

Transparency in the Evidence-Based-Policy Pipeline: Challenges and Solutions

Data H195A, UC Berkeley

Fernando Hoces de la Guardia, BITSS
November 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 and policy analysis.

Core Team

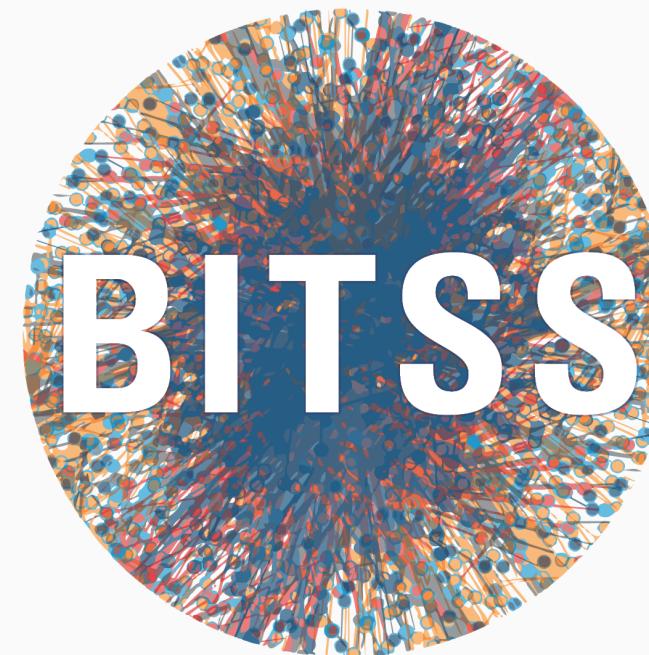
Aleks Bogdanoski

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

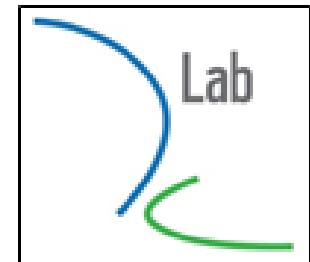
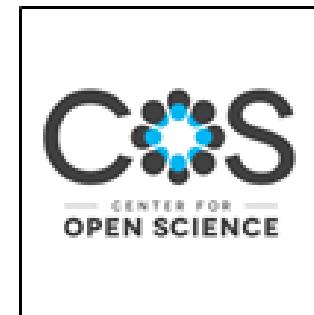
CEGA staff

Undergrad and Graduate RAs

Catalysts

Outside Collaborators
(Researchers, Programmers)

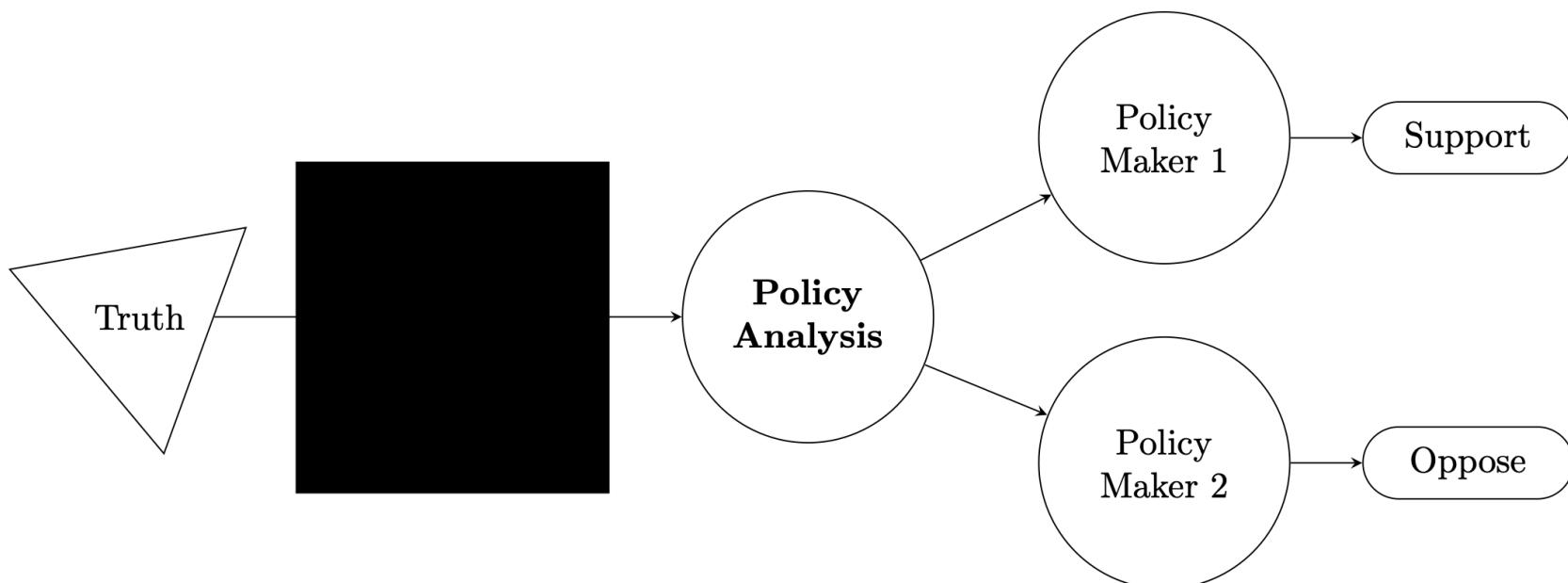
Part of the much larger Open Science Community



ICPSR

Today's Presentation

- **Part I: Transparency in Research**
- Part II: Transparency in Policy Analysis



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

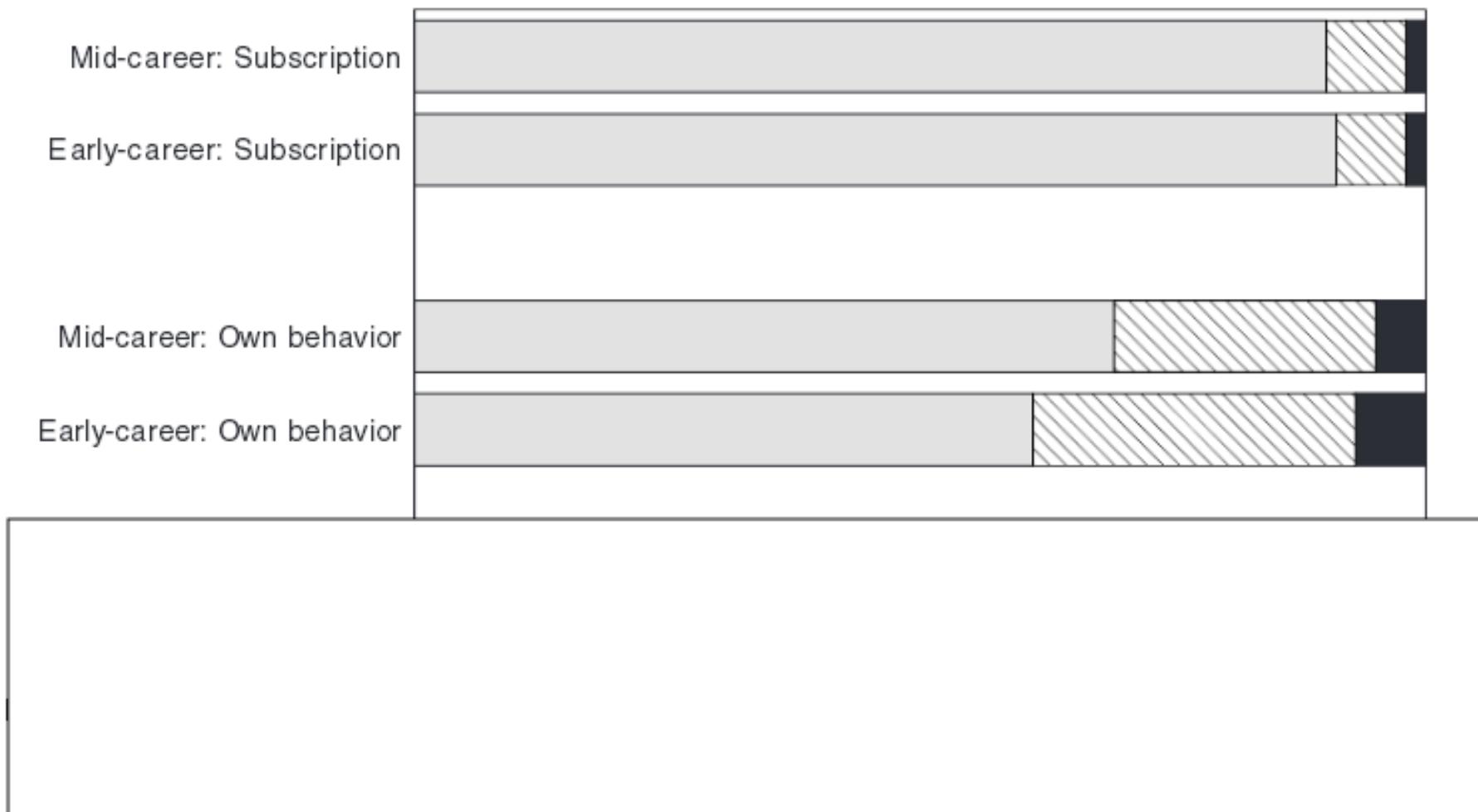


FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

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

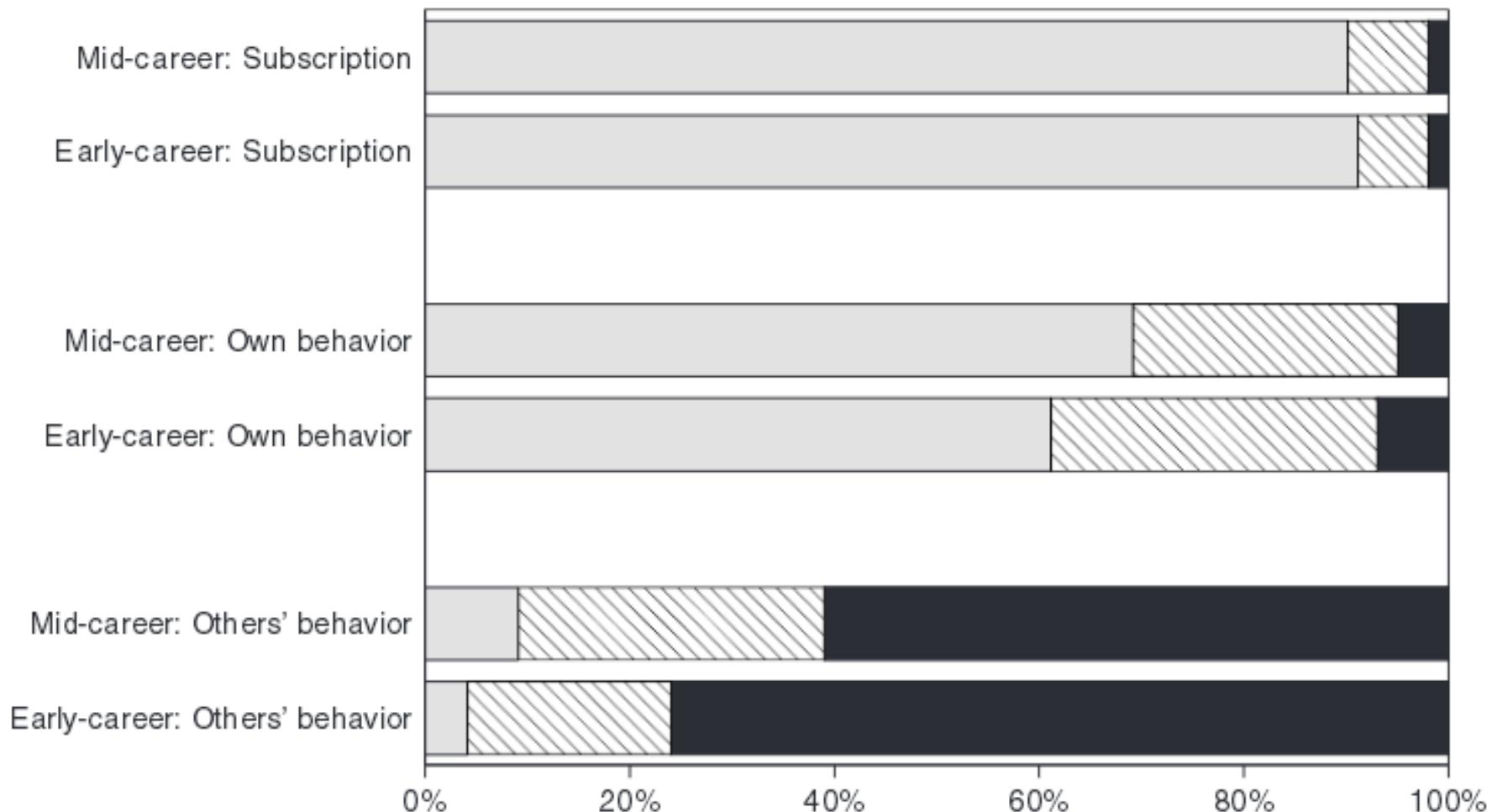
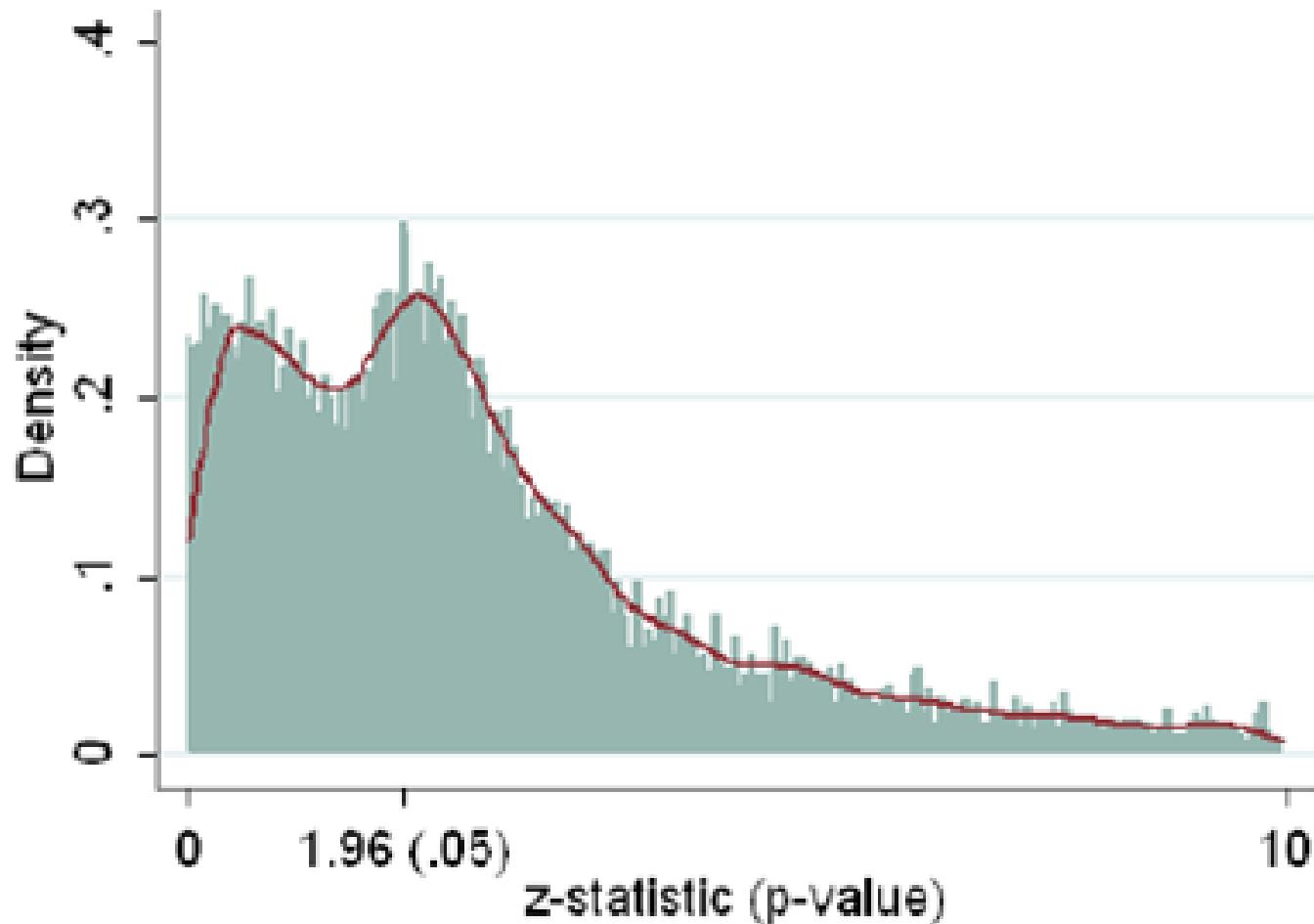


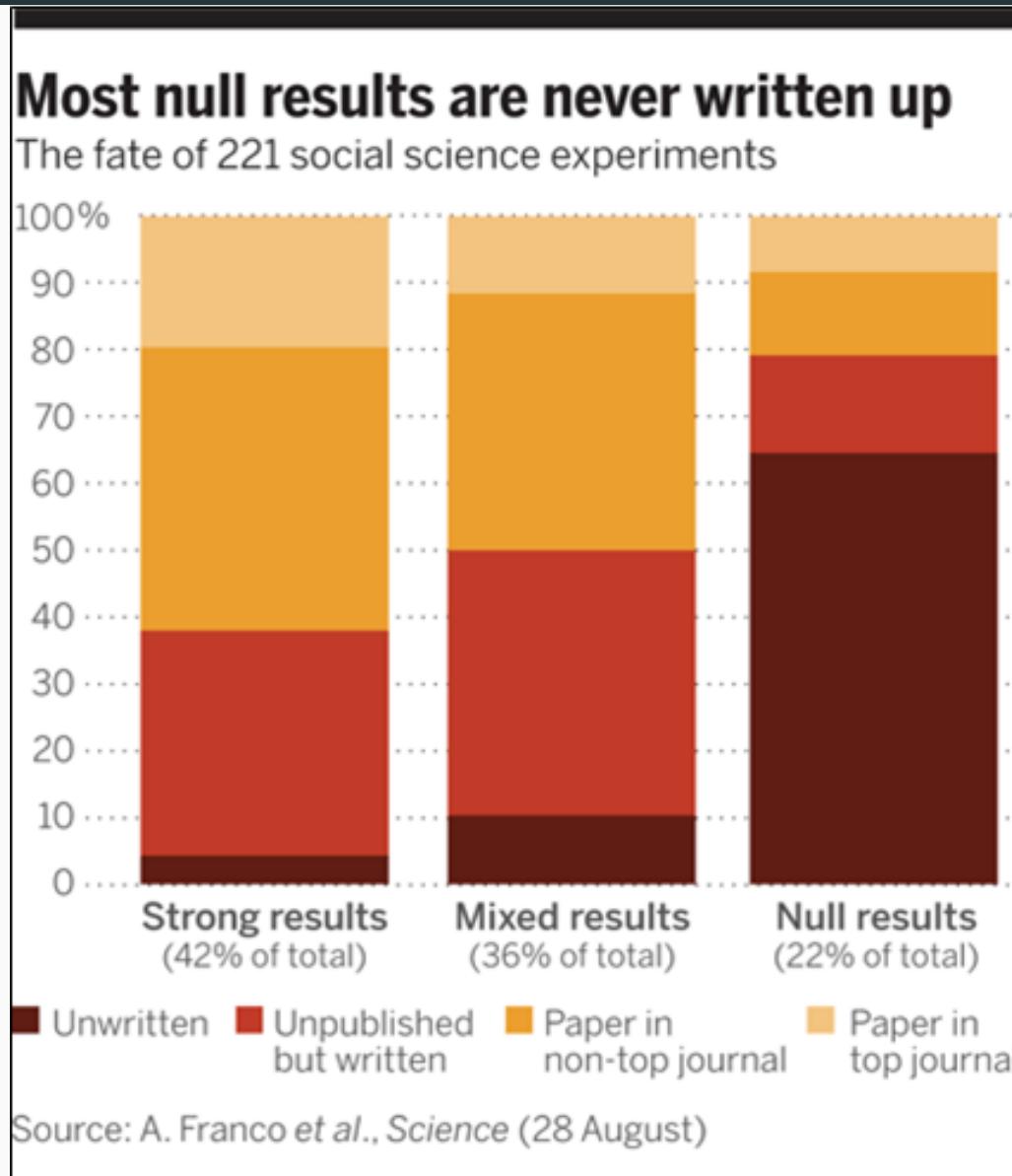
FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

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

(b) Unrounded distribution of z-statistics.



Publication Bias (Franco et. al. 2014)



Low Replicability and Reproducibility ("Reproducibility Crisis")

Replication in Social Sciences (same method, different sample)	Reproduction in Economics (same data and methods)
OSC (2015): 30%-60%	Chang & Li (2015): 43%
Camerer et. al. (2016): ~60%	Gertler et. al. (2017): 14%
Nosek & Camerer et. al. (2018): ~60%	Kingi et. al. (2018): 43%
Klein et. al. (2018): 50%	Wood et. al. (2018): 25%

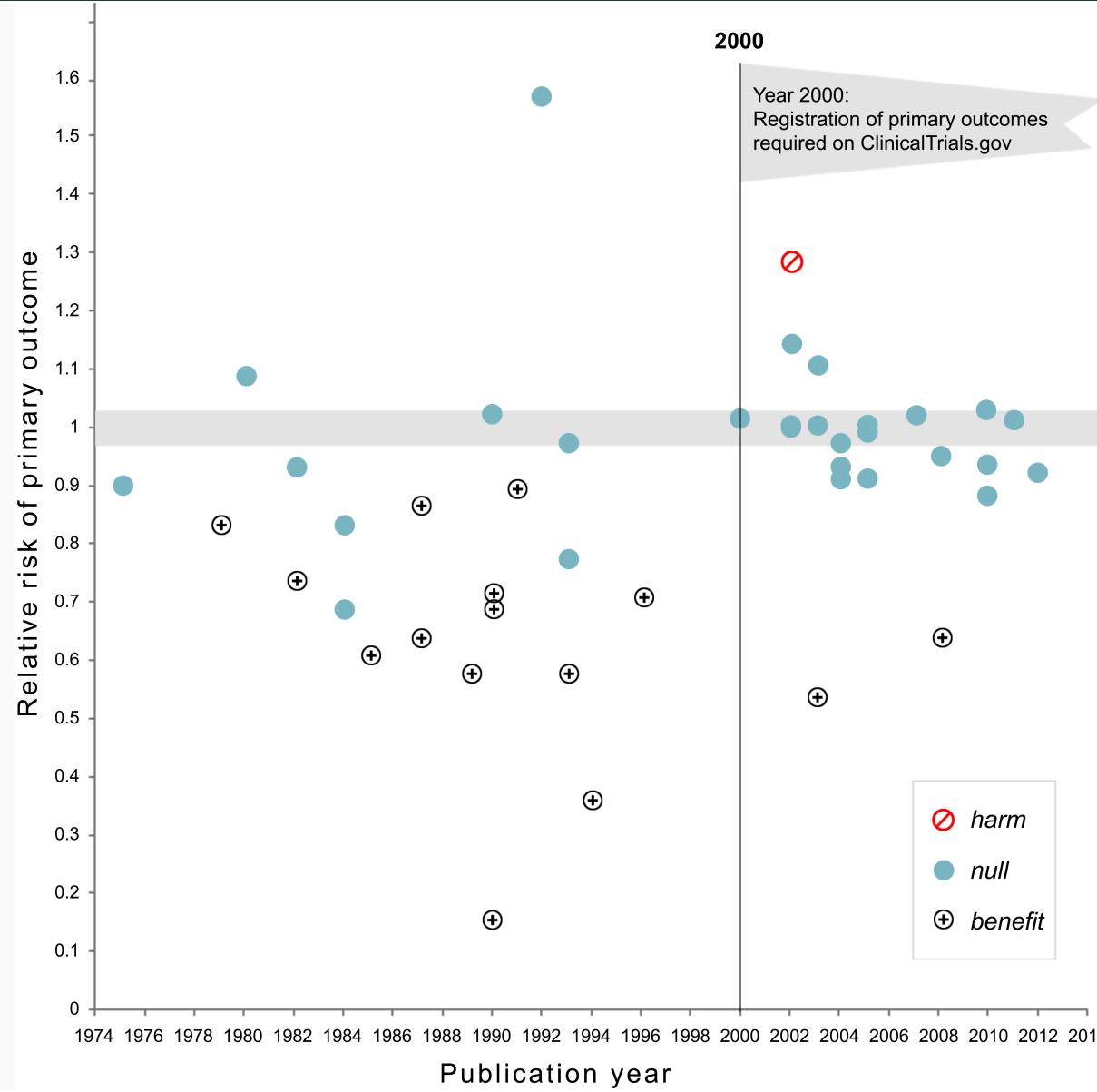
Main Solutions

1. Registrations
2. Pre Analysis Plans
3. Computational Reproducibility
4. Others: Reporting Guidelines, Pre-prints, etc.

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

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).

Computational Reproducibility: Why?

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.”

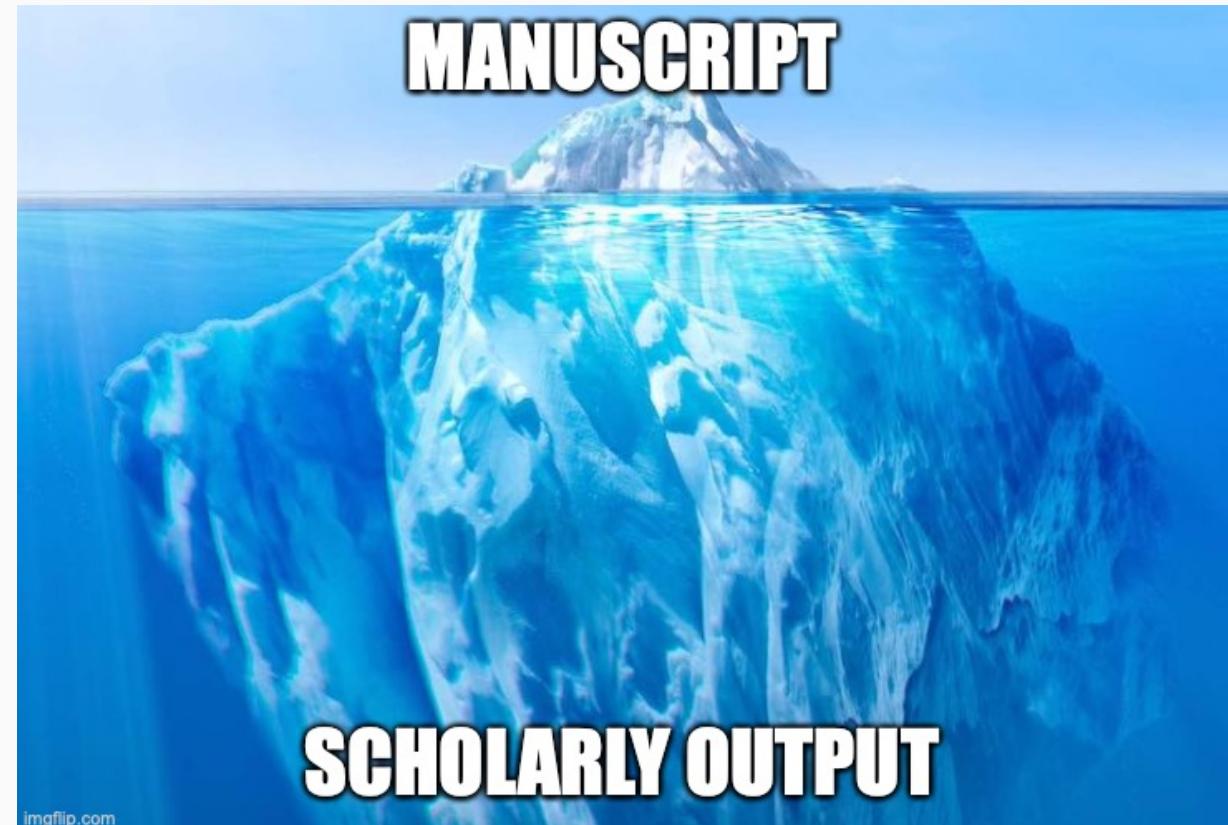
Buckheit and D.L. Donoho (1995, 2009)

Computational Reproducibility: Why?

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)



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

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

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.

Context for ACRE

- American Economics Association (AEA) creates first data policy in 2006.
 - Must publish some data (waivers available)
- AEA updates **policy in 2019.**
 - Must post all data and code. Publication is conditional on verifying reproducibility (if confidential: must document extensively)
 - A new requirement is to post all cleaning code, even for data that is not public
- We should expect high levels of computational reproducibility after 2019 (AEA).
- We should not demand 100% reproducibility before, but we could identify the gaps and try to improve some.

Accelerating Computational Reproducibility: Framework

Beyond Binary Judgments

Reproductions can easily gravitate towards adversarial exchanges.

- Early career researchers (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"

ACRe Framework

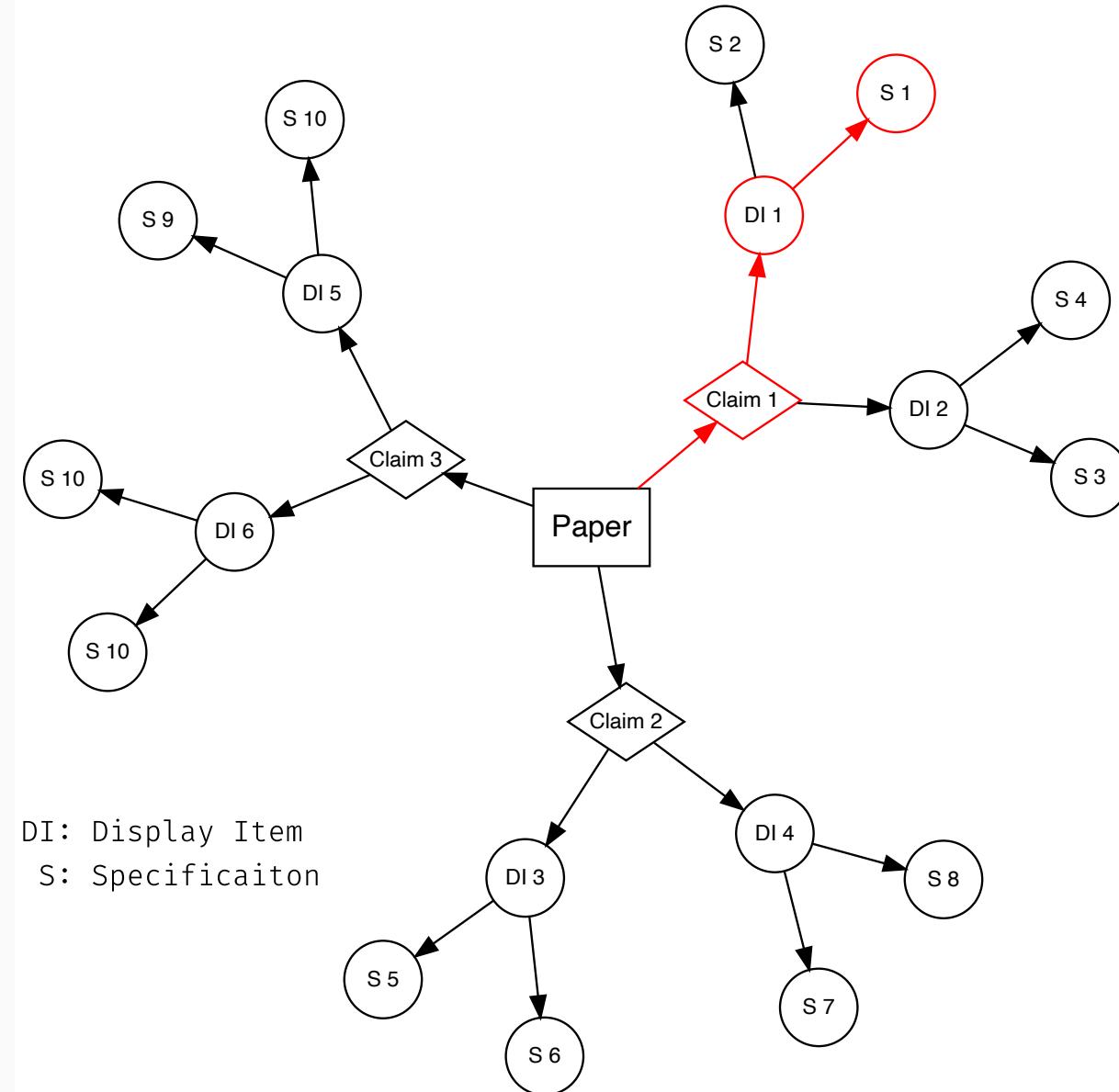
Each **reproduction attempt** is centered around scientific **claims** (following **SCORE**).

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.

Key challenge: **standardization** of concepts and formats.



Demo: socialsciencereproduction.org

Home

About

Events

Guide

Forum

Metrics

Welcome to the Social Science Reproduction Platform

Accelerating computational
reproducibility in the social sciences

GET STARTED!

Search reproductions by paper title, DOI, or author name



Select a paper

Article Information

Abstract

How large are the benefits of transportation infrastructure projects, and what explains these benefits? This paper uses archival data from colonial India to investigate the impact of India's vast railroad network. Guided by four results from a general equilibrium trade model, I find that railroads: (1) decreased trade costs and interregional price gaps; (2) increased interregional and international trade; (3) increased real income levels; and (4) that a sufficient statistic for the effect of railroads on welfare in the model accounts well for the observed reduced-form impact of railroads on real income in the data.

Citation

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*, 108 (4-5): 899-934.

DOI: 10.1257/aer.20101199

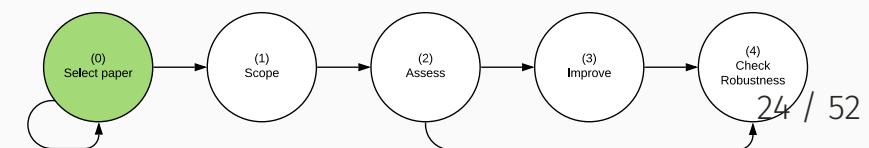
Choose Format: ▾

Additional Materials

Data Set (17.80 MB)

Online Appendix (247.97 KB)

Author Disclosure Statement(s) (63.46 KB)



Scope

Go to page 19

C. Results

Table 2 presents ordinary least squares (OLS) estimates of equation (12). In column 1 I estimate the effect of the lowest-cost route effective distance on trade costs when the relative costs of each mode (α) are set to observed historical relative freight rate estimates. I use the relative per unit distance freight rates described in Section IB (at their midpoints): $\alpha^{road} = 4.5$, $\alpha^{river} = 3.0$, and $\alpha^{coast} = 2.25$ relative to the freight rate of railroad transport, normalized to 1). Column 1 demonstrates that the elasticity of trade costs with respect to the lowest-cost route effective distance, calculated at observed freight rates, is 0.088, and this is statistically significant at the 5 percent level.

However, as argued in Section IB, it is possible that these observed relative freight rates do not capture the full benefits (such as increased certainty or savings) of railroad transport relative to alternative modes of transportation. For this reason the NLS specification in column 2 estimates the relative freight rates (i.e., the parameters α) that minimize the sum of squared residuals in equation (12).

Column 2 is my preferred specification. When the mode-wise distance costs (i.e., α) are not restricted to be equal to the observed freight rates, the estimated elasticity of trade costs with respect to effective distance (i.e., δ) rises to 0.169. Even when controlling for all unobserved, time-constant determinants of trade costs between all salt sources and destinations, as well as unrestricted shocks to the source price of each salt type, reductions in trade costs along lowest-cost routes (estimated from railroad-driven time variation in these routes alone) have a large effect on reducing salt price gaps over space.

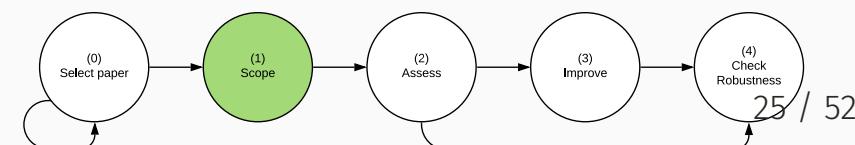
Col 1: OLS with alpha at historical rates. 1% inc in effc dist (in km) -> 0.088% inc in trade costs

Col 2: NLS with alpha chosen to minimize SSR of col1. effc dist (in km) -> 0.169% inc in trade costs

TABLE 2—RAILROADS AND TRADE COSTS: STEP 1

	(1)	(2)
Dependent variable: log salt price at destination		
log effective distance to source, along lowest-cost route (at historical freight rates)	0.088 (0.028)	0.169 [0.062, 0.296]
log effective distance to source, along lowest-cost route (at estimated mode costs)		
Estimated mode costs per unit distance:		
Railroad (normalized to 1)		
Road	2.375 [1.750, 10.000]	N/A
River	2.250 [1.500, 6.250]	
Coast	6.188 [5.875, 10.000]	
Observations	7,345	7,345
R^2	0.946	0.946

Notes: Regressions estimating equation (12) using data on 6 types of salt (listed in online Appendix A), from 133 districts in Northern India, annually from 1861 to 1930. Column 1 and column 2 estimated by OLS and NLS respectively: both include salt type \times year and salt type \times destination fixed effects. “Effective distance to source”

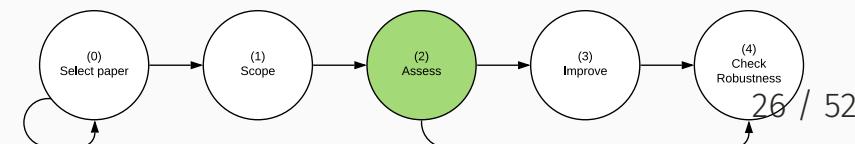


Assess

Table 2

└ TC estimation.do

- └ LCRED_RoX_CoY_RiZ.csv
 - └ TC_est_prep.m
 - └ railways_Dissolve_Simplify2_point2.csv
 - └ bd_ns_boundary2.csv
 - └ rivers_simplepoint2.csv
 - └ rivers_simplepoint2.csv
 - └ coast_simplepoint2.csv
 - └ od_salt_list.csv
 - └ prices_salt.dta
 - └ railways_Dissolve_Simplify2_point2.csv
 - └ bd_ns_boundary2.csv
 - └ rivers_simplepoint2.csv
 - └ coast_simplepoint2.csv
 - └ od_salt_list.csv
- └ RAIL_dummies.dta
- └ railways_D...point2.c
- └ rivers_simplepoint2.c



Assess

Analysis	
descriptives	T1_descriptives.do
income	income estimation.do income_est...temp.dta income_estimation.log
simulation	eq_cond2.m exp_shares.csv func_solve2.m land_area.csv secant_price2.m sim_prep_temp.dta T5_simulation.m
trade costs	haversine.m matlab_bgl_4.0.osxf TC estimation.do
trade flows	gravity estimation.dc gravity_estimation.lo income_temp.dta rainfall_temp.dta TF_est_prep.m
Data	crosswalks district-bl...wdistid.d district-bl...rrelation.i
income	income.dta
maps	bd_ns_boundary2.cs coast_simplepoint2.cs RAIL_dummies.dta railways_D...point2.cs rivers_simplepoint2.cs
Prices_salt	od_salt_list.csv prices_salt.dta
rainfall	crop rainfall.dta
trade flows	freight class.dta trade_data.dta trade_prices_1870.dta weight-value.dta
Readme for data files.pdf	

Table 2

- └ TC estimation.do
 - └ LCRED_RoX_CoY_RiZ.csv
 - └ TC_est_prep.m
 - └ railways_Dissolve_Simplify2_point2.csv
 - └ bd_ns_boundary2.csv
 - └ rivers_simplepoint2.csv
 - └ rivers_simplepoint2.csv
 - └ coast_simplepoint2.csv
 - └ od_salt_list.csv
 - └ prices_salt.dta
 - └ railways_Dissolve_Simplify2_point2.csv
 - └ bd_ns_boundary2.csv
 - └ rivers_simplepoint2.csv
 - └ coast_simplepoint2.csv
 - └ od_salt_list.csv

Table 3.5: Levels of Computational Reproducibility

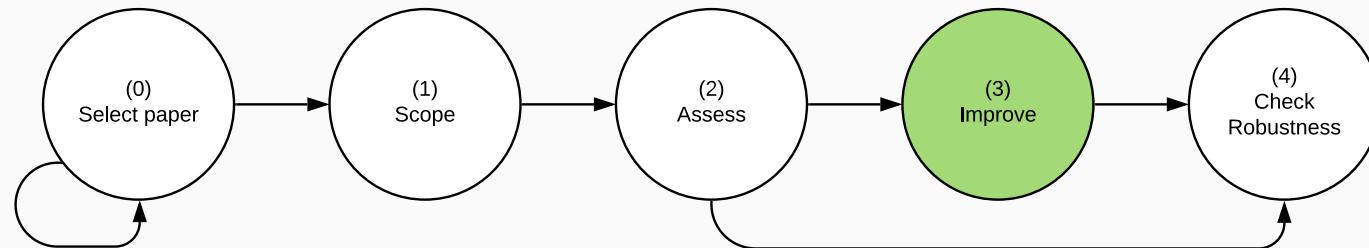
(P denotes “partial,” C denotes “complete”)

	Availability of materials, and reproducibility								CRR
	Analysis Code		Analysis Data		CRA	Cleaning Code		Raw Data	
	P	C	P	C		P	C	P	C
L1: No materials	-	-	-	-	-	-	-	-	-
L2: Only code	✓	✓	-	-	-	-	-	-	-
L3: Partial analysis data & code	✓	✓	✓	-	-	-	-	-	-
L4: All analysis data & code	✓	✓	✓	✓	-	-	-	-	-
L5: Reproducible from analysis	✓	✓	✓	✓	✓	-	-	-	-
L6: All cleaning code	✓	✓	✓	✓	-	✓	✓	-	-
L7: Some raw data	✓	✓	✓	✓	-	✓	✓	✓	-
L8: All raw data	✓	✓	✓	✓	-	✓	✓	✓	✓
L9: All raw data + CRA	✓	✓	✓	✓	✓	✓	✓	✓	-
L10: Reproducible from raw data	✓	✓	✓	✓	✓	✓	✓	✓	✓

a **Computationally Reproducible from Analytic data (CRA):** The output can be reproduced with minimal effort starting from the *analytic* datasets.

b **Computationally Reproducible from Raw data (CRR):** The output can be reproduced with minimal effort from the *raw* datasets.

Improvements

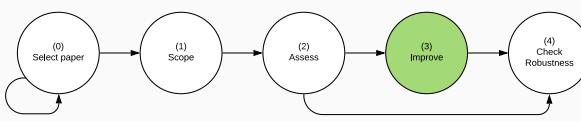


Three types of improvements:

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

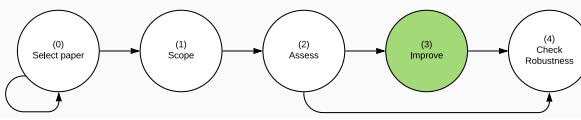
Improvements: Paper-level

- 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](#)).



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 row 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"



Improvements: future possible

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

Example 1:

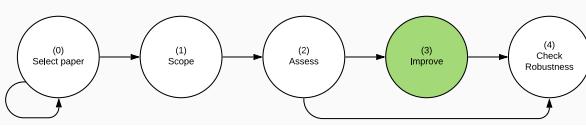
"Revise the .aml and .bat code scripts to reflect reorganized structure"

Example 2:

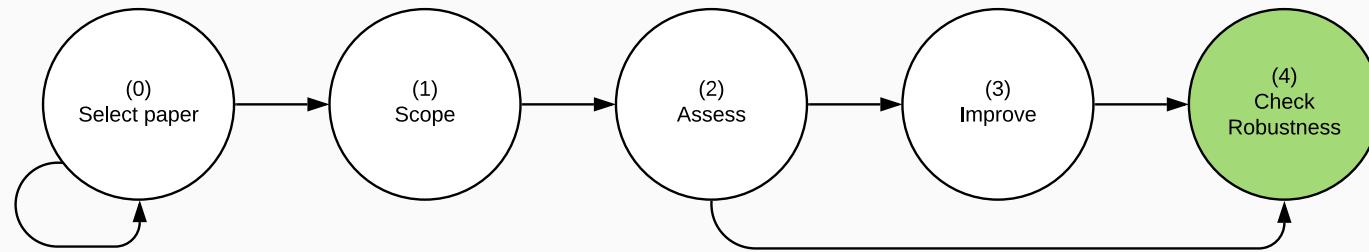
"Provide data and codes generating the other two figures in the paper, which are not given in the replication file."

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

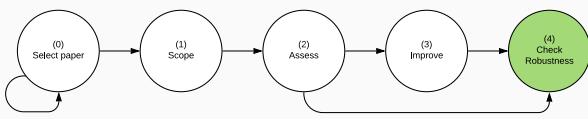
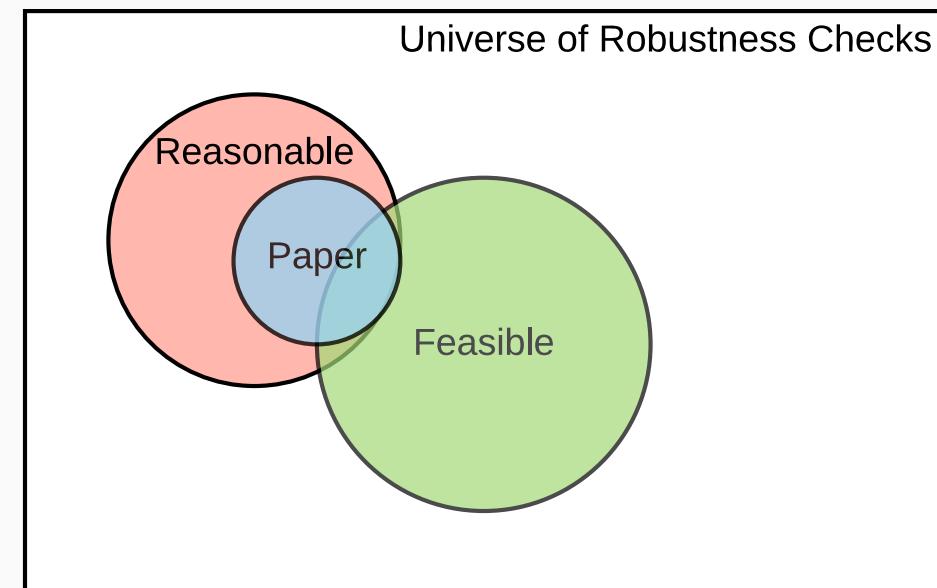
Robustness checks: any possible change in a computational choice, both in data analysis and data cleaning

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



Completed Reproduction: Example

Reproduction of: Railroads of the Raj: Estimating Transportation Infrastructure American Economic Review
<https://doi.org/10.1257/aer.20101199>

Reproducer: Emma Ng
Date created: November 19 2020
Date submitted: August 03 2021

Claims assessed: 1
Display Items (DI) assessed: 1
Improvements: 3

Robustness tests:
• Feasible choices added: 0
• Specifications justified: 0

Claims identified by reproducer:
1. "The paper estimated the value of the extent to which railroads reduced trade costs within colonial India using OLS and NLS methods. I estimate that the average reduction in trade costs from railroads is 10% of the total cost of trade. The reduction in LCRED (lowest-cost route effective distance) is 10% of the total trading cost. This estimate has a 95% confidence interval from 5% to 15%."

Reproducibility of Display Items:
1. Table 2 - Table 2 -- Railroads and Trade Costs: Step 1...

Original reproduction package: <https://www.aeaweb.org/articles?id=10.1257/aer.20101199>

Revised reproduction package: <https://github.com/em-ng21/railroads-of-the-raj>

Original authors' availability for further inquiries: Not sure

0 Select a paper

Declare the paper that you will analyze in the remainder of the exercise and other "candidate" papers for which you were unable to obtain access to a

[VIEW THIS SECTION](#)

1 Scoping

2 Assessment

3 Improvement

4 Robustness

0 Select a paper

Declare the paper that you will analyze in the remainder of the exercise and other "candidate" papers for which you were unable to obtain access to a

Completed Reproduction: Example

Reproduction of: Railroads of the Raj: Estimating Transportation Infrastructure American Economic Review
<https://doi.org/10.1257/aer.20101199>

Reproducer: Emma Ng
Date created: November 19 2020
Date submitted: August 03 2021

Claims assessed: 1
Display Items (DI) assessed: 1
Improvements: 3

Robustness tests:
• Feasible choices added: 0
• Specifications justified: 0

Claims identified by reproducer:

1. "The paper estimated the value of the extent to which railroads reduced transportation costs within colonial India using OLS and NLS methods. I have reproduced the estimates for the costs of different modes of transportation are estimated using LCRED (lowest-cost route effective distance) and the reduction of LCRED (lowest-cost route effective distance) from the original paper. This estimate has a 95% confidence interval of [100, 200] km."

Reproducibility of Display Items:

1. Table 2 - Table 2 -- Railroads and Trade Costs: Step 1...

Original reproduction package: <https://www.aeaweb.org/articles?id=10.1257/aer.20101199>

Revised reproduction package: <https://github.com/em-ng21/railroads-of-the-raj>

Original authors' availability for further inquiries: Not sure

0 Select a paper

Declare the paper that you will analyze in the remainder of the exercise and record other "candidate" papers for which you were unable to obtain access to a reproduction package.

0 Select a paper

Declare the paper that you will analyze in the remainder of the exercise and record other "candidate" papers for which you were unable to obtain access to a reproduction package.

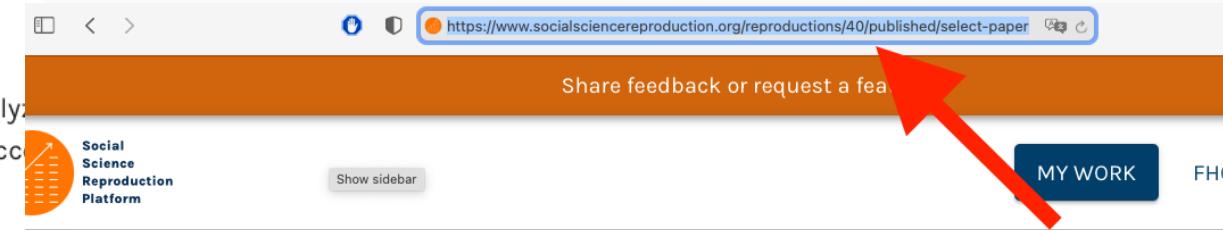
[VIEW THIS SECTION](#)

1 Scoping

2 Assessment

3 Improvement

4 Robustness



The screenshot shows the 'Select a paper' section of the platform. At the top, there's a header with the title 'Select a paper' and a sub-instruction 'Declare the paper that you will analyze in the remainder of the exercise and record other "candidate" papers for which you were unable to obtain access to a reproduction package. See detailed guidance [here](#)'. Below this is a large orange button labeled 'VIEW THIS SECTION'. To the left of the main content area, there are four numbered stages: 'Scoping', 'Assessment', 'Improvement', and 'Robustness', each with a small circular icon. On the right side of the main content area, there's a 'RETURN TO STAGES OVERVIEW' button and a 'Basic information' section. The top right corner of the page includes a 'MY WORK' button and a 'FHG' logo. The overall layout is clean and organized, designed for users to manage their research projects.

Select a paper

Declare the paper that you will analyze in the remainder of the exercise and record other "candidate" papers for which you were unable to obtain access to a reproduction package. See detailed guidance [here](#).

[RETURN TO STAGES OVERVIEW](#)

Basic information

- **Share it:** Original authors, instructors, other researchers, add it to your CV
- **Discuss it:** Discourse forum for reproductions
- **Cite it:** Reproduction DOI coming soon

Promoting a Constructive Exchange

- 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
 - 7 - Responding to hostile responses from original authors
- Additionally: sample responses form authors to reproducers

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!

Best regards,
[Reproducer]

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]

How Do I Get Started?

SIGN IN

About Events Guide Forum Metrics

Welcome to the Social Science Reproduction Platform

Accelerating computational reproducibility in the social sciences

GET STARTED

Search reproductions by paper title, DOI

Introduction

- Beyond binary judgments
- Reproduction stages
- Reproduction strategies

1 Selecting a paper

- 1.1 From candidate to declared p...
- 1.2 Identifying the relevant timeline
- 1.3 Potential sources of papers to...

2 Scoping

- 2.1 Read and summarize the paper
- 2.2 Record a revised reproduction...
- 2.3 Record scope of the exercise

3 Assessment

- 3.1 Describe the inputs
- 3.2 Connect display items to all it...
- 3.3 Assign a reproducibility score.

4 Improvements

- 4.1 Display item improvements
- 4.2 Paper-level improvements
- 4.3 Documenting the improvemen...

5 Checking for Robustness

- 5.1 Feasible robustness checks: i...
- 5.2 Justifying and testing reasona...

Guide for Accelerating Computational Reproducibility in the Social Sciences

BITSS BERKELEY INITIATIVE FOR TRANSPARENCY IN THE SOCIAL SCIENCES

2021-09-22

See a full list of contributors

A companion to the [Social Science Reproduction Platform](#).

(0) Select paper → (1) Scope → (2) Assess → (3) Improve → (4) Check Robustness

(0) Select (1) Scoping (2) Assessment (3) Improvement (4) Robustness

Display-

Use it for your class
or independent project

Consult the supporting guide

How Do I Get Started?

SIGN IN

About Events Guide Forum Metrics

Welcome to the Social Science Reproduction Platform

Accelerating computational reproducibility in the social sciences

GET STARTED

Search reproductions by paper title, DOI

Introduction

- Beyond binary judgments
- Reproduction stages
- Reproduction strategies

1 Selecting a paper

- 1.1 From candidate to declared p...
- 1.2 Identifying the relevant timeline
- 1.3 Potential sources of papers to...

2 Scoping

- 2.1 Read and summarize the paper
- 2.2 Record a revised reproduction...
- 2.3 Record scope of the exercise

3 Assessment

- 3.1 Describe the inputs
- 3.2 Connect display items to all it...
- 3.3 Assign a reproducibility score.

4 Improvements

- 4.1 Display item improvements
- 4.2 Paper-level improvements
- 4.3 Documenting the improvemen...

5 Checking for Robustness

- 5.1 Feasible robustness checks: i...
- 5.2 Justifying and testing reasona...

Guide for Accelerating Computational Reproducibility in the Social Sciences

BITSS BERKELEY INITIATIVE FOR TRANSPARENCY IN THE SOCIAL SCIENCES

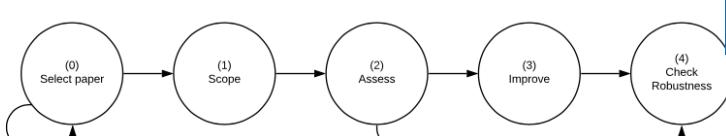
2021-09-22

See a full list of contributors

A companion to the Social Science Reproduction Platform.

(0) Select paper → (1) Scope → (2) Assess → (3) Improve → (4) Check Robustness

(0) Select (1) Scoping (2) Assessment (3) Improvement (4) Robustness



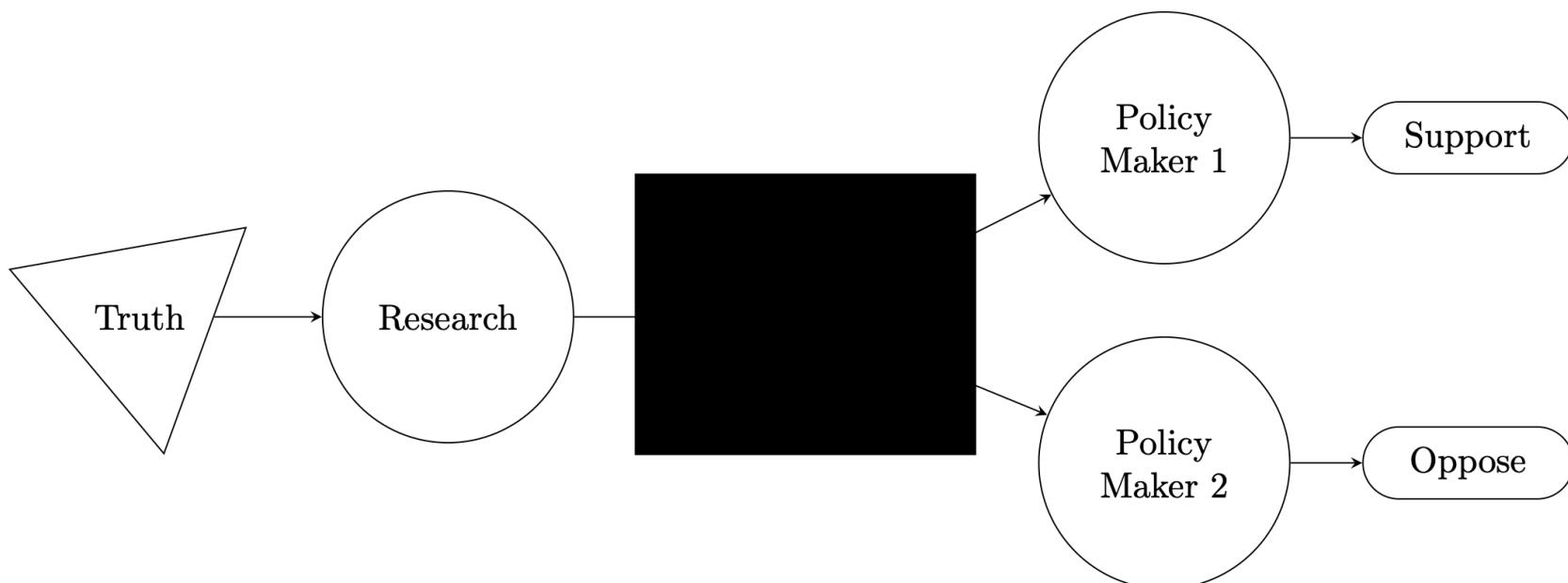
Use it for your class or independent project



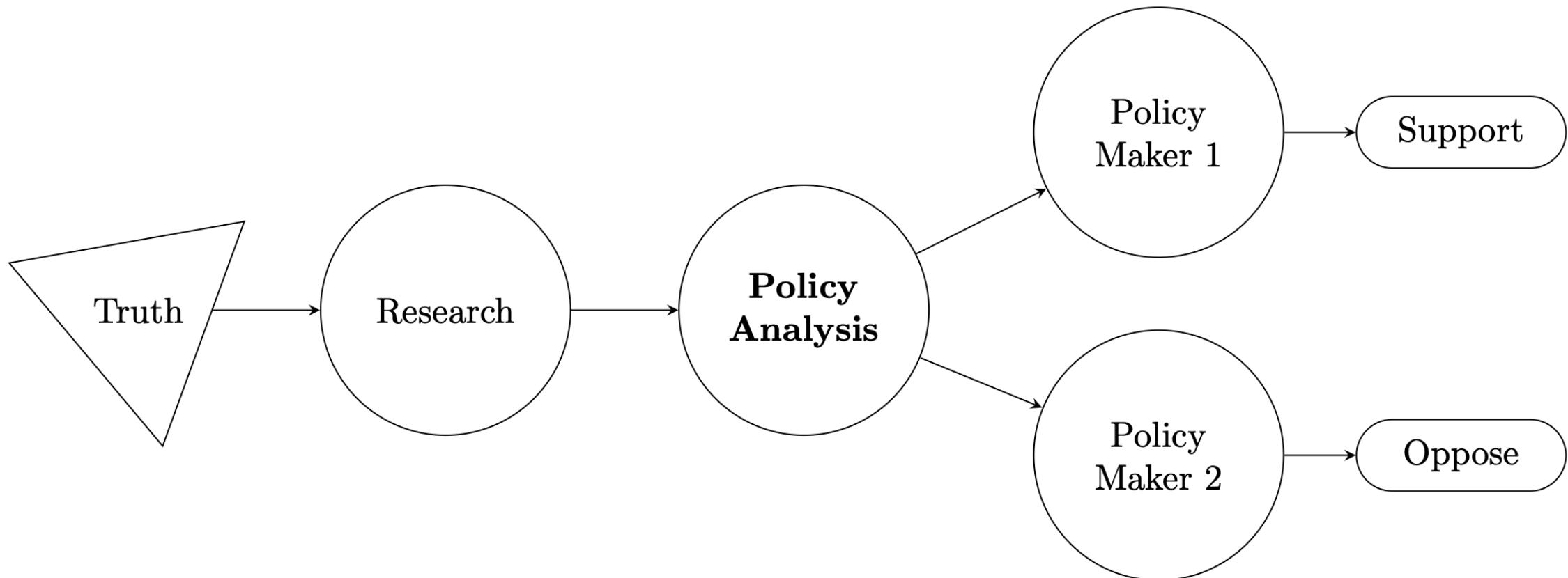
Ask questions

Today's Presentation

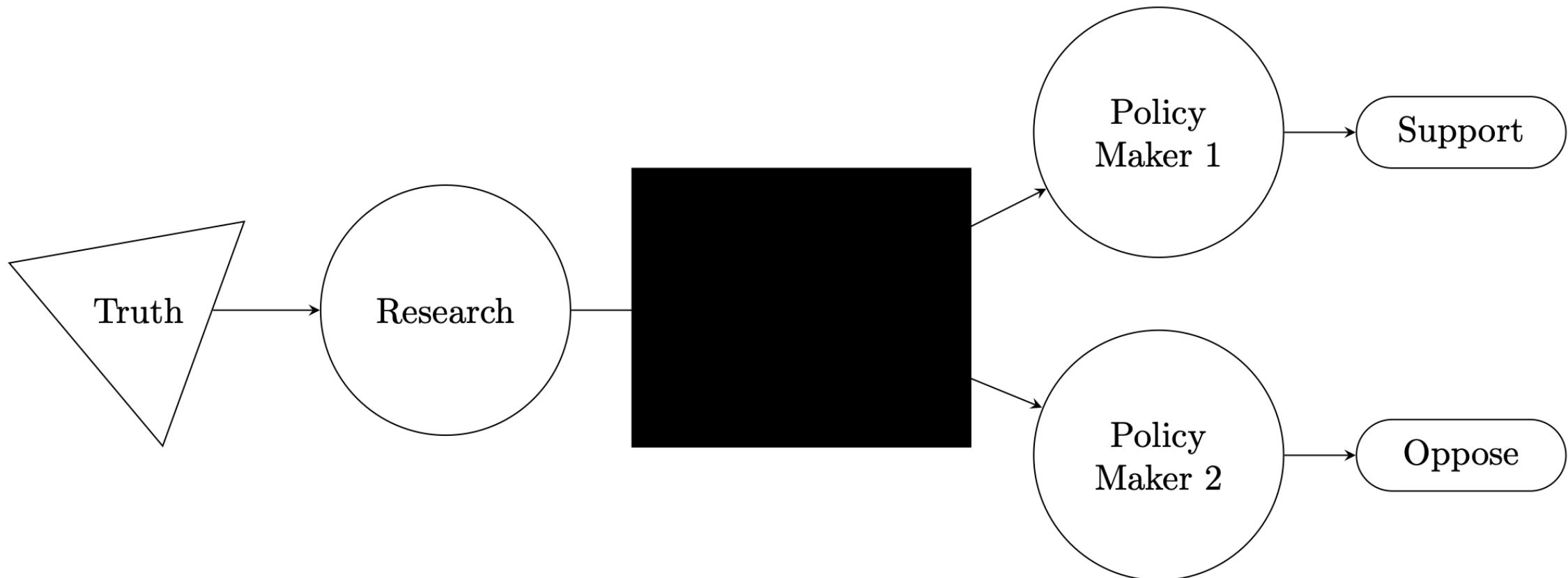
- Part I: Transparency in Research
- **Part II: Transparency in Policy Analysis**



How This Affects The Evidence Based Policy Link?



How This Affects The Evidence Based Policy Link?



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 output, analysis and materials
 - Agenda to implement and document how OPA can be applied into several policy issues
- Here we will review the application of OPA to mass deworming interventions

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

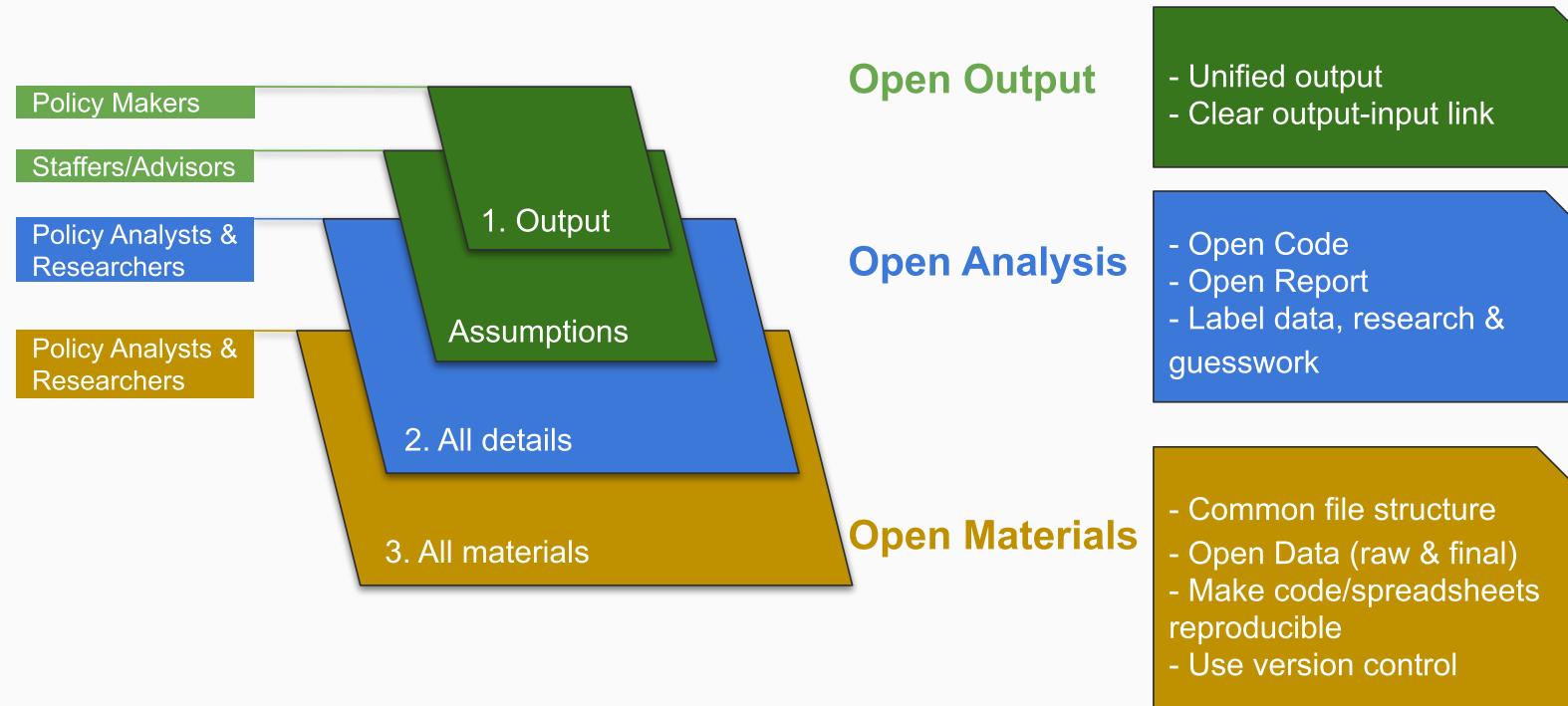
- Context of original study (Kenya, 1998-99) had very high prevalence rates of worm infections
- Implementation costs were very low (\$0.42 per round of treatment)
- Length of treatment was relatively short (2.4 years)
- Current deworming settings have lower prevalence rates, varying implementation costs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

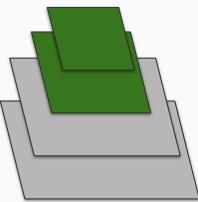


Contributions of OPA to deworming:

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

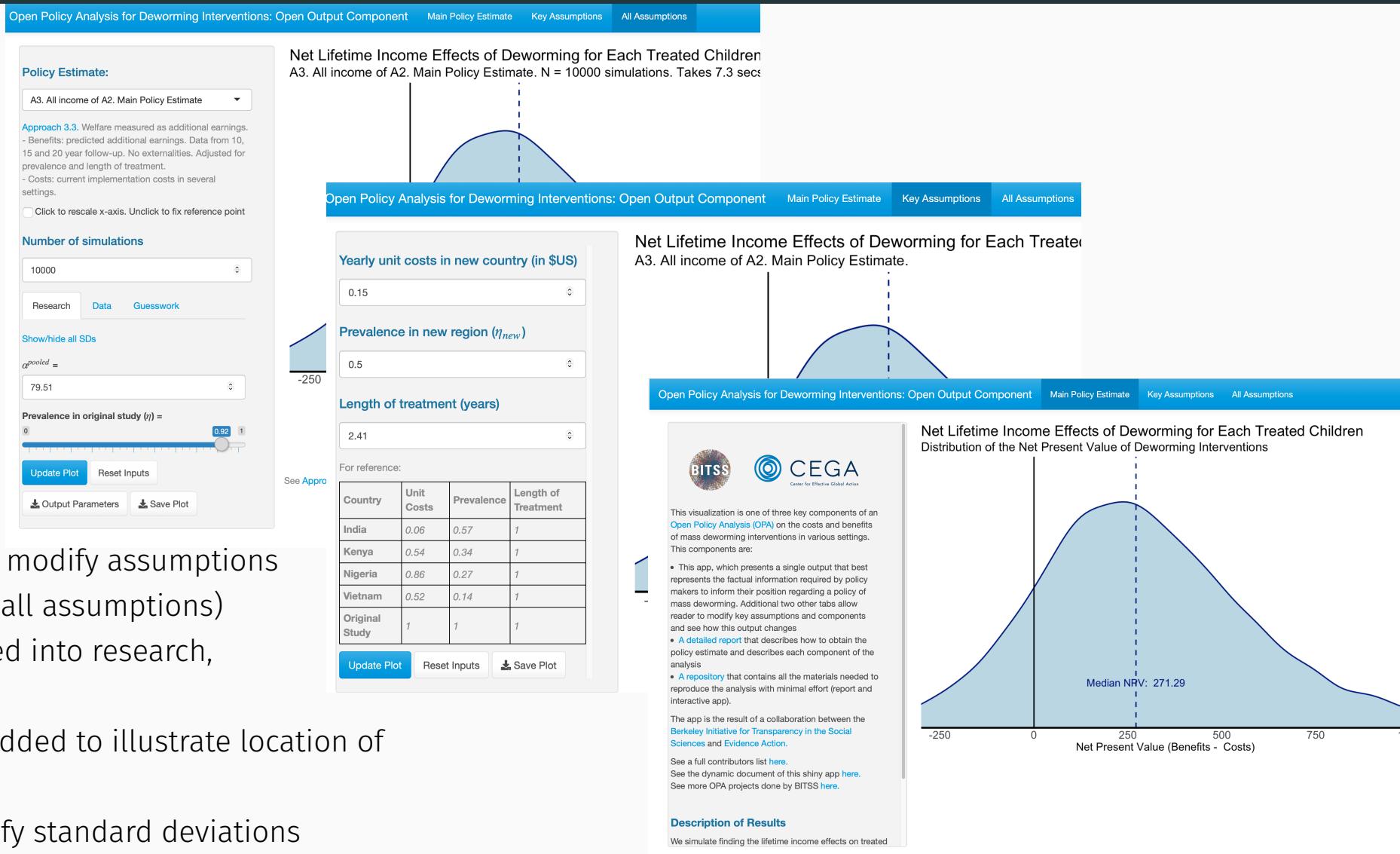
Open Output

Demo



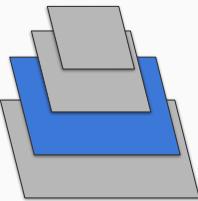
Main features

- 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

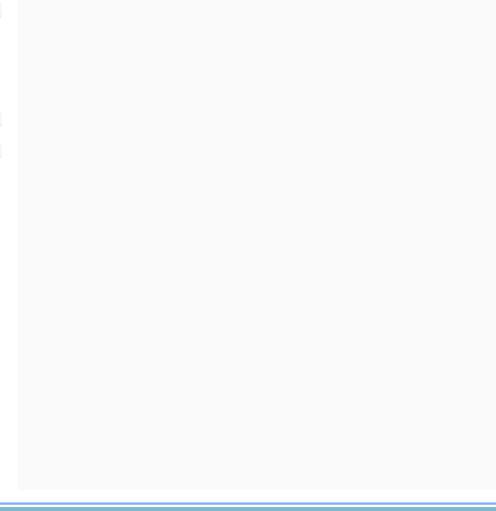
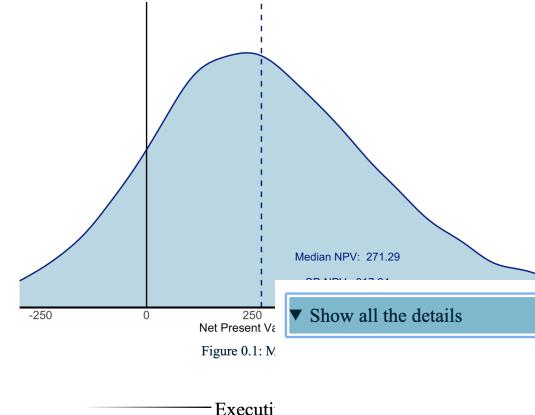
- Complete narrative description of the methodology
- Translation of each narrative step into an equation
- Implementation of each equation into code
- Combine all of the above into using a dynamic document (RMarkdown)
- Presentation of narrative, equations, and code in layered fashion to avoid overwhelming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



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

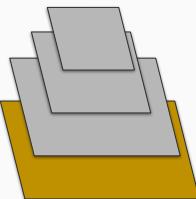
Where:

- 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

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/bitsso_logo_horizontal.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 bibliography: 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.)

Environment History Connections

Import Dataset Global Environment

Files Plots Packages Help Viewer

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

49 / 52

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).

Thank You

fhocea@berkeley.edu