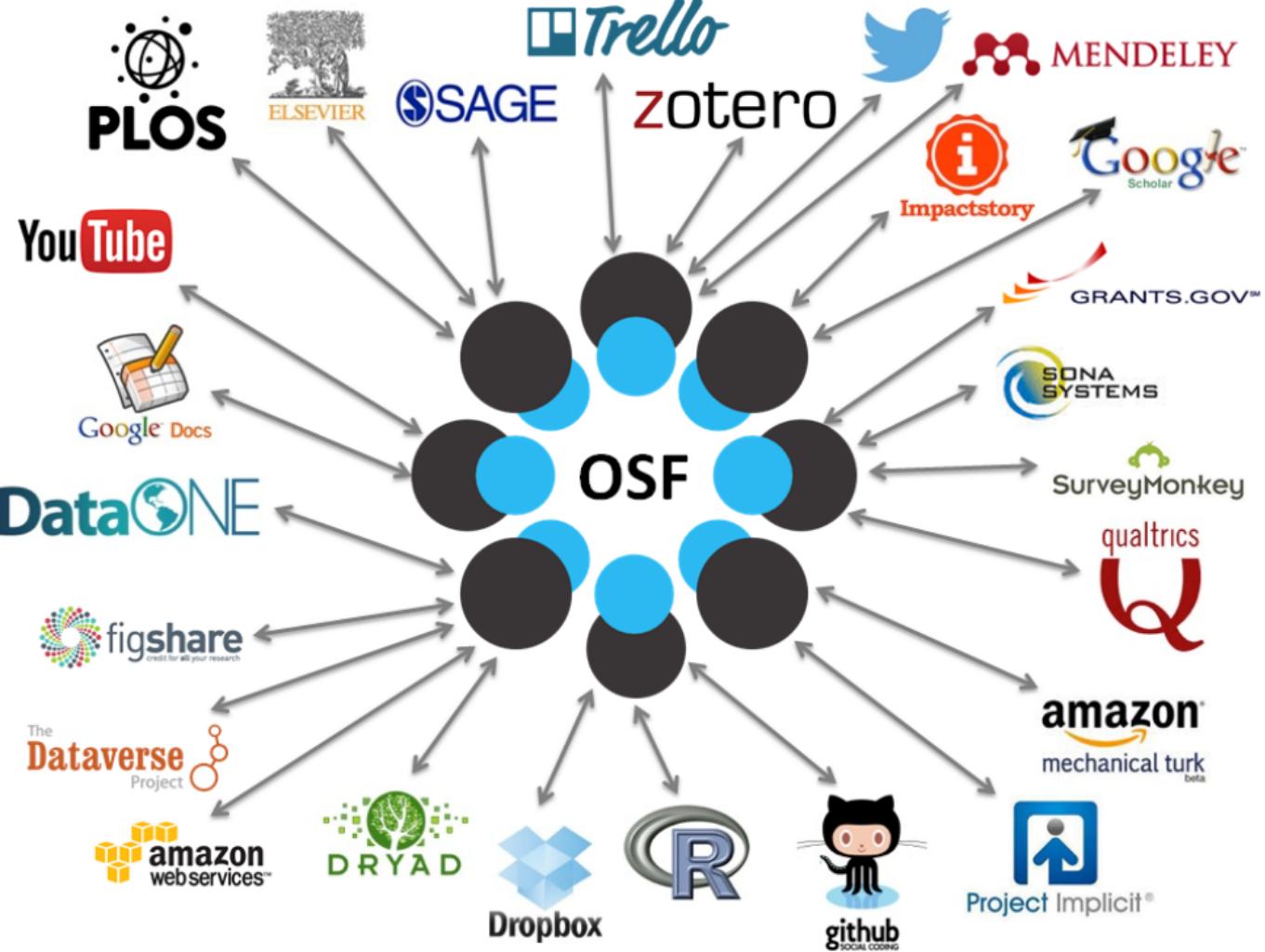
- Using version control (AKA revision control) can help to make your work more reproducible.
- What is version control?

Version control is a system that records changes to a file or set of files over time so that you can recall specific versions later. For the examples in this book you will use software source code as the files being version controlled, though in reality you can do this with nearly any type of file on a computer.

–Git, About Version Control









BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Dynamic Documents

Write your code and your paper in the same file so you won't lose information or make copy and paste mistakes.
Possible in R and Stata.

- Include tables by linking to a file, instead of a static image.
- Include number by linking to a value calculated by an analysis file, instead of a static number typed manually.
- Automatically update tables and numbers.
- Produce entire paper with one or two clicks.

The logo for Jupyter consists of the word "jupyter" in a bold, black, sans-serif font, centered within a white circle. The circle is partially overlaid by two thick, orange, curved bands that intersect at the bottom. Small dark gray circles are positioned at the intersections of these curves.

jupyter





BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Data Sharing

Essential component for research transparency. Deep dive in next session.

- 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.



Outline

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

1 Introduction

2 Ethical Research

3 Publication Bias — Registrations

4 P-Hacking — Pre-Analysis Plans

5 Replications — Guidelines & Tools

6 Conclusion

Conclusion

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

OK, I'm convinced. How do I implement this in my own research?

- Take our MOOC (E. Miguel). [▶ Link](#)
- Apply for our NIH Research Transparency & Reproducibility Training (RT2). [▶ Link](#)
- Subscribe to the BITSS blog & E-mail list [▶ Link](#)
- Read the manual Garret wrote. [▶ Link](#)

Research Transparency & Reproducibility Training (RT2)

Credibility
Crisis and
some
Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Three days of training in April in Amsterdam or September in Los Angeles.





BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Credibility Crisis and some Solutions

Introduction

Ethical
Research

Publication
Bias —
Registrations

P-Hacking —
Pre-Analysis
Plans

Replications
— Guidelines
& Tools

Conclusion

Questions?

Thank you!