

# Manual of Best Practices in Transparent Social Sciences Research

Garret Christensen<sup>1</sup> Courtney Soderberg<sup>2</sup>

<sup>1</sup>Berkeley Initiative for Transparency in the Social Sciences  
UC Berkeley

<sup>2</sup>Center for Open Science

BITSS Annual Meeting, 2014

# Outline

- 1 Motivation
  - Publication Bias
  - P-Hacking
- 2 Solutions
  - Registration
  - Pre-Analysis Plan
- 3 Implementation
- 4 Conclusion

# Publication Bias

- The distribution of published p-values jumps around .05 (Brodeur et al 2013).
- There is a higher fraction of rejected hypothesis tests in the social sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- 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 massive. (Franco, Malhotra, Simonovits 2014—see tomorrow)

# Publication Bias

- The distribution of published p-values jumps around .05 (Brodeur et al 2013).
- There is a higher fraction of rejected hypothesis tests in the social sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- 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 massive. (Franco, Malhotra, Simonovits 2014—see tomorrow)

# Publication Bias

- The distribution of published p-values jumps around .05 (Brodeur et al 2013).
- There is a higher fraction of rejected hypothesis tests in the social sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- 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 massive. (Franco, Malhotra, Simonovits 2014—see tomorrow)

# Publication Bias

- The distribution of published p-values jumps around .05 (Brodeur et al 2013).
- There is a higher fraction of rejected hypothesis tests in the social sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- 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 massive. (Franco, Malhotra, Simonovits 2014—see tomorrow)

# Publication Bias

- The distribution of published p-values jumps around .05 (Brodeur et al 2013).
- There is a higher fraction of rejected hypothesis tests in the social sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- 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 massive. (Franco, Malhotra, Simonovits 2014—see tomorrow)

# 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.



# P-Hacking

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

# P-Hacking

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

# P-Hacking

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



# Definition and Origins

# Contents

# Contents

# AEA Trial Registration



# EGAP Registration

# 3ie Registration

# Open Science Framework

# Conclusion