

Research Transparency in Social Science

Prof. Theodoridis Seminar, UMass

Fernando Hoces de la Guardia, BITSS

March 2021 | [slides](#)

BITSS

The Berkeley Initiative for Transparency in the Social Sciences works to improve the credibility of science by advancing transparency, reproducibility, rigor, and ethics in research.

Core Team

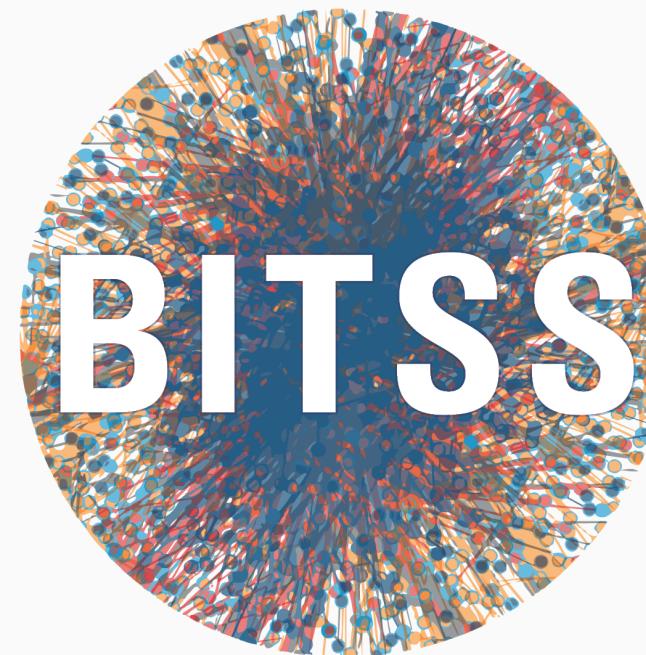
Aleks Bogdanoski

Fernando Hoces de la Guardia

Katie Hoeberling

Edward Miguel

We are part of the Center for Effective Global Action ([CEGA](#)).



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Many Others

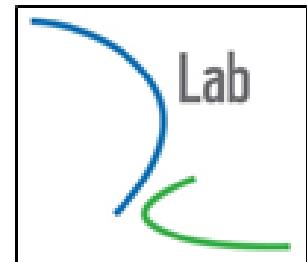
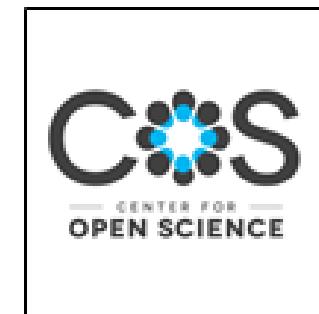
CEGA staff

Undergrad and Graduate RAs

Catalysts

Outside Collaborators
(Researchers, Programmers)

Part of the much larger Open Science Community



ICPSR

Table of Contents

1. **Problems**
2. Solutions
3. Accelerating Computational Reproducibility
4. Open Policy Analysis

Researchers and Scientific Norms (Anderson et. al., 2007)

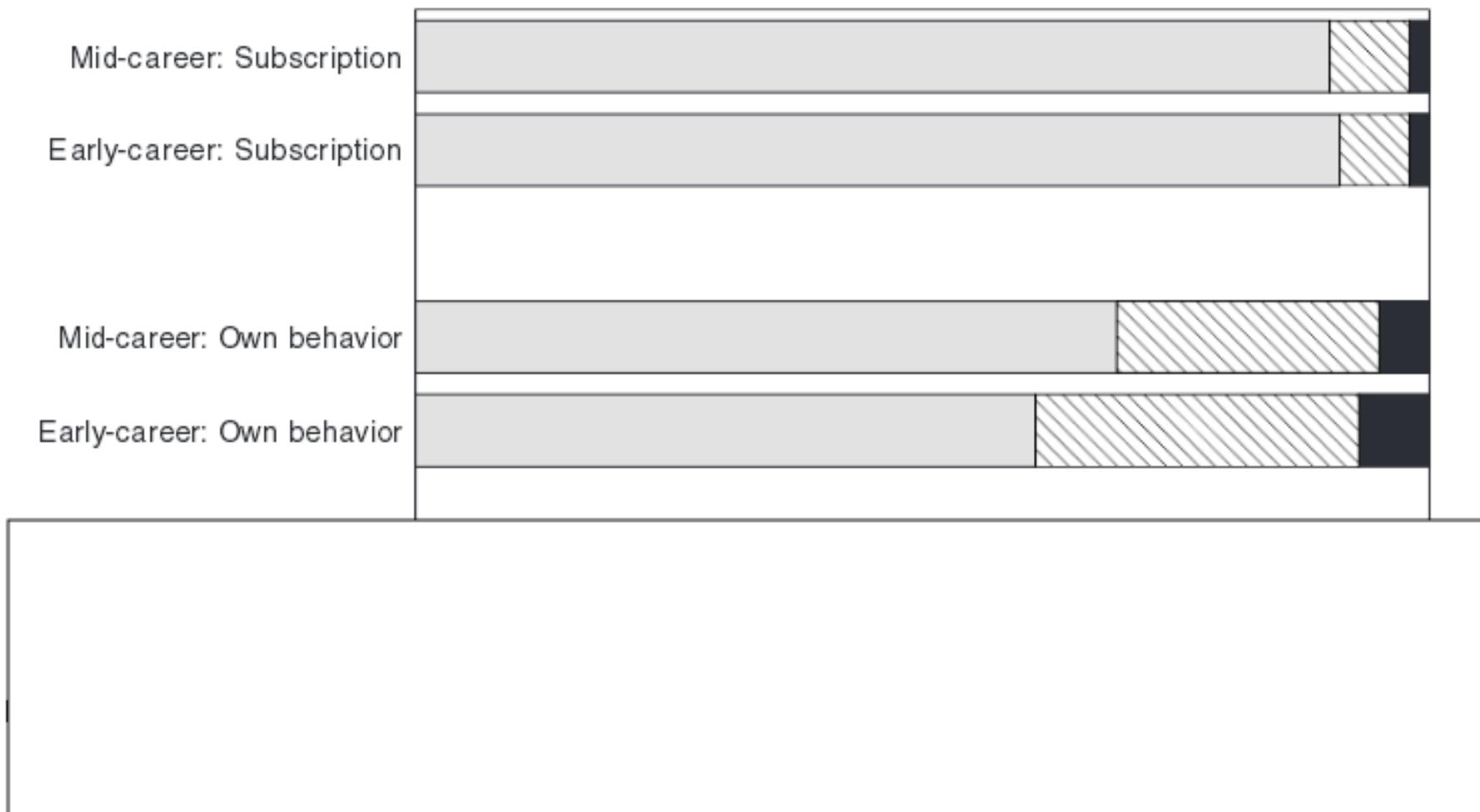


FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

Researchers and Scientific Norms (Anderson et. al., 2007)

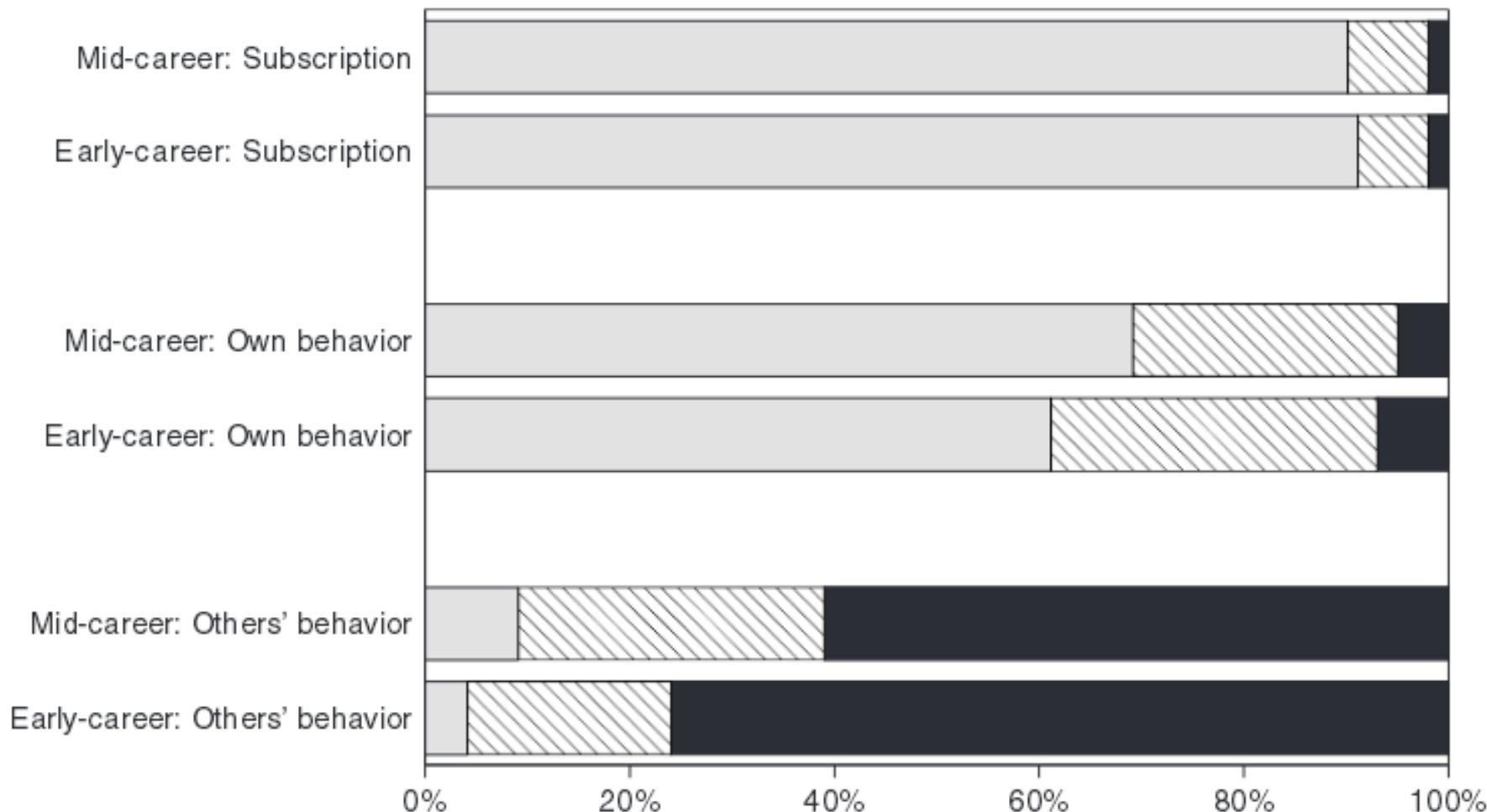
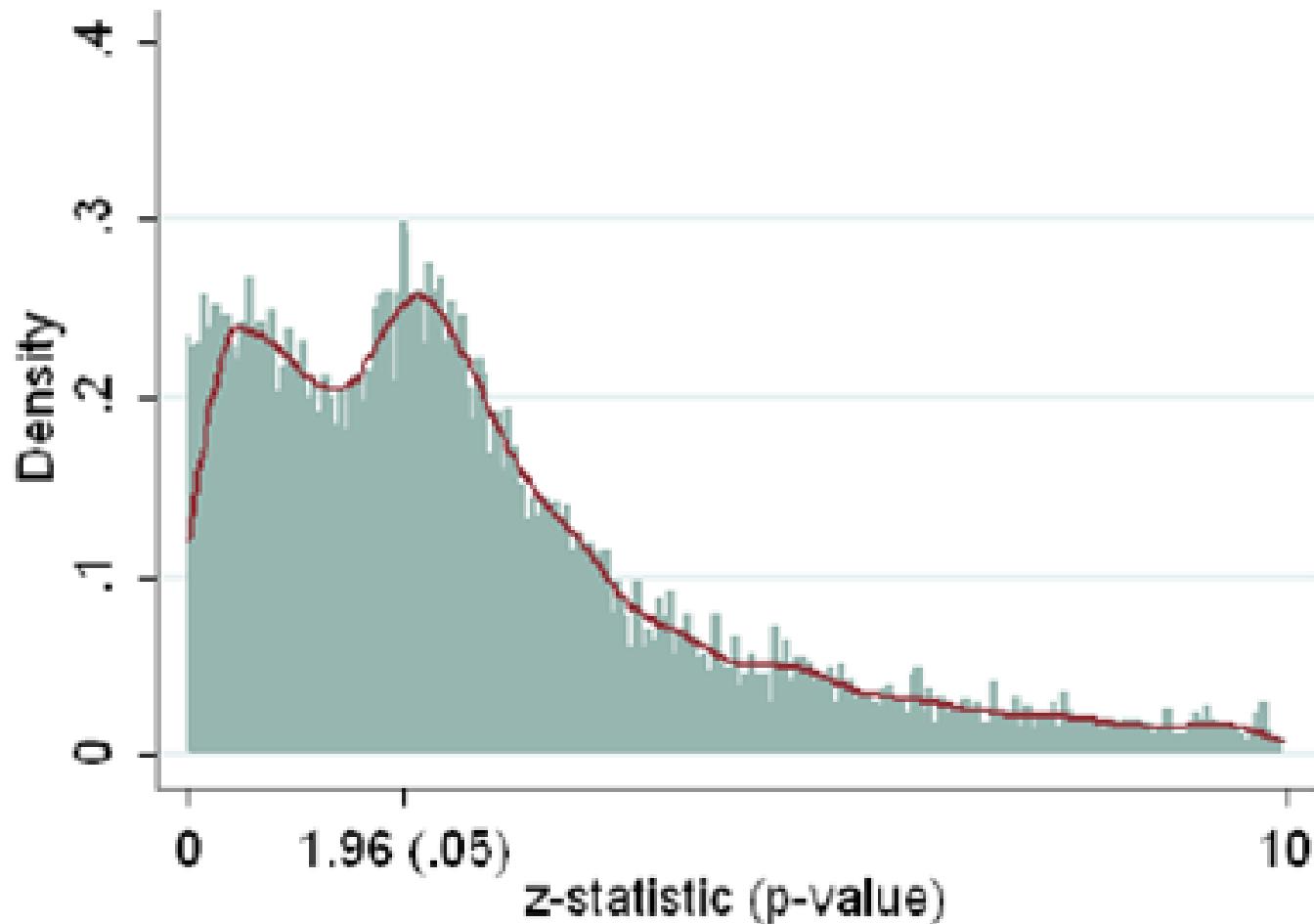


FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

P-hacking (for Economics: Brodeur et. al 2016, 2020)

(b) Unrounded distribution of z-statistics.



Publication Bias (Franco et. al. 2014)

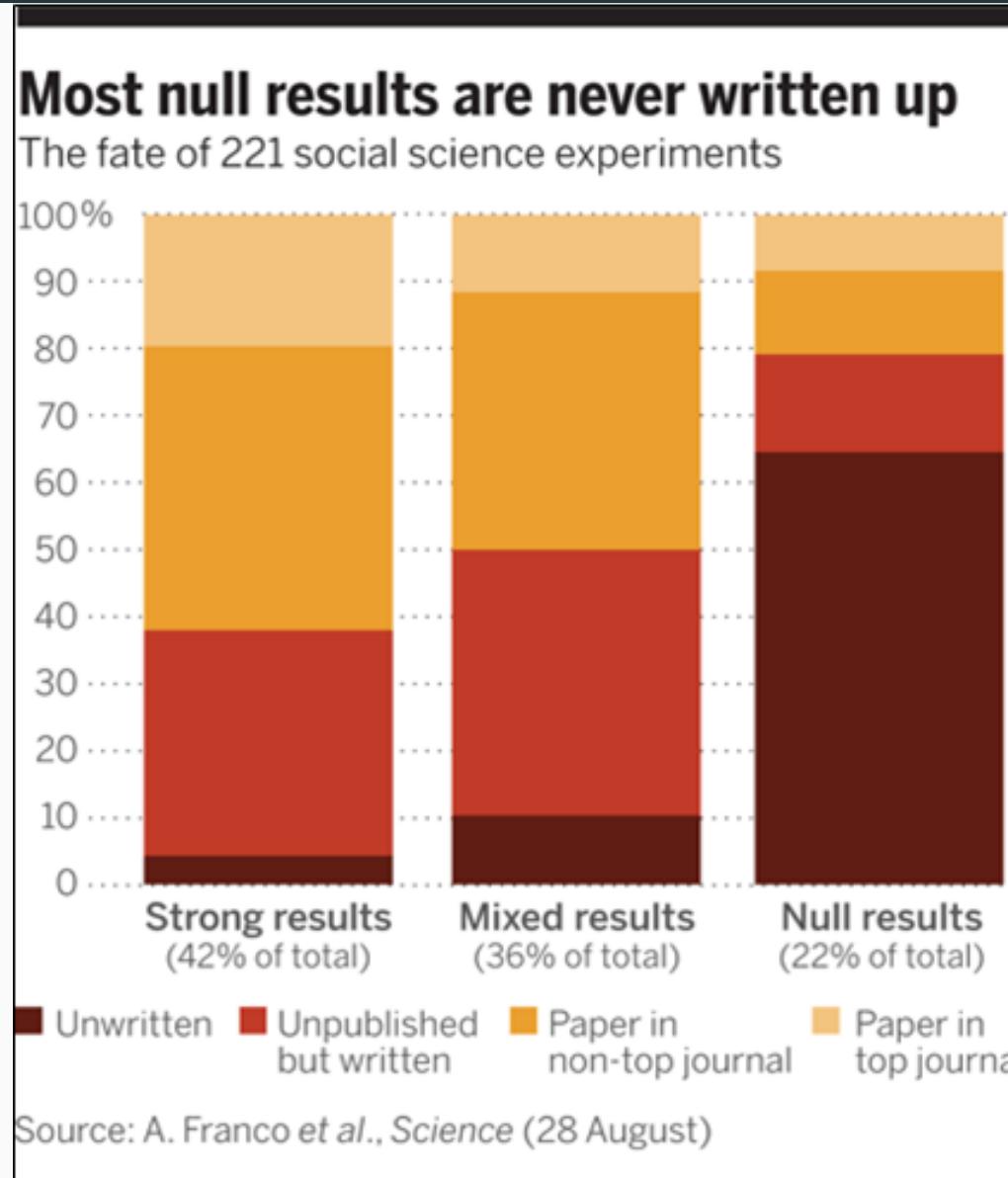


Table of Contents

1. Problems
2. **Solutions**
3. Accelerating Computational Reproducibility
4. Open Policy Analysis

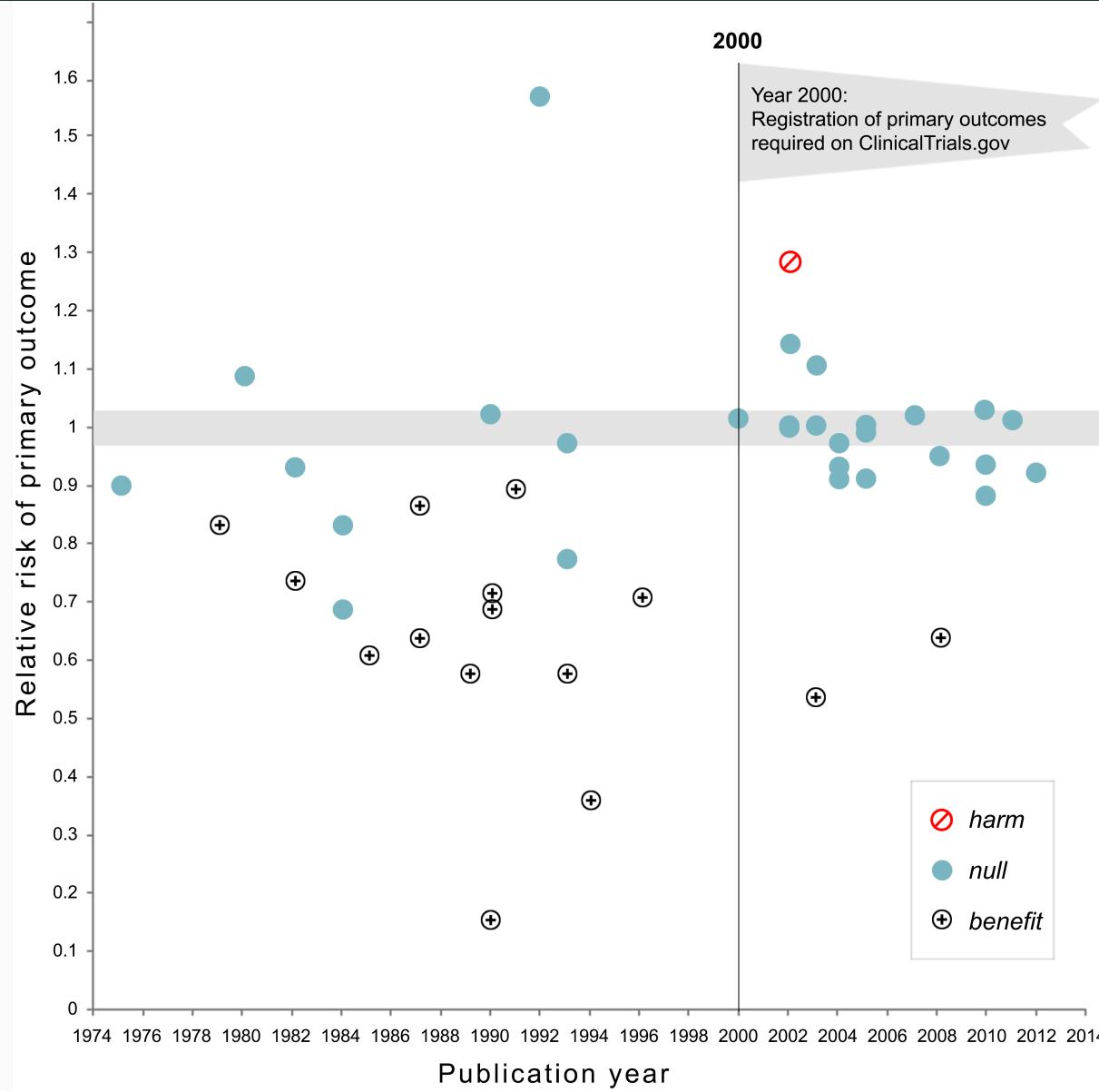
Main Solutions

1. Registrations
2. Pre Analysis Plans
3. Computational Reproducibility

Registrations

- A registration is a record that contains minimal information about a study: title, authors, study country, status, keywords, abstract, start and end dates, outcomes, intervention information, basic research design, whether or not treatments are clustered (when performing an RCT), and Institutional Review Board (IRB) information.
- Preferably, it should be recorded before analyzing data
- **The main goal:** track the complete body of knowledge in a topic of research, regardless of the direction and magnitude of the results.

Why Register: Kaplan and Irvin (2015)



Pre-Analysis Plans

- PAPs are **extensive** methodological descriptions of the analysis to be performed before the endline data is collected
- Helps to prevent p-hacking
- Only way to guarantee accurate statistical testing
- Distinguishes confirmatory from exploratory analysis

How to do a PAP

Christensen, Freese and Miguel (2020) suggest detailed information on:

- | | |
|-------------------------------------|--|
| 1 - Detail study design | 7 - Direction of the effect (optional) |
| 2 - Sampling details | 8 - Exact statistical specification |
| 3 - Outcome Measures | 9 - Structural model (optional) |
| 4 - Index construction | 10 - Timestamp |
| 5 - Multiple hypothesis adjustments | |
| 6 - Subgroup analysis | |

Common Concerns About PAPs

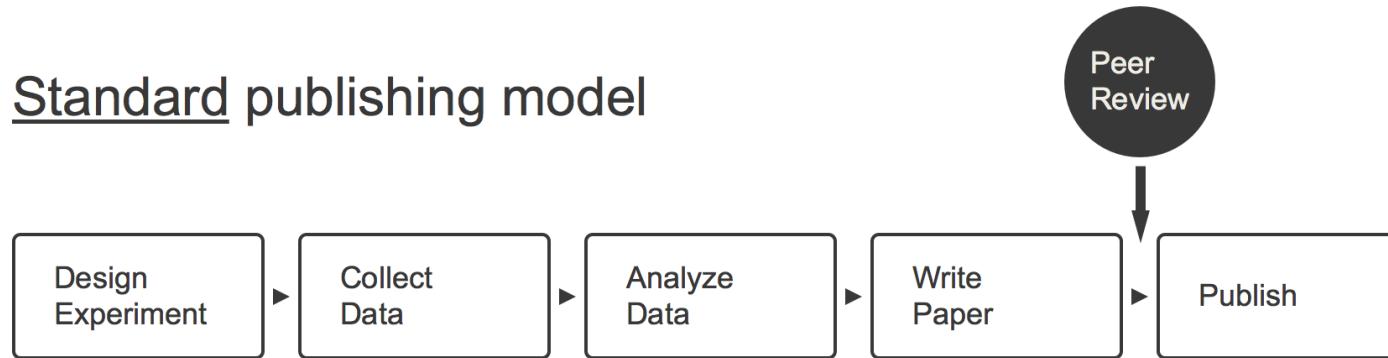
| Critique | Response |
|--|---|
| PAPs take too much time and are too difficult (Olken 2015) | A PAP changes the timing of the analytic component, not clear that it increases it |
| Scientific discovery often comes from surprises. PAPs stifle discovery (Olken 2015) | PAPs do not prevent researchers from doing exploratory work; they only require researchers to be clear about the objectives of their analyses (Ofosu and Posner 2020). |
| If replications are cheap they will rule out false positives, making PAPs irrelevant. (Coffman and Niederle 2015) | Very few experiments are inexpensive as to perform many replications. Moreover, most of the false positives have been identified where experiments are least expensive (lab experiments). |

Why do a PAP: Casey et. al. (2012)

| Outcome variable | (1) | (2) |
|---|-------------------|--------------------|
| | Mean for controls | 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* |

Another Reason: Register Reports

Standard publishing model



Results-blind review publishing model

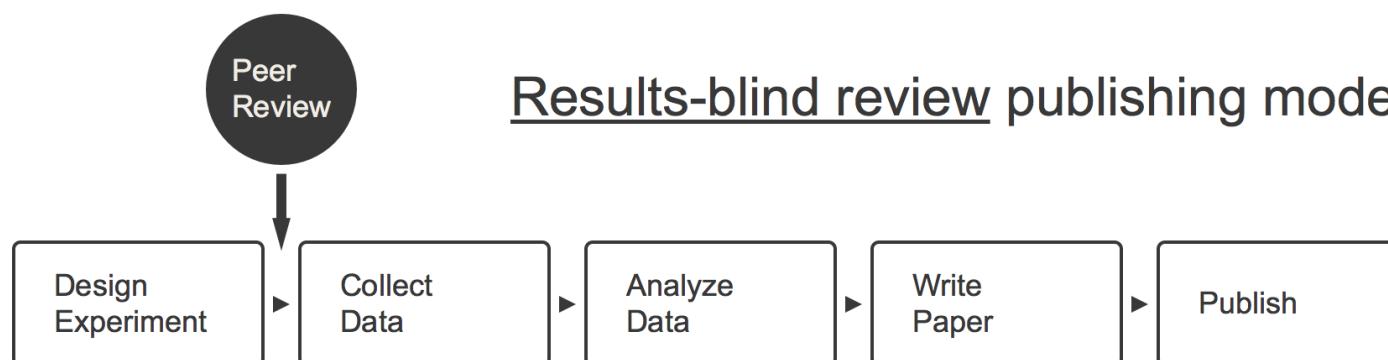


Table of Contents

1. Problems
2. Solutions
3. **Accelerating Computational Reproducibility**
4. Open Policy Analysis

Accelerating Computational Reproducibility (ACRe)

In addition to reproducibility problems, there are two additional motivations for this projects:

Motivation 2: More Inclusive Concept Scholarly Output

Clarebout Principle:

“An article about computational science in a scientific publication is not the scholarship itself, it’s merely scholarship advertisement. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.”

Buckheit and D.L. Donoho (1995, 2009)

Accelerating Computational Reproducibility (ACRe)

In addition to reproducibility problems, there are two additional motivations for this projects:

Motivation 2: More Inclusive Concept Scholarly Output

Clarebout Principle:

“An **article** about computational science in a scientific publication is not the scholarship itself, it’s merely scholarship advertisement. The actual **scholarship is the complete software development environment and the complete set of instructions which generated the figures.**”

Buckheit and D.L. Donoho (1995, 2009)

M2: More Inclusive Concept Scholarly Output

Potential benefits of following the Clarebout Principle

Well discussed potential positive effects on:

- Pedagogy
- Incremental generation of knowledge

Under discussed:

- Possible positive effect on diversity, equity and inclusion: no connections or language skills ("appropriate politeness") required to obtain materials

M3: Prevent Loss of Knowledge

Every semester, graduate students around **the world** take an Empirical/Applied [...] Economics course. A typical assignment consists of reproducing the results of a paper and, possibly, testing the robustness of its results.

| Stage | New Knowledge |
|------------------------------|--|
| Scope (select and verify) | Data and code exist? |
| Assess | Degree of reproducibility for specific part of the paper |
| Improve | E.g. fixed paths, libraries, added missing files, etc. |
| Test robustness | Results are robust to additional specifications |

M3: Prevent Loss of Knowledge

- In each of the previous steps there was valuable new knowledge
- This knowledge is not usually disseminated
 - End of semester report, or presentation that is not published.
 - Language across courses is not standard (reproduction? replication?)
 - Current attempts to record this are closer to a full paper format (see [Replication Wiki](#))
- As a result there is a large amount of missed knowledge.
 - Each new semester, new students repeat similar excesses and have to reinvent the wheel.
 - Lost opportunity for building on top of previous exercises.
 - Lost opportunity for aggregating the new knowledge.

Beyond Binary Judgments

Reproductions can easily gravitate towards adversarial exchanges.

- Early career researcher (ECR) have incentives to emphasize unsuccessful reproductions
- Original authors have a more senior position and can use it to deter in-depth reproductions from ECRs.
- The media also focuses on eye-catching headlines

Our approach:

We do not want to say

"Paper X is (ir)reproducible"

We do want to say

"Result Y in paper X has a high/low **level** of reproducibility according to **several** reproduction attempts. Moreover, **improvements** have been made to the original reproduction package, **increasing** its reproducibility to a higher level"

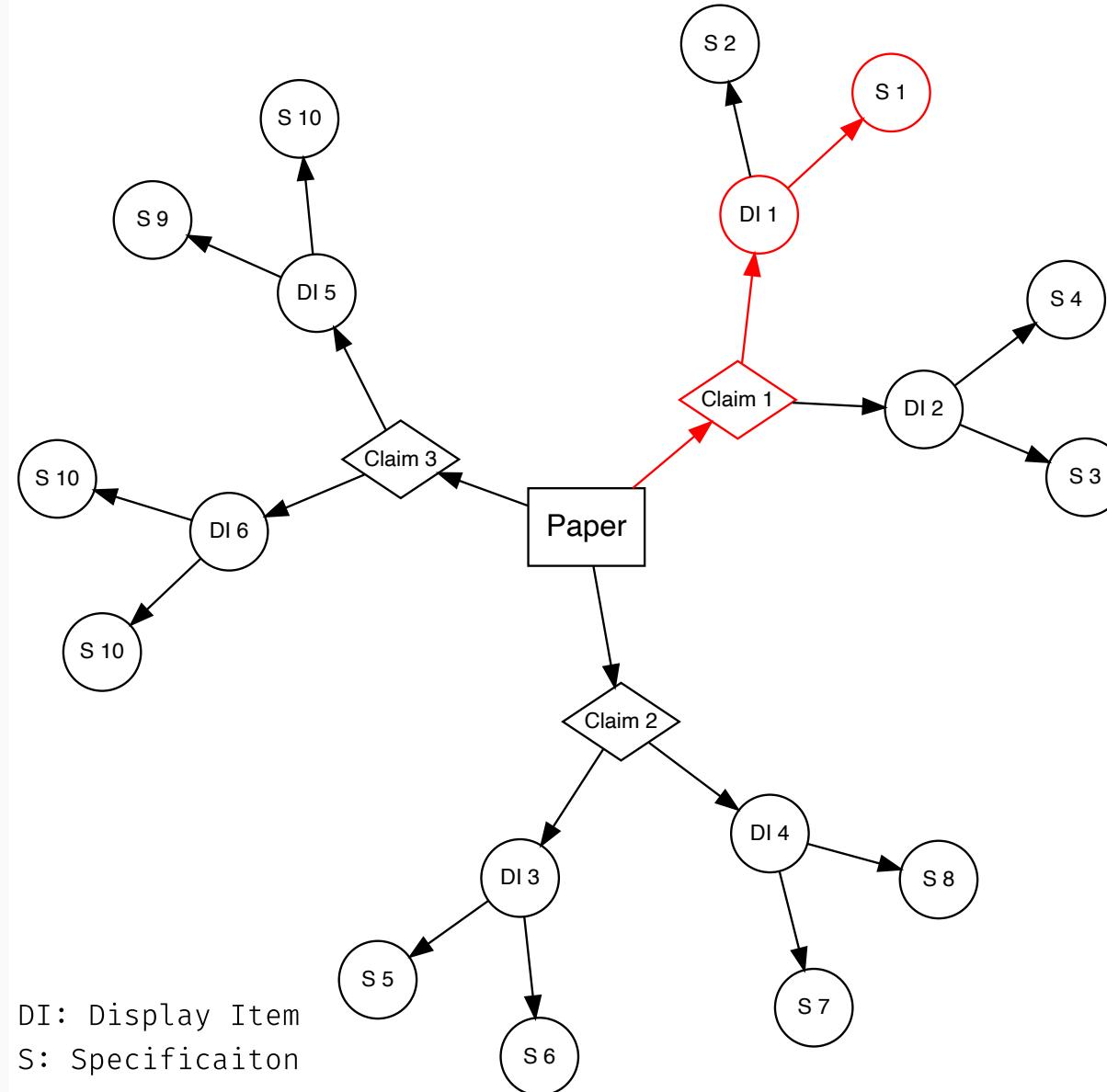
Our Framework

Each **reproduction attempt** is centered around scientific **claims**

One paper can contain several claims.

Each claim may be supported by various **display items**: tables, figures & inline results.

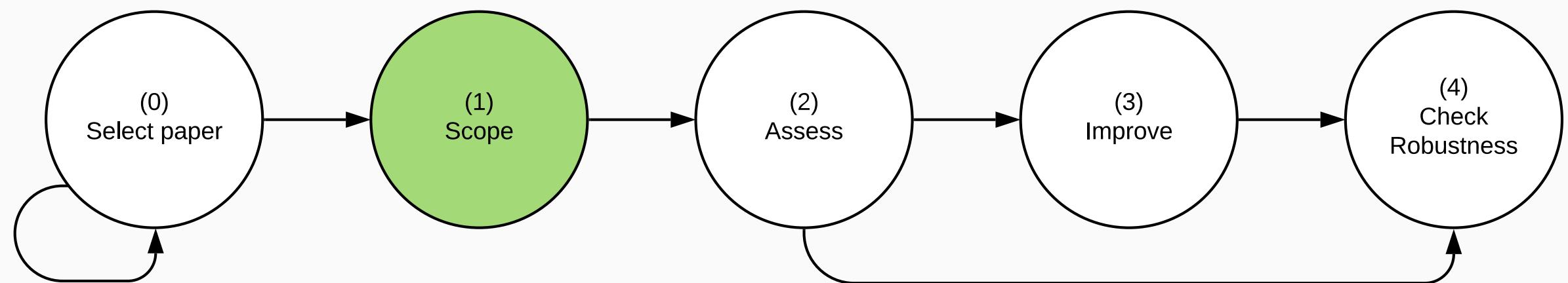
A reproduction attempt is at the claim level, and reproducers must record their **specifications** of interest.



Large part of this exercise is about standardization

- **Computational Reproduction (or Reproduction)**
- **Replication (will not mention this term again!)**
- Reproduction attempt
- **Reproduction package**
- Claim
- Display item
- Specification
- Preferred specification
- **Raw data**
- **Analysis data**
- **Candidate paper**
- **Declared paper**
- Reproduction tree
- Complete Workflow
- Computationally Reproducible from Analytic data (CRA)
- Computationally Reproducible from Raw data (CRR)
- Reasonable test
- Feasible test
- Minimal effort

Stages



Scoping

1. Select or be assigned a candidate paper
2. Check ACRE Platform for previous entries and verify availability of reproduction package (RP)
3. If not RP, leave a short record, and repeat with a different candidate paper
4. Once RP is found then candidate becomes declared paper
5. Only then: read the paper and select claim(s), display items and specification to reproduce

Box 1: Summary Report Card for ACRE Paper Entry

Title: Sample Title

Authors: Jane Doe & John Doe

Original Reproduction Package Available: URL/No
[If "No"]

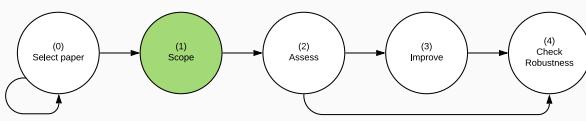
Contacted Authors?: Yes/No
[If "Yes(contacted)"]

Type of Response: Categories (6).

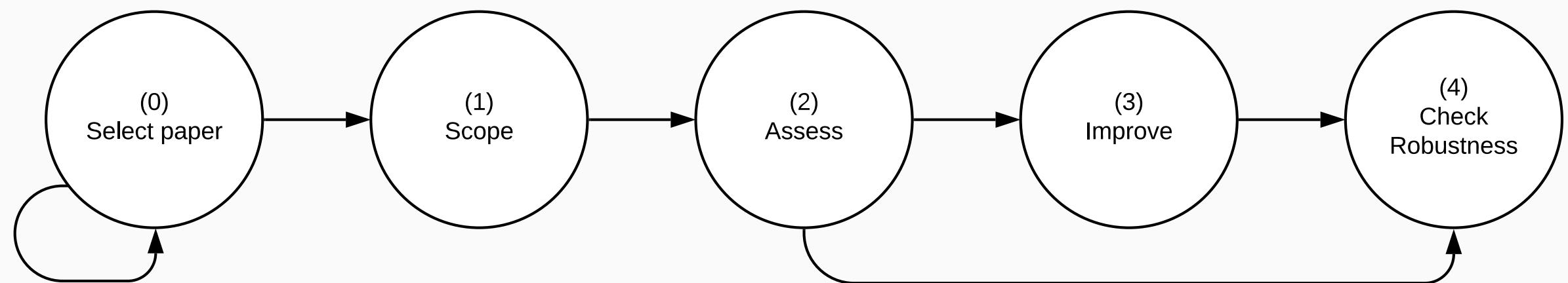
Additional Reproduction Packages: Number (eg, 2)

Authors Available for Further Questions for ACRE

Reproductions: Yes/No/Unknown



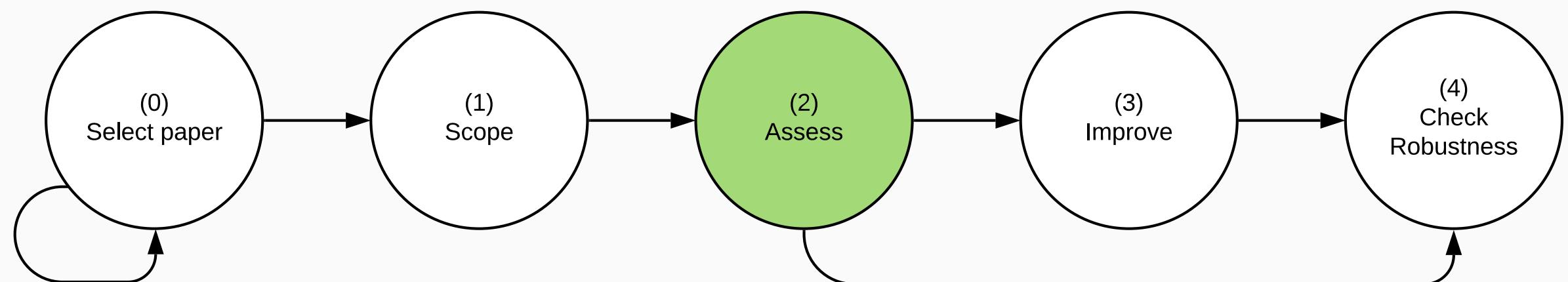
Stages



Assessment

Two main parts for assessment:

1. Find all the elements behind a display item
2. Score the reproducibility of that display item



Identify All the Elements Behind a Display Item

Reproducers will be asked to draw a clear connection to the raw data sources mentioned in the paper and the display item under reproduction.

Data sources

Connect the data sources in the paper's text with specific raw data files.

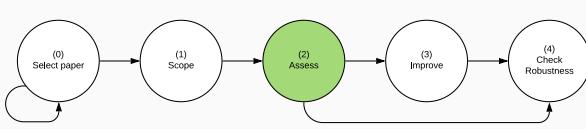
Analytic data sets

Describe each analytic data file.

Code files

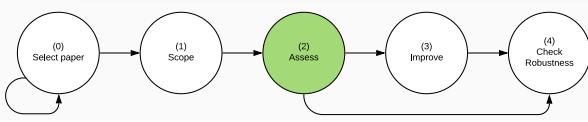
Inspect all code files and record all their inputs and outputs.

With all the information recorded above, reproducers can use the **ACRE Diagram Builder** to generate a **reproduction tree**.



ReproducibiliTREE

```
table1.tex
|___[code] analysis.R
|___analysis_data.dta
|___[code] final_merge.do
|___cleaned_1_2.dta
|   |___[code] clean_merged_1_2.do
|   |___merged_1_2.dta
|   |___[code] merge_1_2.do
|   |___cleaned_1.dta
|   |   |___[code] clean_raw_1.py
|   |   |___raw_1.dta
|   |___cleaned_2.dta
|   |___[code] clean_raw_2.py
|   |___raw_2.dta
|___cleaned_3_4.dta
|___[code] clean_merged_3_4.do
|___merged_3_4.dta
|___[code] merge_3_4.do
|___cleaned_3.dta
|   |___[code] clean_raw_3.py
|   |___raw_3.dta
|___cleaned_4.dta
|___[code] clean_raw_4.py
|___raw_4.dta
```

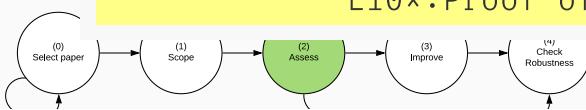


Levels: Proprietary/Confidential Data

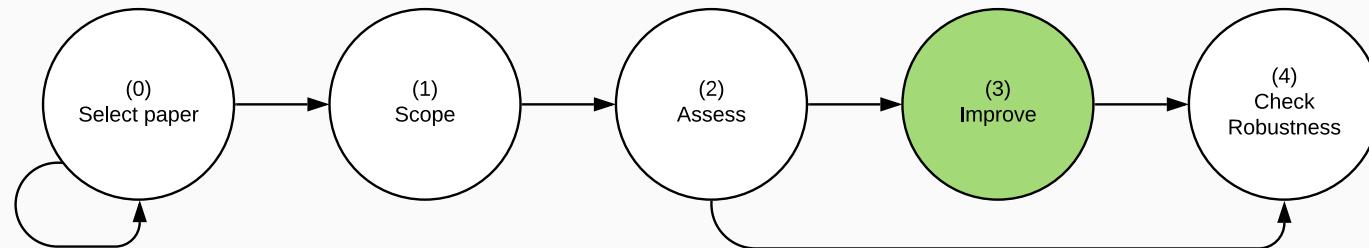
Levels of Computational Reproducibility
with Proprietary/Confidential Data

(P denotes "partial", C denotes "complete")

| | Availability of materials, and reproducibility | | | | | | | | | | | |
|--|--|------|--------|------|----------|-----|----------|---|----------|---|-----|---|
| | Instr. | | Instr. | | Analysis | | Analysis | | Cleaning | | Raw | |
| | Code | Data | CRA | Code | Data | CRR | P | C | P | C | P | C |
| L1: No materials..... | - | - | - | - | - | - | - | - | - | - | - | - |
| L2: Only code | ✓ | ✓ | - | - | - | - | - | - | - | - | - | - |
| L3*: Partial analysis data & code..... | ✓ | ✓ | ✓ | - | - | - | - | - | - | - | - | - |
| L4*: All analysis data & code..... | ✓ | ✓ | ✓ | ✓ | - | - | - | - | - | - | - | - |
| L5*: Proof of third party CRA..... | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - | - | - | - | - | - |
| L6: Some cleaning code..... | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - | - | - | - | - |
| L7: All cleaning code..... | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - | - | - | - |
| L8*: Some instr. for raw data..... | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - | - |
| L9*: All instr. for raw data..... | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | - |
| L10*:Proof of third party CRR..... | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |



Improvements

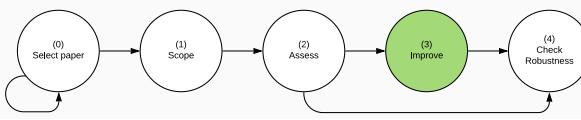


Three types of improvements:

1. Improvements at the paper level
2. Improvements at the display-item level
3. Specific future improvements

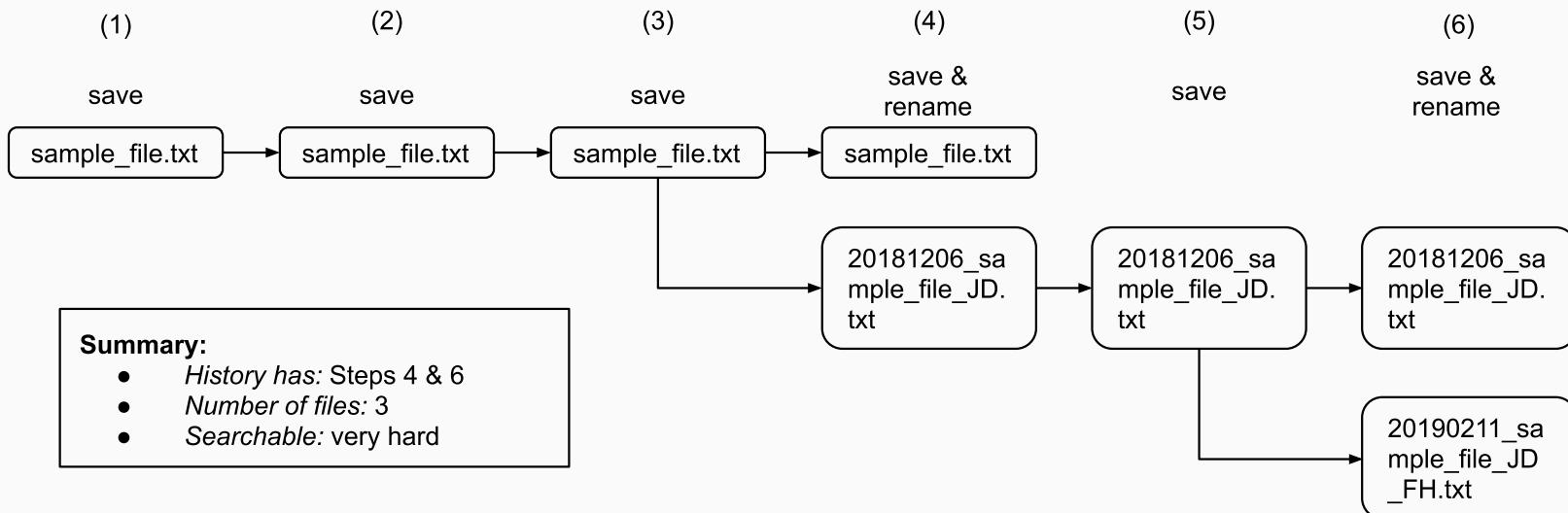
Improvements: Paper-level

- Use version control software (Git/Github).
- Improve documentation: comments, indentations, object names, etc.
- Re-organize the reproduction package into a set of folders and sub-folders that follow **standardized best practices**, and add a master script that executes all the code in order, with no further modifications. [See AEA's reproduction template](#).
- Literate programming environment (e.g., Jupyter notebooks, RMarkdown)
- Re-write code using a different statistical software (ideally open source, like R, Python, or Julia).
- Set up a computing capsule (e.g., [Binder](#) and [Code Ocean](#)).

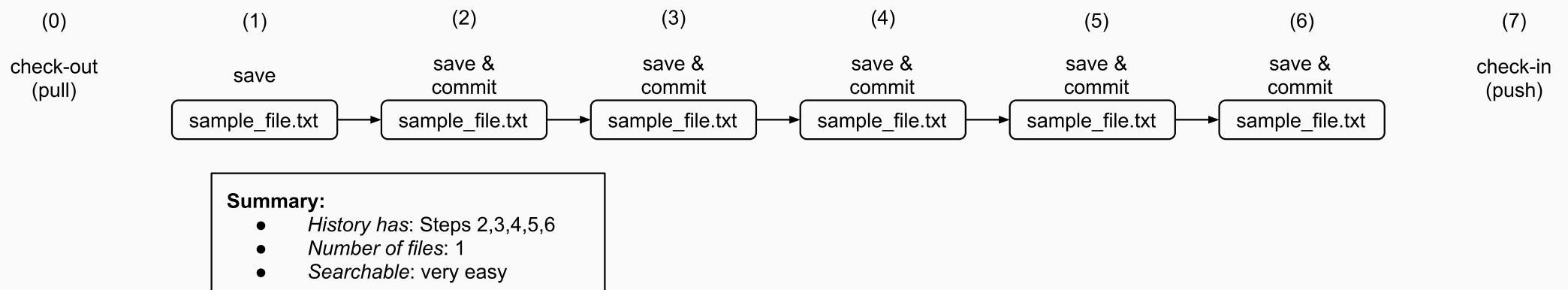


Version Control Software: Comparison of Workflows

Strategy 1: Renaming



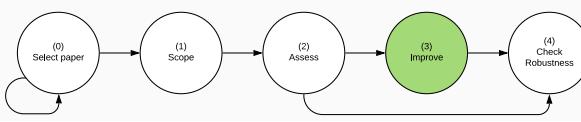
Strategy 2: Version Control Software



Improvements: Display item-level

- Adding missing raw data: files or meta-data
 - Example: "Add raw temperature and relative humidity data"
- Adding missing analytic data files
 - Example: "Copy the raw files from Data folder into new Analysis\trade cost\Input"
- Adding missing analysis or cleaning code
 - Example: "Replaced broken Wald bootstrap code with updated code/command"
- Debugging code
 - Example: "was counting each group 4 times in round 1, so fixed that"

Key challenge: constructive exchanges with original authors and other reproducers



Improvements: future possible

We ask reproducer to leave concise and actionable tasks for other reproducers in the future.

Example 1:

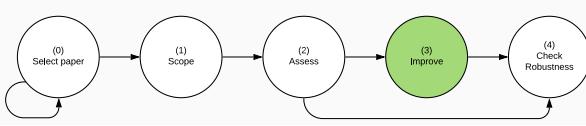
"Revise the .aml and .bat code scripts to reflect reorganized structure"

Example 2:

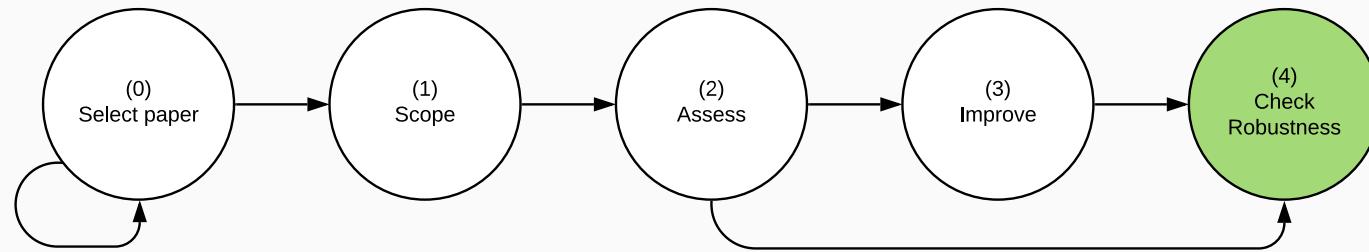
"Provide data and codes generating the other two figures in the paper, which are not given in the replication file."

Example 3:

"Table 3 can be reproduced identically from the [...] analytic data files. I was not able to reproduce the analytic data files due to lack of access to ArcGIS software, but the code scripts and raw data files [...] are included in the reproduction package."



Robustness Checks



Two main parts for robustness:

1. Increase the number of robustness checks
2. Justify the appropriateness of a specific test

Robustness

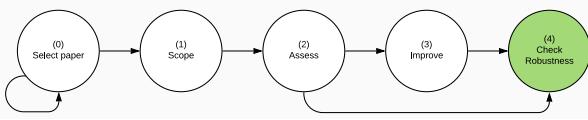
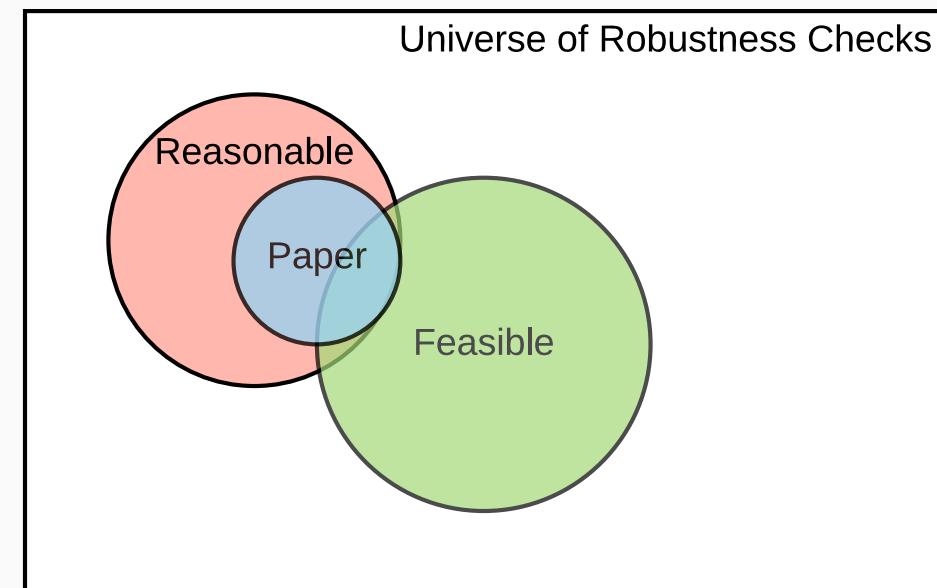
Robustness checks: any possible change in a computational choice, both in data analysis and data cleaning

Reasonable specifications (Simonsohn et. al., 2018):

1. Sensible tests of the research question
2. Expected to be statistically valid, and
3. Not redundant with other specifications in the set.

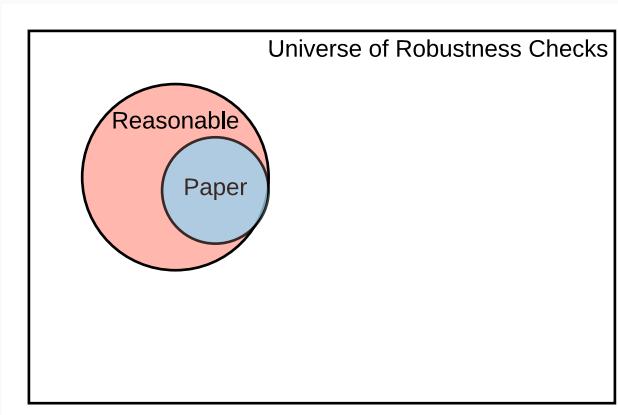
Reproducers will be able to record two types of contributions:

- Mapping the universe of robustness checks
- Proposing a specific robustness check

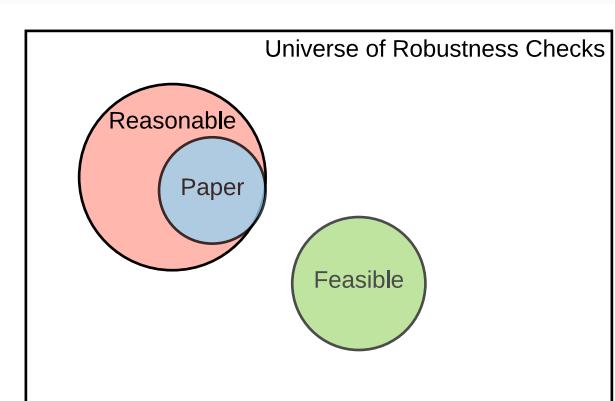


Robustness & Reproducibility

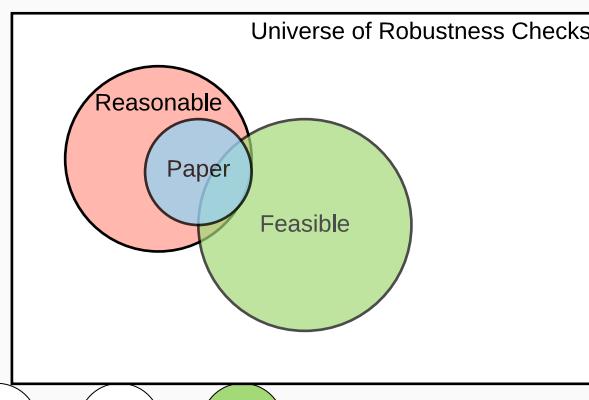
Robustness with level 1



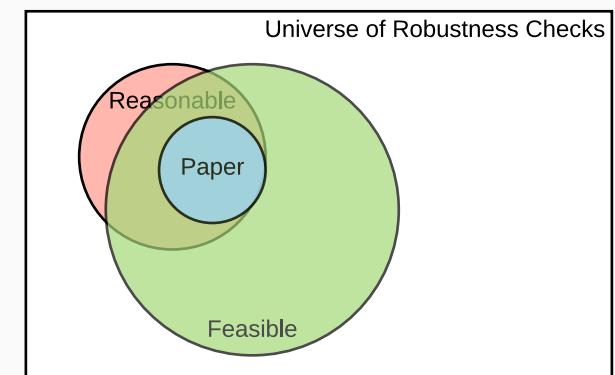
Robustness with levels 2-4



Robustness with levels 5-9



Robustness with level 10



Promoting a Constructive Exchange

For reproducers addressing original authors

- 1 - Contacting the original author(s) when there is no reproduction package
- 2 - Contacting the original author(s) to request specific missing items of a reproduction package
- 3 - Asking for additional guidance when some materials have been shared
- 4 - Response when the original author has refused to share due to *undisclosed reasons*
- 5 - Response when the original author has refused to share due to legal or ethical restrictions of the data
- 6 - Contacting the original author to share the results of your reproduction exercise

For original authors responding to requests from reproducers

- 1 - Responding to a repeated request that has been addressed in an earlier interaction
- 2 - Acknowledging that the author no longer has access to certain part(s) of the reproduction package
- 3 - Acknowledging that some material is still embargoed for future research
- 4 - Responding to incomplete/unclear requests

Example 1: There is no reproduction package

Subject: Reproduction package for ["Title of the paper"]

Dear Dr. [Lastname of Corresponding Author],

I am contacting you to request a reproduction package for your paper titled [Title] which was published in [Reference]. A reproduction package may contain (raw and/or analytic) data, code, and other documentation that makes it possible to reproduce paper. Would you be able to share any of these items?

I am a [position] at [Institution], and I would like to reproduce the results, tables, and other figures using the reproduction materials mentioned above. I have chosen this paper because [add context ...]. **Unfortunately, I was not able to locate any of these materials on the journal website, Dataverse [or other data and code repositories], or in your website.**

I will record the result of my reproduction attempt on ACRE [...]. With your permission, I will also record the materials you share with me, which would allow access for other reproducers and avoid repeated requests directed to you. Please let me know if there are any legal or ethical restrictions that apply to all or parts of the reproduction materials so that I can take that into consideration during this exercise.

In addition to your response above, would you be available to respond to future (non-repetitive) inquiries from me or other reproducers conducting an ACRE excercise? **Though your cooperation with my and/or any future request would be extremely helpful, please note that you are not required to comply.**

Since I am required to complete this project by [date], I would appreciate your response by [deadline].

Let me know if you have any questions. Please also feel free to contact my supervisor/instructor [Name (email)] for further details on this exercise. Thank you in advance for your help!

Example 1: Following up on additional materials

Template email:

Subject: Clarification for reproduction materials for [“Title of the paper”]

Dear Dr. [Lastname of Corresponding Author],

Thank you for sharing the materials. They have been immensely helpful for my work.

Unfortunately, I ran into a few issues as I delved into the reproduction exercise, and I think your guidance would be helpful in resolving them. **[Describe the issues and how you have tried to resolve them. Describe whatever files or parts of the data or code are missing. Refer to examples 1 and 2 below for more details]**.

Thank you in advance for your help.

Best regards,

[Reproducer]

An example of well described issues:

Specifically, I am attempting to reproduce OUTPUT X (e.g., table 1, figure 3). I found that the following components are required to reproduce to reproduce OUTPUT X:

```
OUTPUT X
└── [code] formatting_table1.R
    ├── output1_part1.txt
    │   └── [code] output_table1.do
    │       └── [data] analysis_data01.csv
    │           └── [code] data_cleaning01.R*
    │               └── [data] UNKNOWN
    └── output1_part2.txt
        └── [code] output_table2.do
            └── [data] analysis_data02.csv
                └── [code] data_cleaning02.R
                    └── [data] admin_01raw.csv*
```

I have marked with an asterisk (*) the items that I could not find in the reproduction materials: **data_cleaning01.R** and **admin_01raw.csv**. After accessing these files, I will also be able to identify the name of the raw data set required to obtain output1_part1.txt. This is to let you know that I may need to contact you again if I cannot find this file (labeled as **UNKNOWN** above) in the reproduction materials.

I understand that this request will require some work for you or somebody in your research group, but I want to assure you that I will add these missing files to the reproduction package for your paper on the ACRE platform. **Doing this will ensure that you will not be asked twice for the same missing file.**

Ok, I get it. But what is in for me?

- Standardized homework/project: everything is set up in terms of structure and deliverables.
- Easy to grade (homework format).
- Easy to guide and oversee (undergraduate dissertation format).
- Easy to setup as an independent study.
- Reduces duplication of requests to authors.
- Facilitates a constructive exchange of ideas.
 - When emailing authors.
 - When discussion reproduction attempts.
- Personal satisfaction that you're contributing a public good to the profession!

Suggested timelines for different formats

| | 2 weeks (~10 days) | | 1 month (~20 days) | | 1 semester (~100 days) | |
|-------------|-----------------------|----------|-----------------------|----------|---------------------------|----------|
| | analysis data | raw data | analysis data | raw data | analysis data | raw data |
| Scoping | 10% (1 day) | | 5% (1 day) | | 5% (5 days) | |
| Assessment | 35% | | 25% | | 15% | |
| Improvement | 25% | 0% | 40% | | 20% | 30% |
| Robustness | 25% | 5% | 25% | | 25% | |

Easy to grade: report 1

This browser does not support PDFs. Please download the PDF to view it: [Download PDF](#).

Easy to grade: report 1

This browser does not support PDFs. Please download the PDF to view it: [Download PDF](#).

This browser does not support PDFs. Please download the PDF to view it: [Download PDF](#).

Suggested timelines for different formats

| | 2 weeks (~10 days) | | 1 month (~20 days) | | 1 semester (~100 days) | |
|-------------|-----------------------|----------|-----------------------|----------|---------------------------|----------|
| | analysis data | raw data | analysis data | raw data | analysis data | raw data |
| Scoping | 10% (1 day) | | 5% (1 day) | | 5% (5 days) | |
| Assessment | 35% | | 25% | | 15% | |
| Improvement | 25% | 0% | 40% | | 20% | 30% |
| Robustness | 25% | 5% | 25% | | 25% | |

The screenshot shows a web browser window with the URL <https://www.socialsciencereproduction.org> in the address bar. The page is titled "BETA - Site is under development. Please provide feedback and bug reports to acre+feedback@berkeley.edu". The main content area has an orange header bar with three links: "Events", "for Reproducers", and "for Instructors". Below this is a dark blue section containing the text "Welcome to the Social Science Reproduction Platform" and "Accelerating computational reproducibility in economics -- one paper at a time!". A "GET STARTED!" button is located at the bottom of this section. The overall design is clean and modern.

Purpose

On the Social Science Reproduction Platform, you can record and review **verifications and improvements** to the **computational reproducibility** of published social science work.

This open source platform was developed by the Berkeley Initiative for Transparency in the Social Sciences ([BITSS](#)) in collaboration with the [American Economic Association Data Editor](#).

Table of Contents

1. Problems
2. Solutions
3. Accelerating Computational Reproducibility
4. **Open Policy Analysis**

Open Policy Analysis (OPA)

Core BITSS Team

Aleks Bogdanoski

Fernando Hoces de la Guardia

Katie Hoeberling

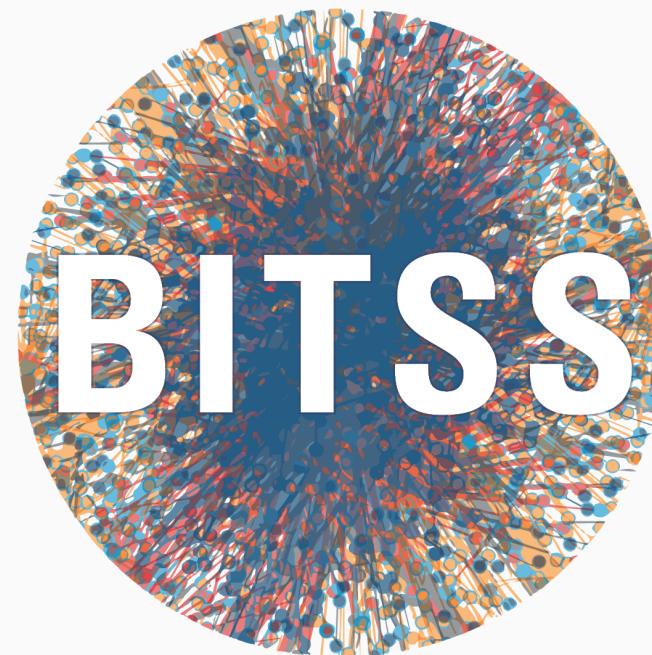
Teng Aleksandra Ma

Edward Miguel

Undergraduate RAs

Sophia Bai

Orianna Jia



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Keanu Lim
Charles Love
Emma Ng
Ryan Scholes
Jiayu Shi

Evidence Action

Grace Hollister

Grace Morgan

BITSS is part of the Center
for Effective Global Action (CEGA).



Motivation: Rise of Alternative Facts



Senator *discussing facts* on unemployment insurance

Interviewer: “We I have looked at what economist are saying, and [...] there is no measurable evidence that people are staying at home because of [\$600 unemp. insurance]”

Congressman: “[scoffs] I don’t know which economist you are talking about, but ...”

Possible Mechanism: Opaqueness of Policy Analysis

- Incredible Certitudes (Manski, 2011)
- Report wars (Wesselink et al, 2013)
- Low overall credibility of PA → credibility based on reputation ("serious") → assignment of reputation varies across political positions
- Parallels to "Reproducibility Crisis" and Open Science reponse (Hoces de la Guardia, Grant, Miguel 2020). Propose:
 - Core principles for Open Policy Analysis (OPA) in outputs, analysis and materials (reproducibility)
 - Agenda to implement and document how OPA can be applied into several policy issues
- Growing adoption of open science methods into policy analysis

Deworming Interventions

- Parasitic worm infections are endemic in many countries, disproportionately affecting the poor
- They interfere with regular bodily processes by decreasing nutrient uptake and can thus lead to serious consequences on human health, education outcomes, and long-term economic well being
- Mass deworming interventions, at school level, have been proposed as a cost-effective approach to tackle this problem

Different settings for deworming

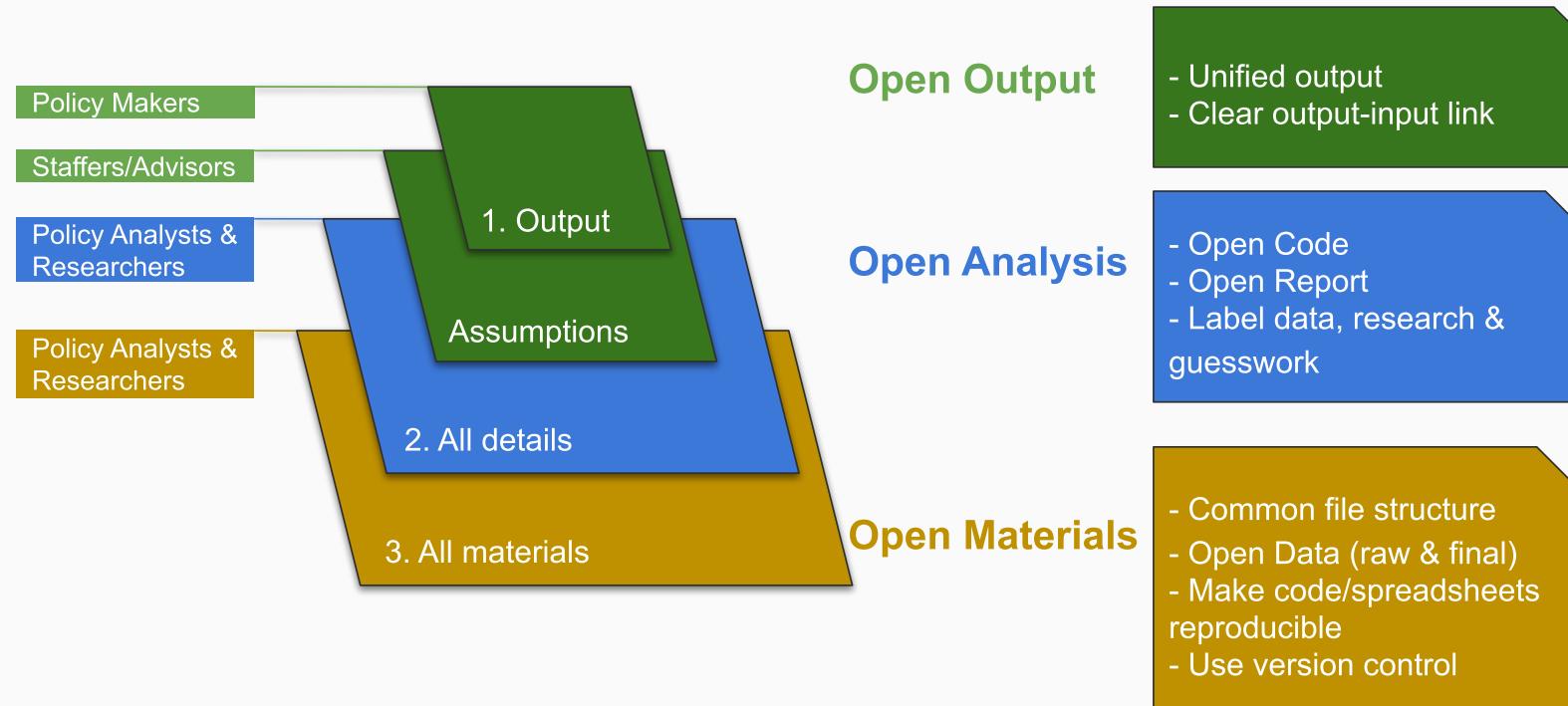
- Context of original study (Kenya, 1998-99) had very high prevalence rates of worm infections
- Implementation costs were very low (\$0.42 per round of treatment)
- Length of treatment was relatively short (2.4 years)
- Current deworming settings have lower prevalence rates, varying implementation costs and length of treatments

Strong debate around initial results

- Ozier (2020) summarizes differences between original findings (Miguel and Kremer 2004) and a re-analysis (Aiken et al, 2015). Emphasizes the role of communication of results in a reanalysis.
- This type of debate (result/re-analysis) can be seen in several other topics. For example: minimum wage, immigration, taxation.
- OPA need not guarantee agreement on key research finding, but should help avoid multiple policy reports

A Framework for Open Policy Analysis

Hoces de la Guardia,
Grant, Miguel (2020)

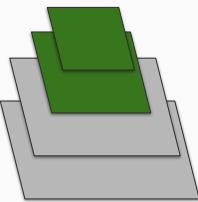


Contributions of OPA to deworming:

1. Selected one policy estimate among several alternatives and establish a clear link between it and underlying assumptions
2. Added documentation to increase reproducibility
3. Created a public repository with all materials for one-click reproducibility

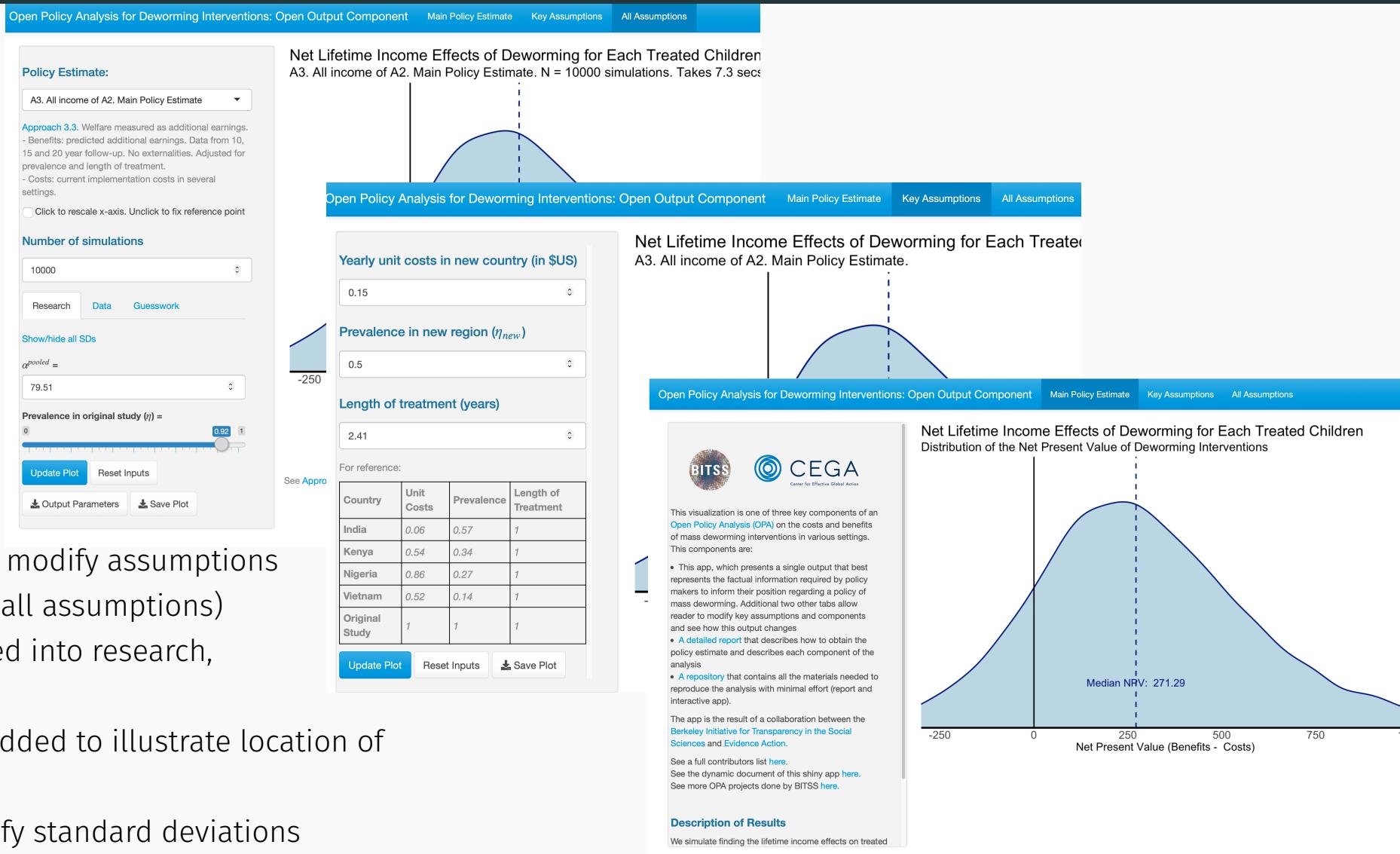
Open Output

Demo



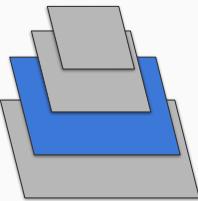
Main features

- One clear output previously agreed in consultation with policy partner
- Two additional tabs to modify assumptions (key assumptions and all assumptions)
- Each source is classified into research, data, or guesswork
- High level equations added to illustrate location of components
- Added feature to modify standard deviations
- Track values of each component



Open Analysis

Demo



Main features

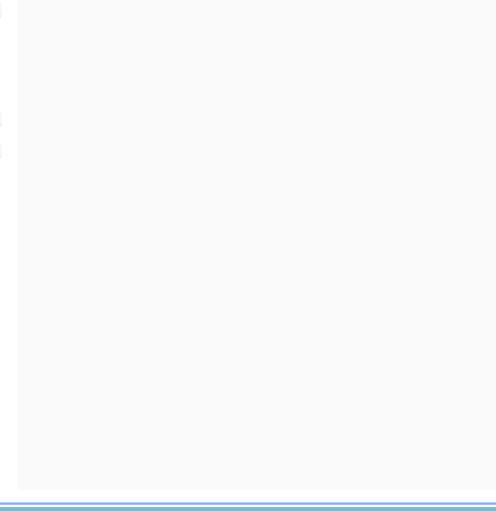
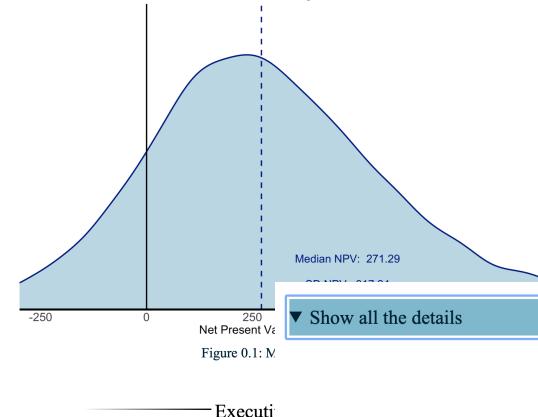
- Complete narrative description of the methodology
- Translation of each narrative step into an equation
- Implementation of each equation into code
- Combine all of the above into using a dynamic document (RMarkdown)
- Presentation of narrative, equations, and code in layered fashion to avoid overwhelming the reader
- Icon figure

BITSS CEGA
Open Policy Analysis
1 Introduction
2 Methodology
3 Main Results
References

OPEN POLICY ANALYSIS FOR DEWORMING

18 December, 2020

Net Lifetime Income Effects of Deworming for Each Treated Children
Distribution of the Net Present Value of Deworming Interventions



$$B = \sum_{t=0}^{50} \left(\frac{1}{1+r} \right)^t E_t \quad (1)$$

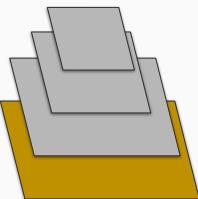
Where:

- E_t : earnings individuals are expected to generate at period t
- r : real interest rate as the discounting rate
- t : period t . Period 0 represents time of intervention. Individuals are assumed to enter the labor market 9 years after treatment.

```
# - inputs: stream earnings, discounting rate, number of periods
# - outputs: function that computes the present value of benefits
chunk_benefits <- function(){
#####
pv_benef_f <- function(
  earnings_var = earnings_in,
  interest_r_var = interest_in,
  periods_var = periods_so
) {
  index_t <- 0:periods_var
  res1 <- sum( ( 1 / (1 + interest_r_var) )^index_t * earnings_var )
  return(res1)
}
```

Open Materials

Demo



Main features

- One-click reproducible documentation and app
- Extensive readme files
- Clear folder structure
- Version controlled
- Open data
- Acknowledgment to all contributors

BITSS-OPA / [opa-deworming](#)

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

master 21 branches 1 tag Go to file Add file Code

fhoces Change title of readmen file 53bb6f1 1 minute ago 728 commits

.binder update install.R 2 months ago

code Merge branch 'master' of <https://github.com/fhoces/opa-deworming> 1 hour ago

data

docs

rawdata

.gitignore

contributors.R

contributors.csv

opa-deworming.Rproj

readme.Rmd

readme.md

RStudio File Edit Code View Plots Session Build Debug Profile Tools Window Help

OS_final_opa.Rmd

```
1 ---  
2 title: "<center><div class= 'mytitle'>Open Policy Analysis for Deworming</div></center>"  
3 date: "<center><div class='mysubtitle'>r format(Sys.time(), '%d %B, %Y')<br><img height='80px' src='./shiny_app/www/bits_logo_horizontal.png'><br><img height='80px' src='./shiny_app/www/CEGA_logo.png'></div></center>"  
4 editor_options:  
5 chunk_output_type: console  
6 output:  
7 bookdown::html_document2:  
8 code_download: yes  
9 code_folding: hide  
10 css: style.css  
11 highlight: tango  
12 includes:  
13   after_body: footer.html  
14 keep_md: yes  
15 number_sections: yes  
16 smooth_scroll: no  
17 theme: cerulean  
18 toc: yes  
19 toc_collapsed: no  
20 toc_depth: 3  
21 toc_float: yes  
22 html_document:  
23 df_print: paged  
24 toc: yes  
25 toc_depth: 3  
26 word_document: null  
27 link_citations: yes  
28 pdf_document:  
29   extra_dependencies: xcolor  
30   fig_caption: no  
31 biblio: bibliography.bib
```

readme.md

Open Policy Analysis of Deworming

BITSS BERKELEY INITIATIVE FOR TRANSPARENCY IN THE SOCIAL SCIENCES

R version 4.0.0 (2020-04-24) -- "Arbor Day"
(Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin17.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
(type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.)

OPEN PO

BITSS CEGA Center for Effective Global Action

Open Policy Analysis

1 Introduction
2 Methodology
3 Main Results
References

Net Lifetime Income Effect
Distribution of the Net Present Value

60 / 66

What happens if one principle fails: Opaque Output

- Policy estimates of approach 2 of deworming (Baird et al, 2016)
- Assume Open Analysis and Open Materials
- Good for researchers
- Unclear for policy analysts
- Allows policy makers to cherry pick results
- Does not disclose uncertainty
- Unclear separation of roles between policy analyst and policy maker (Truman's request for a "one-handed" economist)

TABLE V
FISCAL IMPACTS OF DEWORMING SUBSIDIES

| | No subsidy | Partial subsidy | Full subsidy | Notes |
|---|------------|-----------------|--------------|---|
| Panel A: Calibration parameters | | | | |
| Size of subsidy: S | \$0.00 | \$1.15 | \$1.42 | From Deworm the World; Kremer and Miguel (2007) |
| Take-up rate: $Q(S)$ | 5% | 19% | 75% | From Kremer and Miguel (2007) |
| Average per-person cost: $SQ(S)$ | \$0.00 | \$0.22 | \$1.07 | Subsidy \times take-up rate |
| Mean per person increase in work hours/week: λ_1 | 0.00 | 0.44 | 1.75 | Men: increase of 3.49 hours/week; women: no change (Table III). Partial subsidy multiplied by $\frac{Q(S)}{Q(\text{full})}$ |
| Mean increase in work hours/week from externality: $p\lambda_2$ | 0.00 | 1.76 | 5.21 | 10.20 (Table III) \times Coverage of treatment school students within 6 km (R , 68.1%) $\times [Q(S)$ for full subsidy, $\frac{Q(S)}{Q(\text{full})}$ for partial subsidy] |
| Mean increase in schooling costs | 0.00 | 2.71 | 10.71 | NPV of (additional secondary schooling costs per pupil-year (\$116.85) \times direct increase in secondary schooling). Partial subsidy multiplied by $\frac{Q(S)}{Q(\text{full})}$. |
| Mean increase in schooling costs from externality | 0.00 | 3.40 | 13.42 | NPV of (additional secondary schooling costs per pupil-year (\$116.85) \times externality increase in secondary schooling). Partial subsidy multiplied by $\frac{Q(S)}{Q(\text{full})}$. |
| Panel B: no health spillovers | | | | |
| Annual increase in per person earnings | \$0.00 | \$3.91 | \$15.44 | $\lambda_1 \times$ starting wage \times 52 |
| NPV increase in per person earnings (relative to no subsidy) | — | \$36.08 | \$142.43 | 9.85% annual (real) interest rate in Kenya |
| NPV increase in per person government revenue | — | \$3.27 | \$12.90 | NPV earnings \times 16.575% tax rate – Direct schooling costs |
| Panel C: With health spillovers | | | | |
| Annual increase in per person earnings | \$0.00 | \$26.77 | \$83.11 | $(\lambda_1 + \frac{p}{R} \lambda_2) \times$ starting wage \times 52 |
| NPV increase in per person earnings (relative to no subsidy) | — | \$246.99 | \$766.81 | 9.85% annual (real) interest rate in Kenya |
| NPV increase in per person government revenue | — | \$34.83 | \$102.97 | NPV earnings \times 16.575% tax rate – (Direct externality schooling costs) |

Opaque Analysis/Materials

- Assume open output and open materials
- In this scenario we can have two policy analyses: one from advocates and one from opponents
- Each analysis can claim to be open source. Code is available, reproducible, and each as an interactive app.
- But each analysis can bury in their code analytical choices that move the final policy estimate in their favor
- Same though exercise can be done with opaque materials

Lessons for future OPA projects: Timeline

1. Reviewed existing documentation
2. Review code/spreadsheets and write down missing documentation.
3. Add equations to reflect any additional clarifying steps.
4. Translate analysis to code scripts (ideally open source software)
5. Check for computational reproducibility and consult with original analysts for discrepancies/further questions
6. Discuss with policy partner on target policy estimate
7. Draft doodle of app
8. Incorporate any potential new analysis (repeat steps 1 - 3)
9. Present draft of OPA (DD and doodle of app) to policy partner
10. Incorporate comments into documentation of the OPA
11. Build app based on analysis of dynamic document
12. Incorporate any additional features requested by policy partner
13. Verify that all three components produce the same output
14. Publish v1.0

Useful practices developed during the project

Documentation of analysis and app are connected

- Each analytic step is declared in the documentation and wrapped in a function
- All these steps are used to produce the results in the documentation
- These steps are also stored in a file that is later sourced in the app.
- If something changes in the report, it will update the app.

Style guide

- Group objects according to use: only within the documentation, for simulations, and for the app
- Use diagram trees to depict nested relationships

Lessons for future OPA projects: Costs

- Costs (approx bandwidth over a year at full time):
 - Principal Investigator 30-50%
 - Research assistant/programmer 100-150%
 - Program Manager 20%-30%
 - Original researcher: 1-5%
- Not all policy analysis justify this level of effort
- Characteristics that might justify an OPA:
 - Topics with strong disagreement on the facts among analysts
 - Recurrent reports (eg. ex-ante economic analysis from development banks/agencies)
 - Topics that have large expected welfare effects (eg. tax reform, social cost of carbon)
- With each new OPA project, templates will emerge and costs will likely fall

Additional Benefits of OPA

Easy to update and reuse

After deworming OPA is released, anybody can modify and improve into a newer version

Clearer connection of how evidence from research is used in policy analysis

Researchers can see clearly where their estimates are being used in a policy analysis. For example, the OPA can be used to justify power calculations of potential new studies.

Connection with forecasting

When there is little information for a parameter used in an OPA, a forecasting exercise can be carried out to elicit expert knowledge (DellaVigna, Pope, Vivaldi 2019).