



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Transparencia en Investigación en las Ciencias Sociales

Fernando Hoces de la Guardia¹

¹UC Berkeley:
Berkeley Initiative for Transparency in the Social Sciences

MIES, Octubre 2018
Slides disponibles en

<https://tinyurl.com/yca6bzw2>

Estructura de la Presentación

Transparencia
en
Investigación

Introducción
Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Introducción**
- 2 Ética en la Investigación Científica**
- 3 Sesgo de Publicación — Registros**
- 4 P-Hacking — Planes de Pre-Análisis**
- 5 Replicaciones Fallidas — Pautas, Protocolos y Herramientas**
- 6 Lecciones Para el Análisis de Políticas Públicas**
- 7 Conclusión**



Outline

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Introducción**
- 2 Ética en la Investigación Científica**
- 3 Sesgo de Publicación — Registros**
- 4 P-Hacking — Planes de Pre-Análisis**
- 5 Replicaciones Fallidas — Pautas, Protocolos y Herramientas**
- 6 Lecciones Para el Análisis de Políticas Públicas**
- 7 Conclusión**



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES



ICPSR



Outline

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Introducción**
- 2 Ética en la Investigación Científica**
- 3 Sesgo de Publicación — Registros**
- 4 P-Hacking — Planes de Pre-Análisis**
- 5 Replicaciones Fallidas — Pautas, Protocolos y Herramientas**
- 6 Lecciones Para el Análisis de Políticas Públicas**
- 7 Conclusión**

Ética en la Investigación Científica

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- Transparencia es un elemento central de la ética del investigador.
- Valores científicos acuñados por Robert Merton (Merton 1942):
 - **Universalismo:** cualquier persona puede presentar un argumento, independiente de su estatus.
 - **Comunismo:** el conocimiento es compartido de manera abierta.
 - **Desinterés:** la verdad como motivación, y no los beneficios monetarios.
 - **Escepticismo Organizado:** revisión a través de pares (peer review), replicación.

Ética en la Investigación Científica

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- Casos de fraude existen (Simonsohn 2013), pero más importante como investigadores tenemos que admitir nuestra condición humana, sujeto a sesgos y razonamiento motivado, transparencia puede ayudar con esto (Nosek, Spies, Motyl 2012).
- Quienes llevamos a cabo experimentos o usamos datos con información identificable a nivel individual, tenemos que tomar con seriedad los Comités de Ética Institucionales (IRBs) (Ch. 11–13 Morton & Williams 2010, Desposato 2014).

Ética en la Investigación Científica

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Por qué nos preocupamos:

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

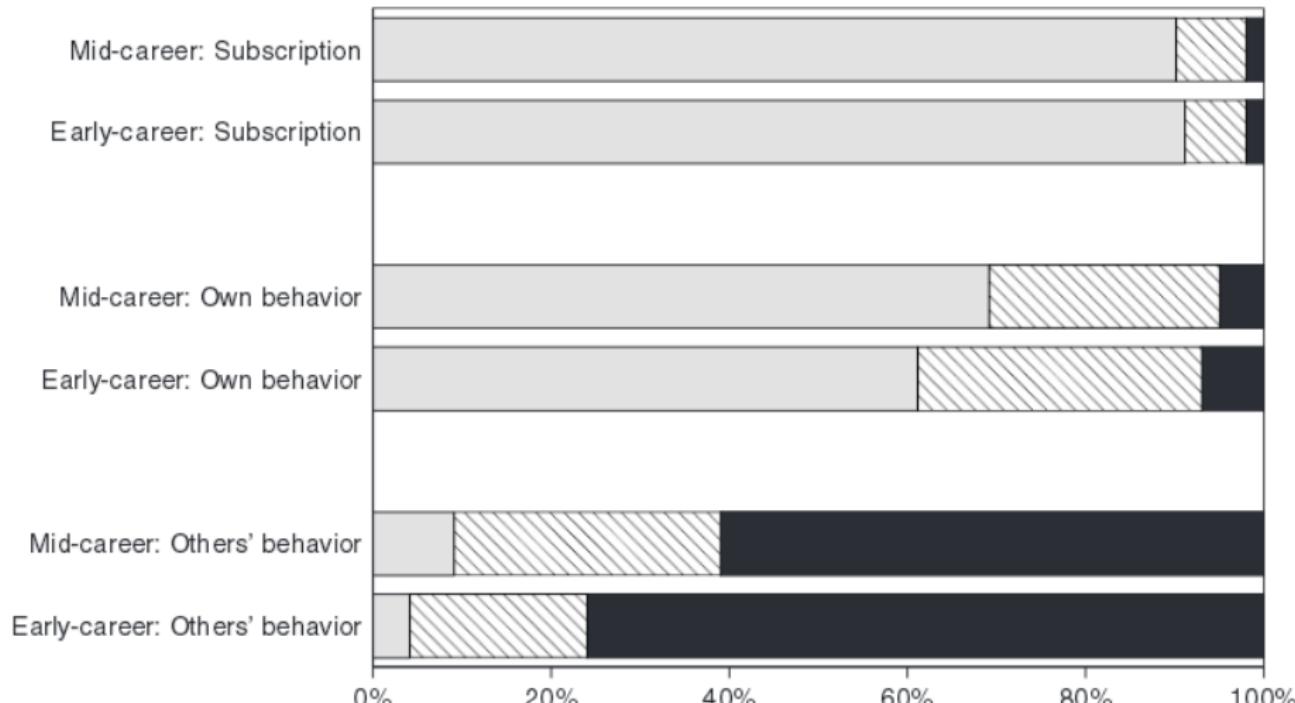


FIG. 3. Norm versus Counternorm Scores: Percent with Norm > Counternorm (dotted), Norm = Counternorm (striped), Norm < Counternorm (solid).

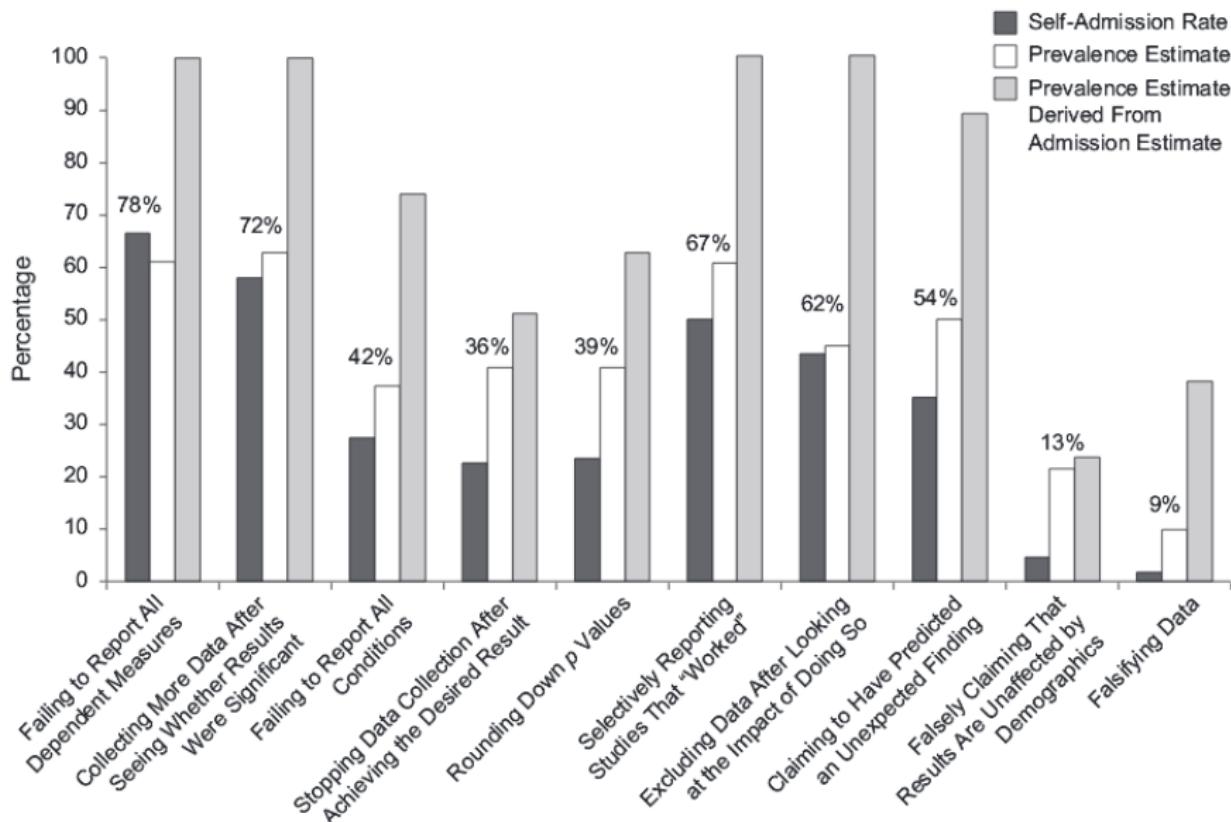


Fig. 1. Results of the Bayesian-truth-serum condition in the main study. For each of the 10 items, the graph shows the self-admission rate, prevalence estimate, prevalence estimate derived from the admission estimate (i.e., self-admission rate/admission estimate), and geometric mean of these three percentages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



Outline

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Introducción**
- 2 Ética en la Investigación Científica**
- 3 Sesgo de Publicación — Registros**
- 4 P-Hacking — Planes de Pre-Análisis**
- 5 Replicaciones Fallidas — Pautas, Protocolos y Herramientas**
- 6 Lecciones Para el Análisis de Políticas Públicas**
- 7 Conclusión**

Sesgo de Publicación

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Existencia del problema:

- El tamaño de los efectos disminuye con el tamaño muestral (Gerber, Green, Nickerson 2001)
- Las Ciencias Sociales muestran una tasa de rechazo de la hipótesis nula mayor que las ciencias duras (Fanelli 2010).
- La publicación de efectos nulos esta desapareciendo en el tiempo, en todas las disciplinas. (Fanelli 2011).
- Estudio que siguió a experimentos completados muestra que aquellos experimentos con fuertes resultados son 40pp más probable de ser publicados, y 60pp más probable de ser escritos. Alto “file drawer problem”. (Franco, Malhotra, Simonovits 2014)

Problema en todas las disciplinas

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

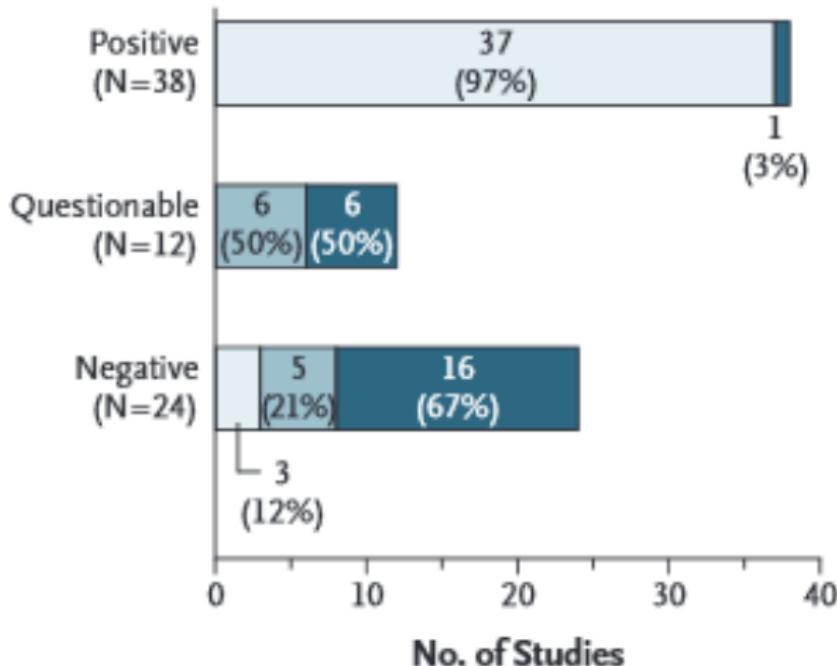
Conclusión

- Medicina: (Turner et al. 2008)
- Ciencias Sociales: (Franco, Malhotra, Simonovits 2014)
- Economía: (Brodeur et al. 2016)
- Sociología: (Gerber and Malhotra 2008)
- Ciencias Políticas: (Gerber and Malhotra 2008)

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

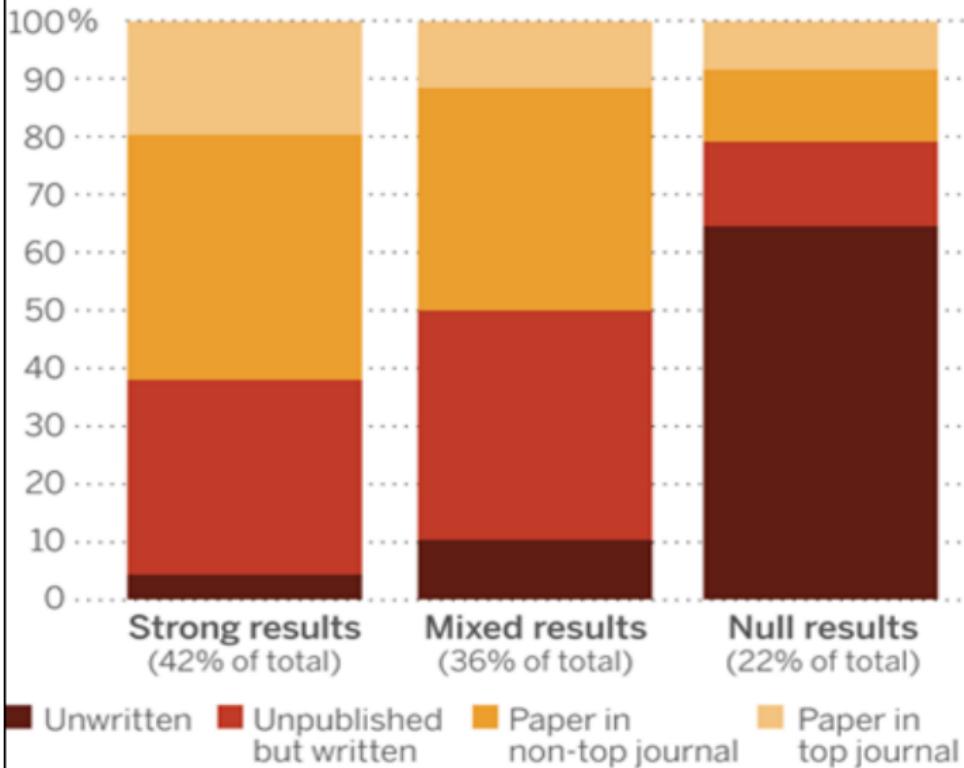
A Studies (N=74)

FDA Decision



Most null results are never written up

The fate of 221 social science experiments



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



Sesgo de Publicación

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Si solo escribimos/publicamos resultados significativos, y no dejamos registro de los no significativos, no tenemos forma de distinguir si nuestros resultados “significativos” son reales, o si son el 5% que deberíamos esperar debido a error estadístico.

Registros

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Pre-Registros como una solución al sesgo de publicación:

- Hacer pública la investigación a ejecutar, publicando por adelantado las hipótesis a testear.
- Adopción casi universal en RCTs en medicina.
Journals top (ICMJE) no publican estudios si no están registrados. <http://clinicaltrials.gov>

Registros

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Nuevos en ciencias sociales, pero:

- Registro de AEA, actualmente solo para RCTs.
<http://socialscienceregistry.org>
- Registro de EGAP
<http://egap.org/design-registration>
- Registro de 3ie, para evaluaciones en países en desarrollo. <http://ridie.3ieimpact.org>
- Open Science Framework
 - Formato abierto
 - Pronto va a estar sincronizado con los de más arriba
- Simple: <http://aspredicted.org>

Meta-Análisis

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Síntesis sistemático de los resultados

Organizaciones:

- Cochrane Collaboration (Medicina)
- Campbell Collaboration (Políticas públicas)
- What Works Clearinghouse (Educación, Gob. US)
- CLEAR (Empleo, Gob. US)
- MAER-NET (Economía)

Meta-Análisis

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Herramientas:

- Funnel plots del tamaño de la muestra vs. tamaño del efecto (Card & Krueger 1995)
- Funnel Asymmetry Test (Stanley & Doucouliagos 2012)
- P-curve (Simonsohn et al. 2014)
 - ▶ Online App
 - Un p-checker para todos ▶ Shiny App

Relevante? Resultados nulos aumentan despues de requerir registros (Kaplan & Irvin 2015)

Transparencia en Investigación

Introducción

Ética en la Investigación Científica

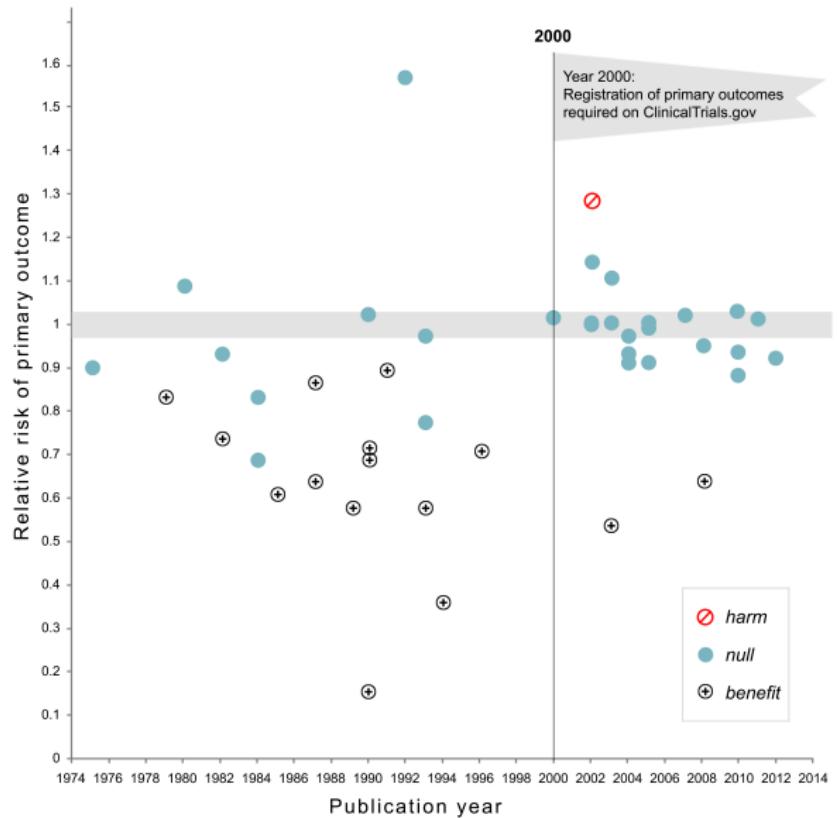
Sesgo de Publicación — Registros

P-Hacking — Planes de Pre-Análisis

Replicaciones Fallidas

Lecciones Para el Análisis de Políticas Públicas

Conclusión





Outline

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Introducción**
- 2 Ética en la Investigación Científica**
- 3 Sesgo de Publicación — Registros**
- 4 P-Hacking — Planes de Pre-Análisis**
- 5 Replicaciones Fallidas — Pautas, Protocolos y Herramientas**
- 6 Lecciones Para el Análisis de Políticas Públicas**
- 7 Conclusión**

P-Hacking

Transparencia
en
Investigación

Introducción
Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Definición del problema:

- Otros nombres: “data-fishing”, grados de libertad del investigador, o “data-mining”.
- Definición: flexibilidad en el análisis de datos permite presentar *casi cualquier resultado* bajo un umbral arbitrario; significancia estadística pierde sentido.
- No es algo único de investigadores con malas intenciones. Es subconsciente, o simplemente una práctica estándar del análisis estadístico (Gelman, Loken 2013).

P-Hacking: La Extension Del Problema (Athey 2018)

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

What We Say v. What We Do (Econometrics)

What We Say

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

What we do

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

P-Hacking: La Practica (BuzzFeed 2018!)

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

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

Hi David,

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

Brief -- American Journal of Public Health

Research Letter – Archives of Pediatric and Adolescent Medicine

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

Best,

Brian



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

P-Hacking: hágalo usted mismo!

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

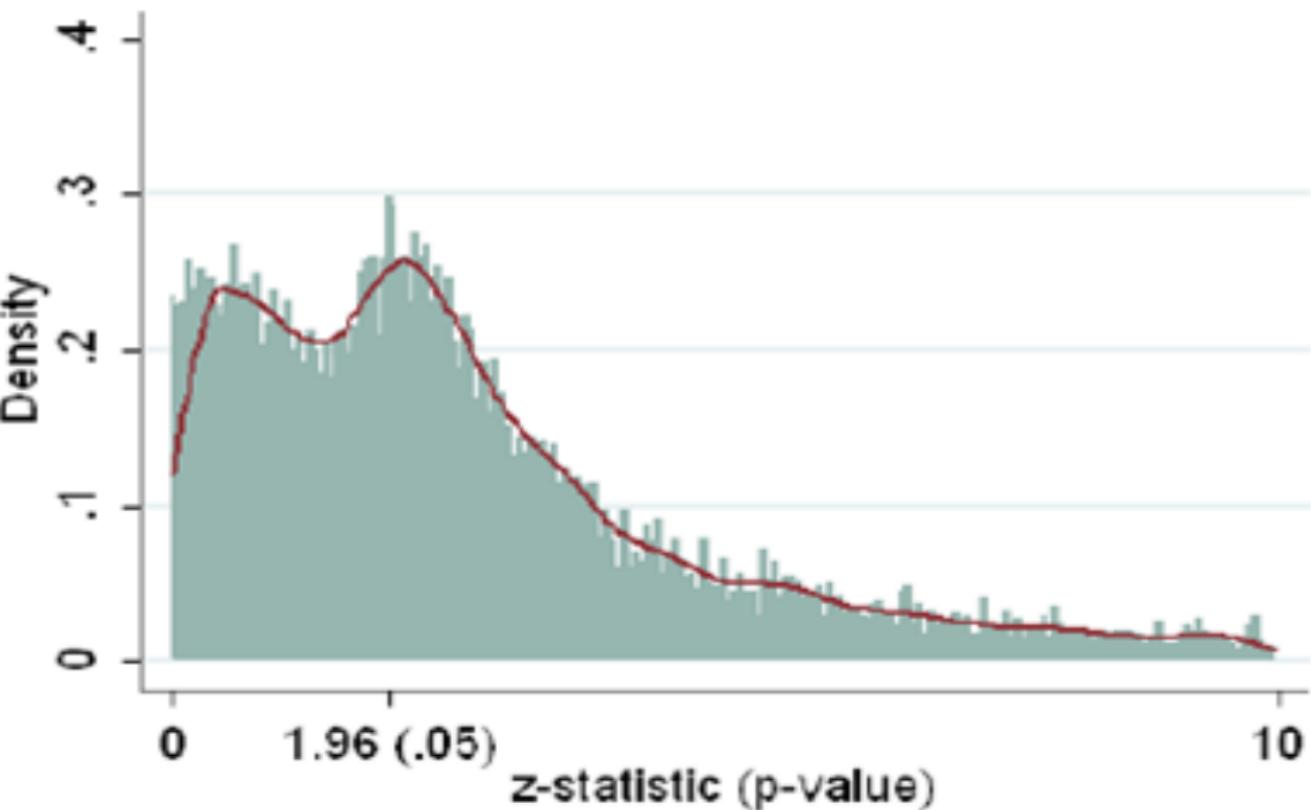
Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

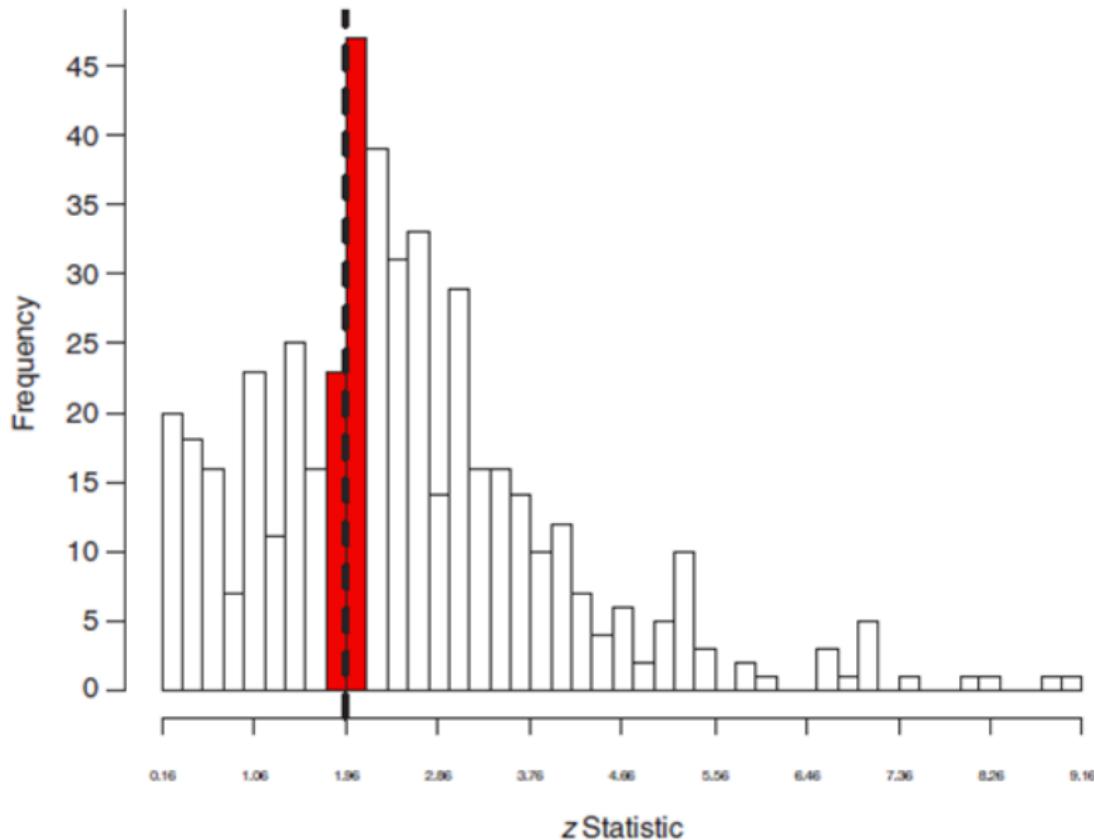
Conclusión

- “Science isn’t Broken” —538 reportaje periodístico con modulo interactivo [Link](#)
- Practique p-hacking en R/Shiny App. [Link](#)
- El Exact Fishy Test [Link](#)

(b) Unrounded distribution of z-statistics.



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



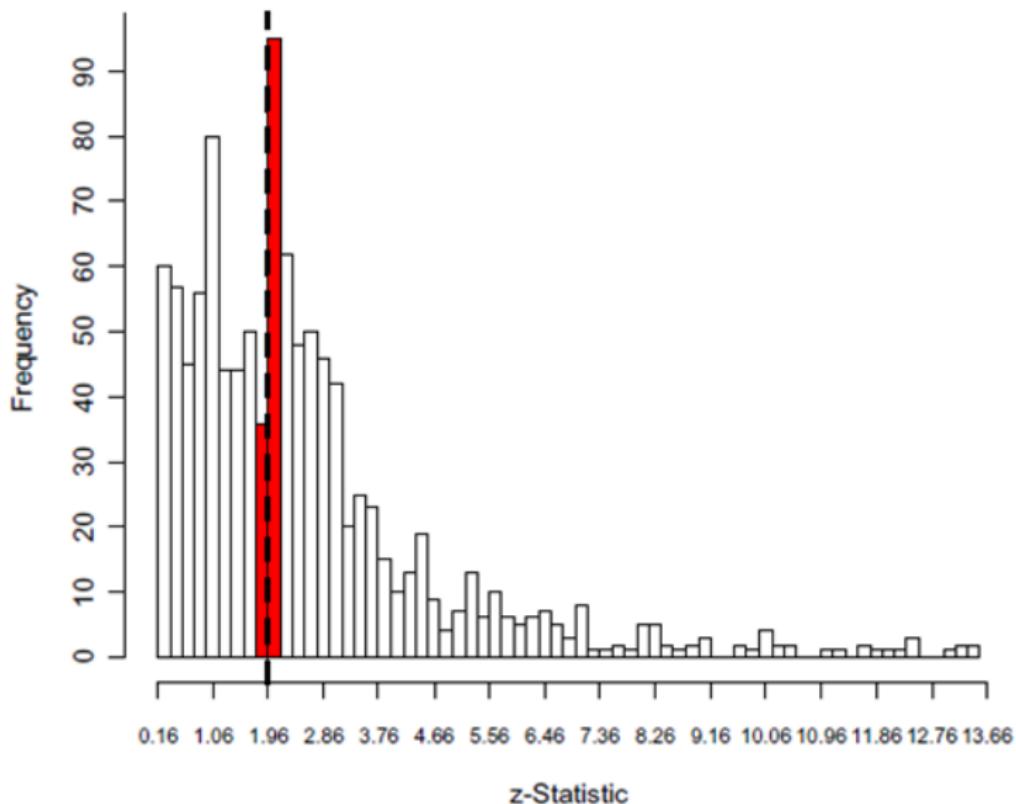


Figure 1(a). Histogram of z -statistics, *APSR* & *AJPS* (Two-Tailed). Width of bars (0.20) approximately represents 10% caliper. Dotted line represents critical z -statistic (1.96) associated with $p = 0.05$ significance level for one-tailed tests.

ONE DATA SET, MANY ANALYSTS

Twenty-nine research teams reached a wide variety of conclusions using different methods on the same data set to answer the same question (about football players' skin colour and red cards).

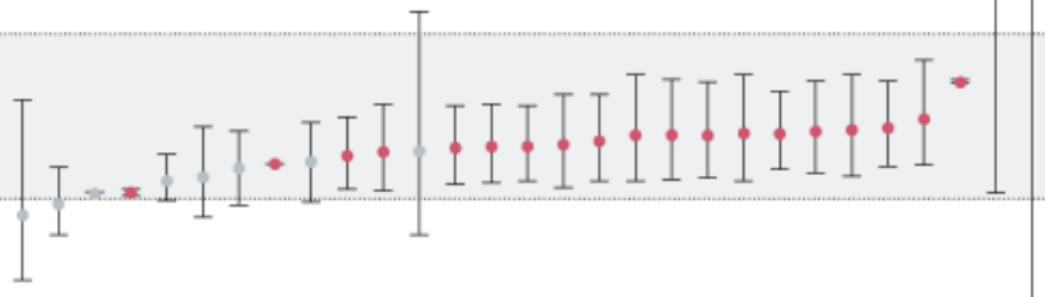
Dark-skinned players four times more likely than light-skinned players to be given a red card.

- Statistically significant effect
- Non-significant effect

Twice as likely

Equally likely

78.7*
115*



Point estimates and 95% confidence intervals. *Truncated upper bounds.



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Planes de Pre-Análisis

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Explicación de la solución, de 3ie:

“Un plan de pre-análisis (PAPs) es una **descripción detallada de los análisis a ser conducidos**. Esta descripción es escrita antes de ver los datos que contienen los impactos del programa bajo evaluación. Puede especificar las **hipótesis a testear, como se construirán las variables, ecuaciones a estimar, controles a utilizar, y otros aspectos del análisis**. Una función clave de los PAPs es incrementar la transparencia en investigación. Al **especificar los detalles por adelantado antes de ver los resultados**, el plan es un resguardo contra data mining y búsqueda de especificaciones. Se les recomienda a los investigadores que desarrollen y suban dichos planes junto con el registro de sus estudios. Sin embargo, esto no es requisito para el registro de un estudio.”

Orígenes: Regulación Farmacéutica en USA

Transparencia
en
Investigación

Introducción
Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

“E9 Statistical Principles for Clinical Trials” (1998) [▶ Link](#)

§V Consideraciones en el Análisis de Datos

- 1 Pre-especificación del análisis
- 2 Grupos de análisis
- 3 Tratamiento de NAs y valores extremos
- 4 Transformaciones a los datos
- 5 Estimación, intervalos de confianza, y testeo de hipótesis
- 6 Ajuste de significancia y niveles de confianza
- 7 Subgrupos, interacciones, y variables de control
- 8 Integridad de los datos y validez del software computacional

Sugerencias de Glennerster, Takavarasha

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Running Randomized Evaluations

- 1 Como se van a medir los principales variables de resultado,
- 2 Dentro de las variables de resultado: cuales son primarias, cuales secundarias?
- 3 Composición exacta de toda la familia de tests utilizados en análisis de efectos promedios
 - Explicación de efectos promedio, FWER, FDR en Anderson (JASA 2008).
- 4 Los subgrupos a ser analizados,
- 5 La dirección esperada de los impactos si queremos usar test de una cola, y
- 6 La especificación primaria a ser utilizada en el análisis.

Sugerencias de McKenzie

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

World Bank Development Impact Blog

- 1 Descripción de la muestra a ser utilizada en el estudio
- 2 Fuentes de datos fundamentales
- 3 Hipótesis a testear a través de la cadena causal
- 4 Especificar como se van a construir las variables
- 5 Especificar la ecuación donde se estimará el efecto del tratamiento
- 6 ¿Cuál es la estrategia para analizar múltiples variables de resultados y múltiples test de hipótesis?
- 7 Procedimiento a utilizar para enfrentar atrición en la muestra
- 8 ¿Qué hará el estudio en caso de limitada variación en las variables de resultados?
- 9 Si hay un modelo a testear, este debe ser incluido
- 10 No olvidar archivarlo (con una fecha estampada digitalmente)

Ejemplos

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- J-PAL Registro de Hipótesis , ver <http://www.povertyactionlab.org/Hypothesis-Registry>
6 ejemplos de papers publicados:
 - Sierra Leone CDD, Oregon Medicare, Turkey Job Training, El Salvador TOMS, dos en Indonesia (Olken et al.)
 - Psicología: Hawkins, Fitzgerald, Nosek—Riesgos de Concepción y Prejuicios

Amplio rango de que es lo que se debe escribir exactamente y cuanto detalle incorporar en el PAP. En un nivel extremo el código estaría listo para ejecutar en cuanto lleguen los datos.

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

PAP-Estudios Observacionales

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

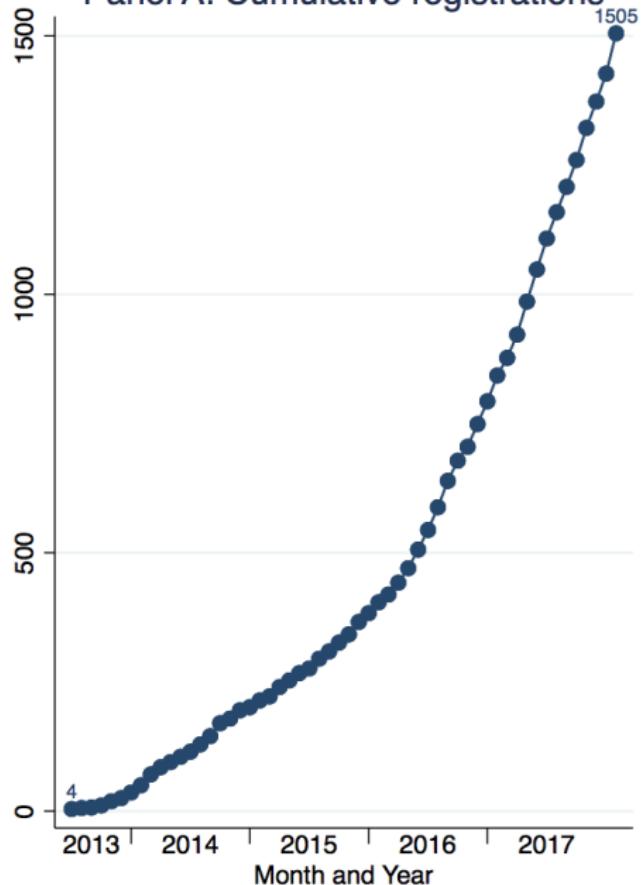
Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

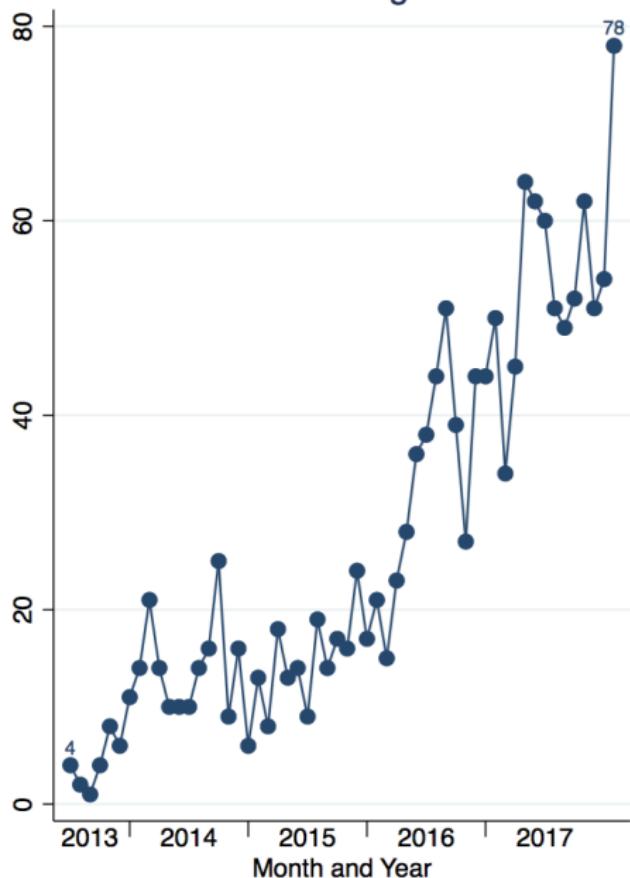
Conclusión

- Actual debate en salud pública/epidemiología.
- Pre-especificar de manera verificable: difícil, pero no imposible.
- Ejemplo: Datos generados periódicamente por el gobierno.
- Ejemplo: Salario Mínimo (Neumark 2001)

Panel A: Cumulative registrations



Panel B: New registrations



Publicaciones Basadas en Diseño de la Investigación

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

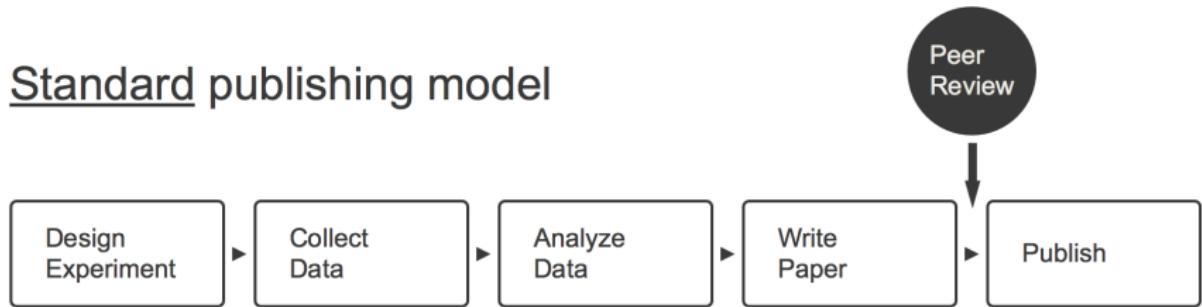
Conclusión

AKA Reportes Registrados, cambia momento de peer review hacia antes del la recolección de datos, análisis y resultados.

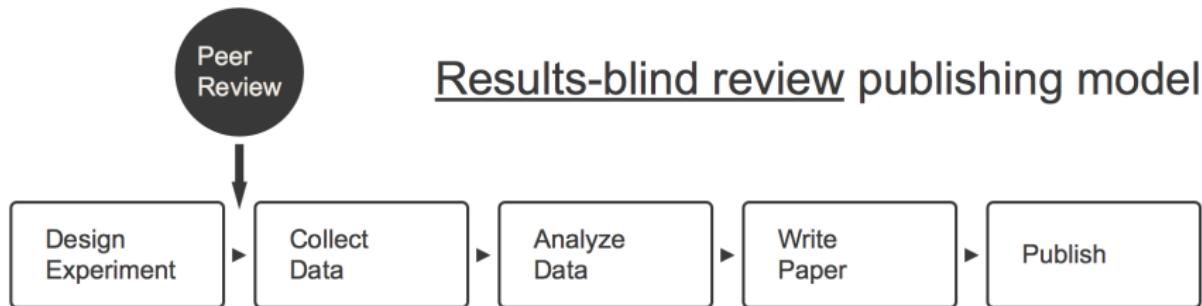
- 1 Diseñar un estudio
- 2 Enviar a un journal
- 3 Revisión basada en la importancia de la pregunta y calidad del diseño
- 4 Obtener aceptación en principio
- 5 Ejecutar el estudio, y publicar incluso con resultados nulos

+100 Journals [▶ Link](#)

Standard publishing model



Results-blind review publishing model





Register Reports at the JDE.

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

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



ISSN: 0304-3878

[Submit Your Paper](#)

[View Articles](#)

[Guide for Authors](#)

[Abstracting/ Indexing](#)

[Track Your Paper](#)

[Order Journal](#)

[Journal Metrics](#)

CiteScore: 2.96

More about CiteScore

Impact Factor: 1.889

5-Year Impact Factor: 3.305

Source Normalized Impact per Paper
(SNIP): 2.732

Journal of Development Economics

Supports Open Access

Editor in Chief: A. Foster

[View Editorial Board](#)

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

[Read more](#)

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

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

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

Informal finance: A theory of moneylenders Andreas Madestam

[View All Most Downloaded Articles](#)

Announcements

NEW: Registered Reports

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

Guidelines and Checklist

Outline

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Introducción**
- 2 Ética en la Investigación Científica**
- 3 Sesgo de Publicación — Registros**
- 4 P-Hacking — Planes de Pre-Análisis**
- 5 Replicaciones Fallidas — Pautas, Protocolos y Herramientas**
- 6 Lecciones Para el Análisis de Políticas Públicas**
- 7 Conclusión**

Replicacion

Transparencia
en
Investigación

Introducción
Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

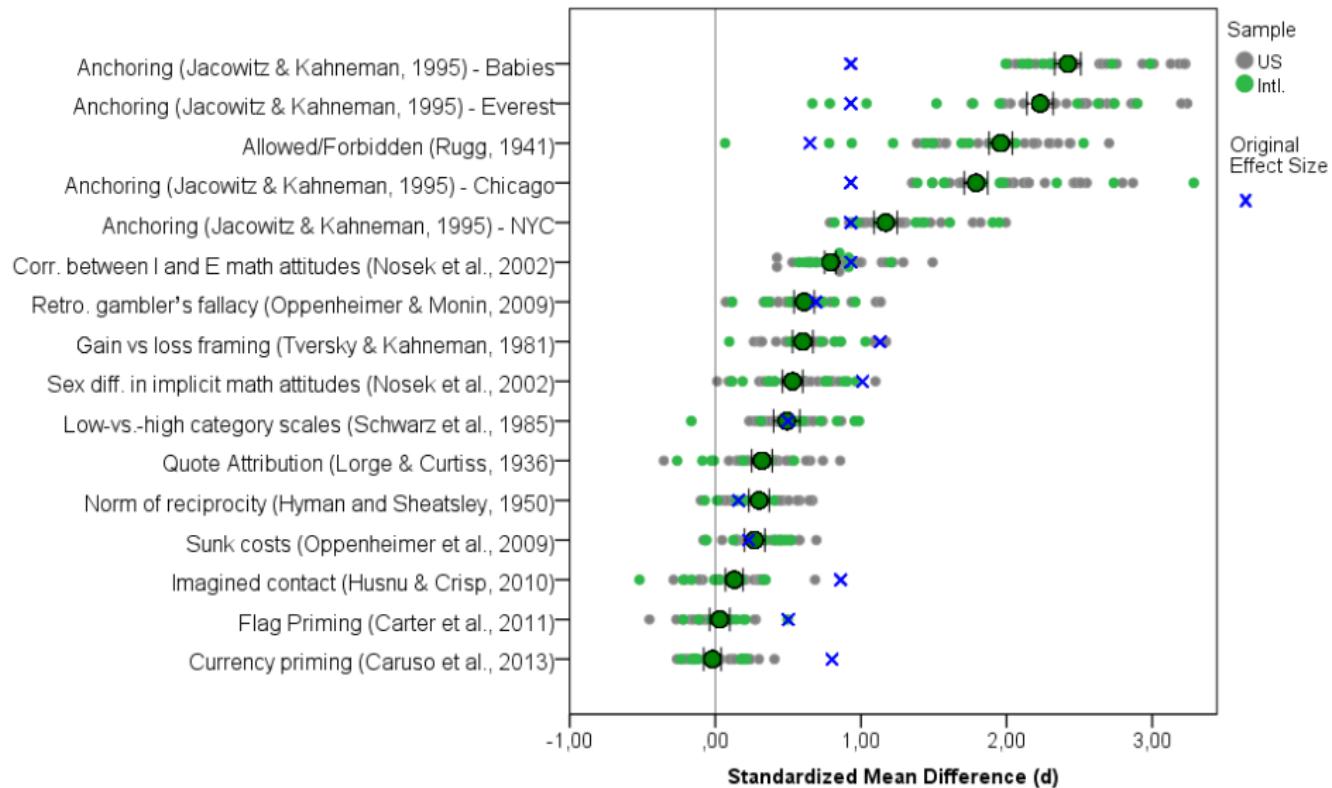
P-Hacking —
Planes de
Pre-Análisis

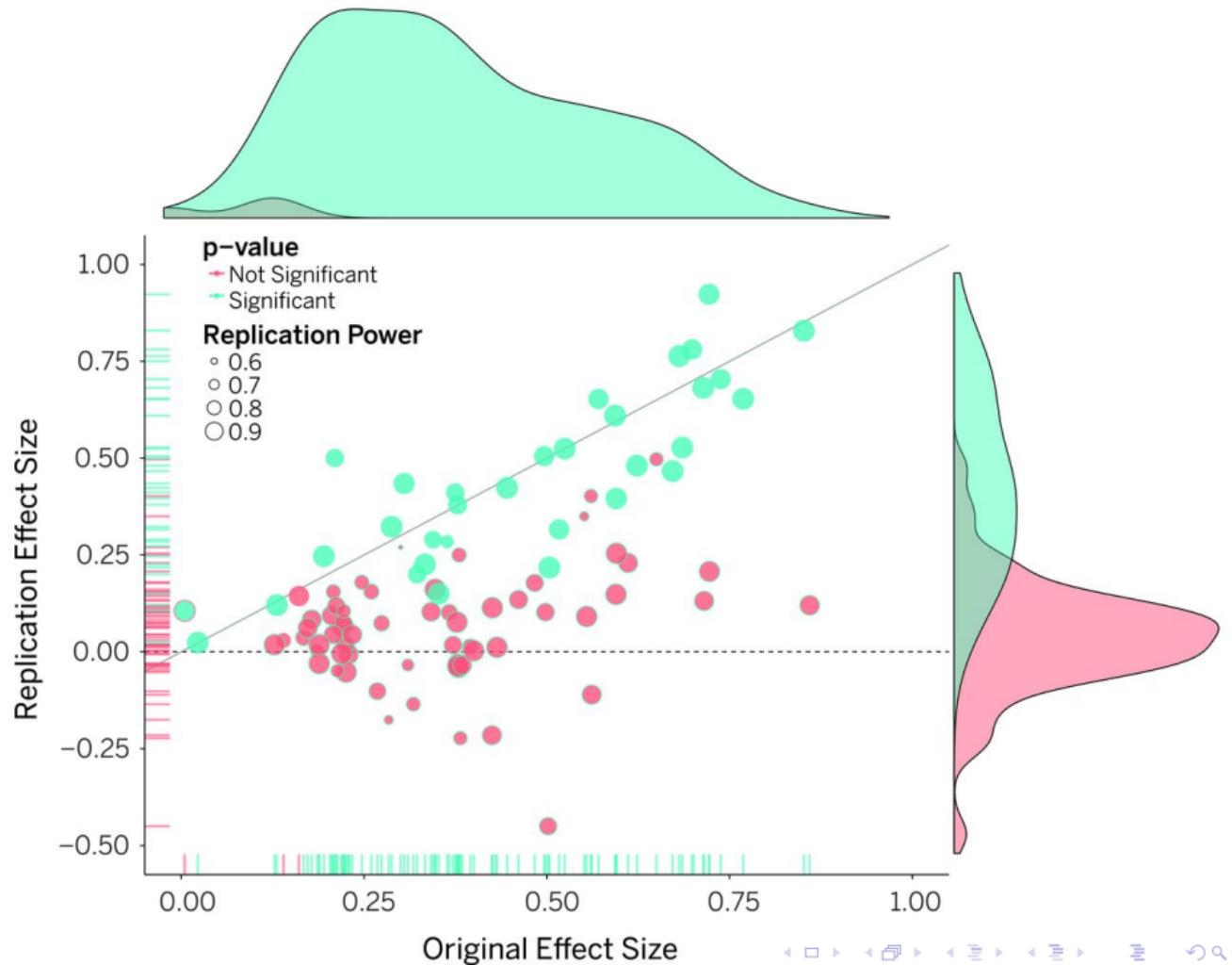
Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Baja replicabilidad de experimentos:** Camerer et al (2016, 2018), Many Labs, Reproducibility Project: Psychology (2015)
- 2 Baja Reproducibilidad Computacional:** JMCCB Project (1986); Chang & Li (2015); Gertler, Galiani, Romero (2018)
- 3 Conjunto de soluciones:**
 - Pautas TOP
 - Flujo de trabajo reproducible
 - Compartir código y datos
 - Guias de estándares de publicación



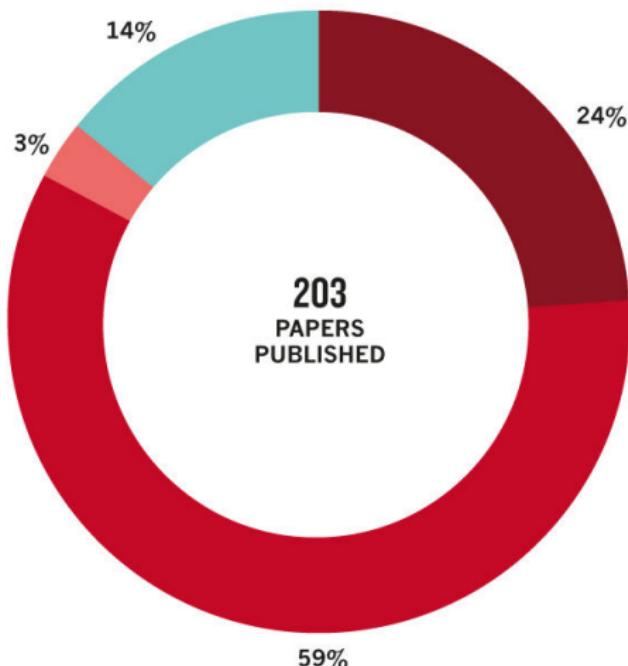


REPLICATION RARELY POSSIBLE

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

ELEMENTS PROVIDED*:

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



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

©nature

Protocolo del Proyecto, Estándares de Información

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Asegurar disponibilidad de todo lo necesario para que otro investigador replique su investigación.

- Encuentre el estándar de información apropiado para su disciplina y sígalo:
<http://www.equator-network.org>
- Informe todos los detalles sobre la implementación de un proyecto en un protocolo detallado:
<http://www.spirit-statement.org>
- Pautas Transparency and Openness Promotion (TOP):
<http://cos.io/top>

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

Promoting an open research culture



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

+ Author Affiliations

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

Flujo de trabajo reproducible: The Claerbout Principle

Transparencia
en
Investigación

Introducción
Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

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

Buckheit & Donoho, (1995)



BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Flujo de trabajo reproducible

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

“Reproducibilidad es simplemente una colaboración con personas que aun no conoces, incluido tu mismo en una semana más’ —Philip Stark, UC Berkeley Statistics

Flujo de trabajo reproducible

- Sugerencias prácticas sobre organización y programación
 - Al realizar cambios a archivos que han sido compartidos/publicados, implica que debe renombrar el archivo.
 - Haga referencia a la versión necesaria para ejecutar funciones.
 - Long (2008) *The Workflow of Data Analysis Using Stata*
- Programación Literaria: comentario extensivo en el código, apuntando a que sea legible por humanos
- Control de Versiones
- Documentos Dinámicos



Control de Versiones

Transparencia
en
Investigación

Introducción
Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

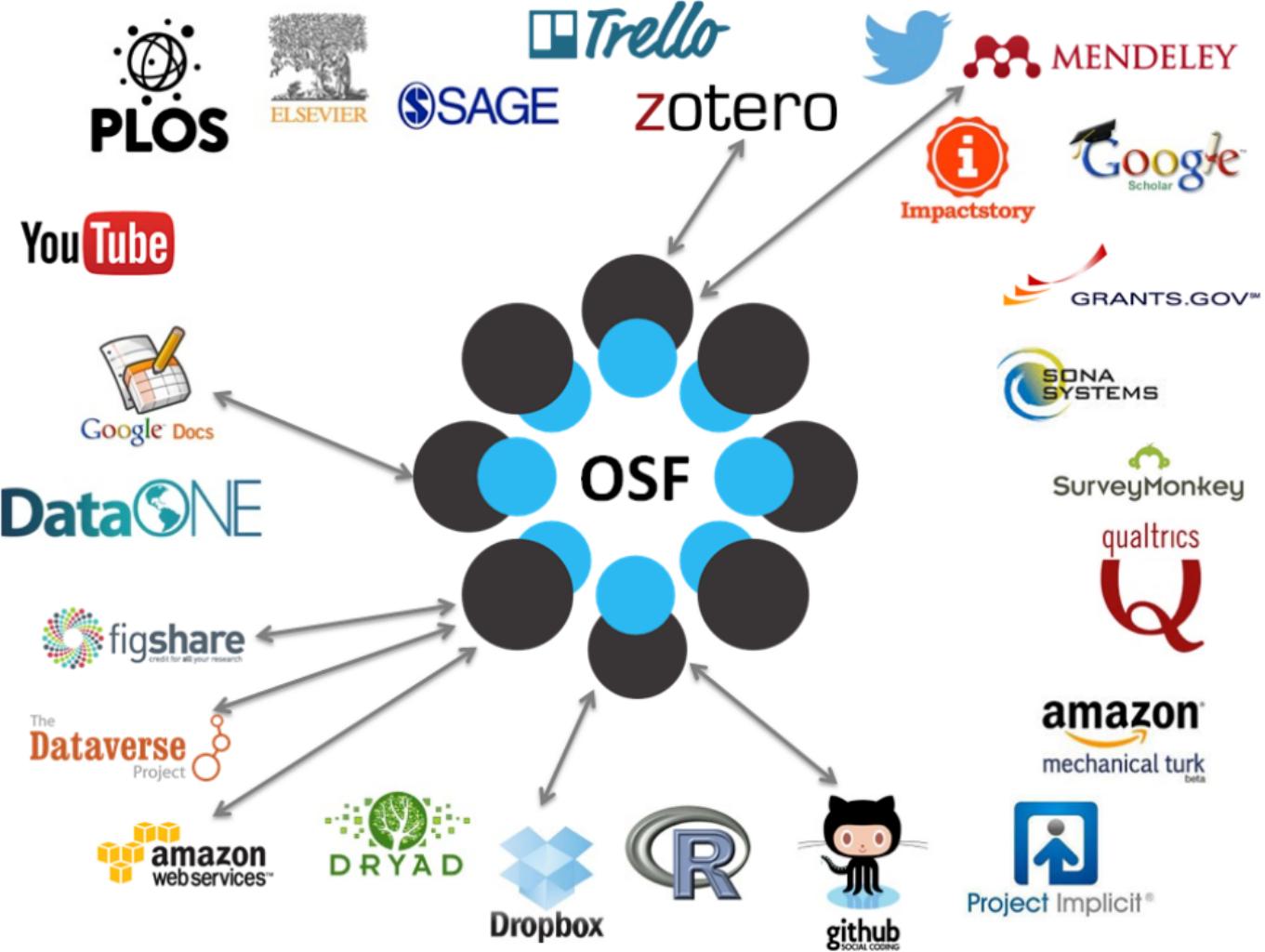
Conclusión

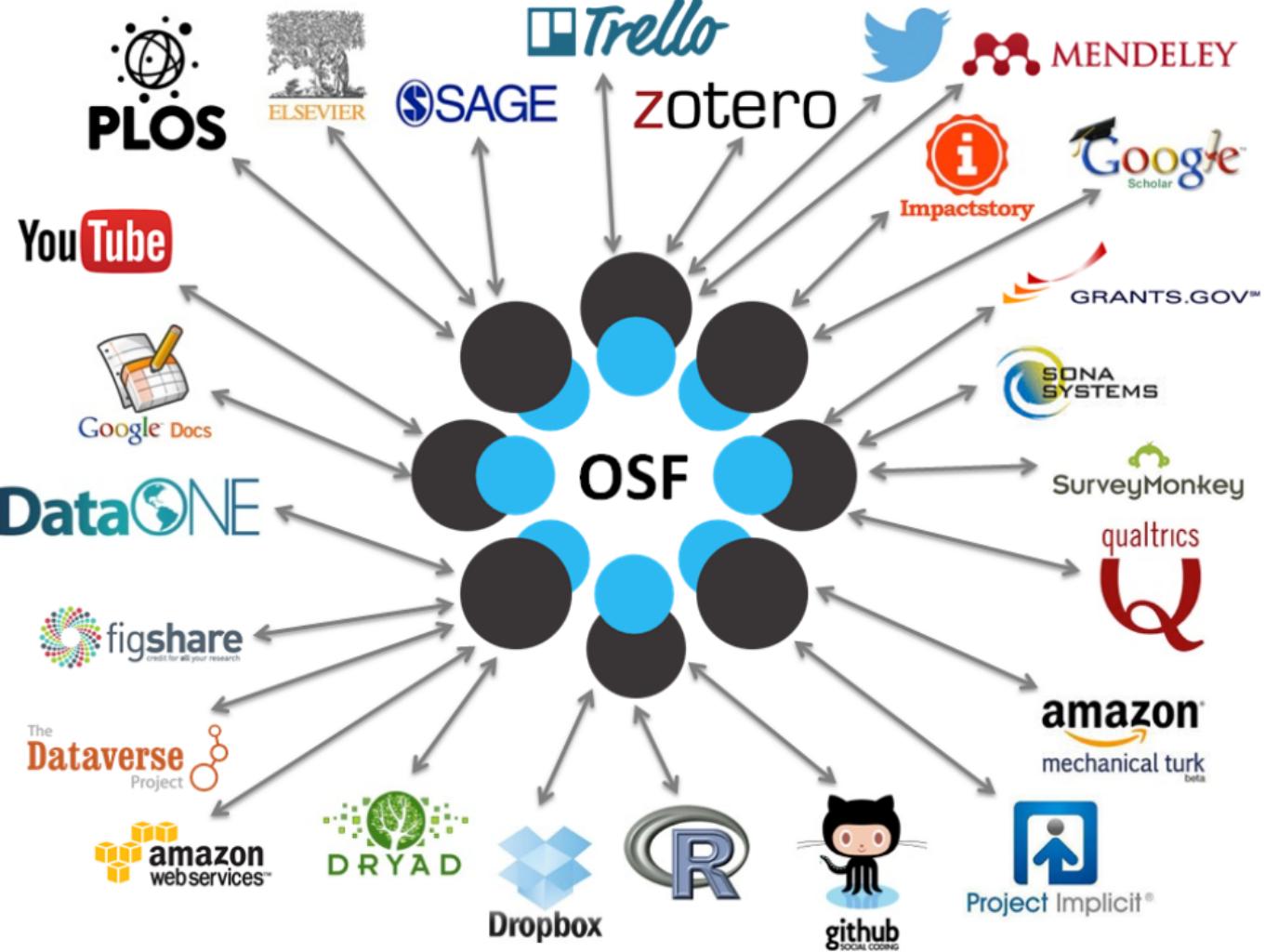
- Utilizar control de versiones puede ayudar a hacer su trabajo más reproducible.
- Qué es Control de Versiones?

Control de versiones es un sistema que registra cambios a uno o varios archivos en el tiempo, de forma tal que versiones específicas pueden ser recuperados más adelante.

–Git, About Version Control









BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Documentos Dinámicos

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Meta aspiracional: Escriba su código y su paper en el mismo archivo. De esta forma se minimiza la perdida de información y se eliminan errores de copiar/pegar.

Establecido en R/Python y comenzando en Stata.

- Incluye tablas dentro del archivo, en lugar de copiar-pegar-formatear elementos estáticos
- Todas las cifras dentro del texto también son calculadas de manera dinámica, eliminando la necesidad de tipeo de cifras manualmente.
- Cifras y tablas se actualizan automáticamente.
- El paper completo es (re)producido con uno o dos clicks

The logo for Jupyter consists of the word "jupyter" in a bold, black, sans-serif font. It is centered within a white circle. Above and below the circle are two thick, orange, curved bands that resemble the top and bottom edges of a smile. Small dark gray circles are positioned at the intersections of these curves with the circle.

jupyter



Compartiendo los Datos

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Publique su código y sus datos en un repositorio público de confianza.

- Encuentre el repositorio apropiado:
<http://www.re3data.org/>
- Los repositorios están diseñados para durar más que su website.
- También están diseñados para facilitar la búsqueda por parte de otros investigadores.
- Los repositorios guardan sus datos en formatos públicos que evitan obsolescencia en el tiempo.



Outline

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Introducción**
- 2 Ética en la Investigación Científica**
- 3 Sesgo de Publicación — Registros**
- 4 P-Hacking — Planes de Pre-Análisis**
- 5 Replicaciones Fallidas — Pautas, Protocolos y Herramientas**
- 6 Lecciones Para el Análisis de Políticas Públicas**
- 7 Conclusión**

Lecciones Para el Análisis de Políticas Públicas

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- Link Evidencia-PP con **alta** transparencia.
- Link Evidencia-PP con **baja** transparencia.
- Nuestra propuesta en BITSS.

Una forma en la cual la evidencia afecta las políticas públicas

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

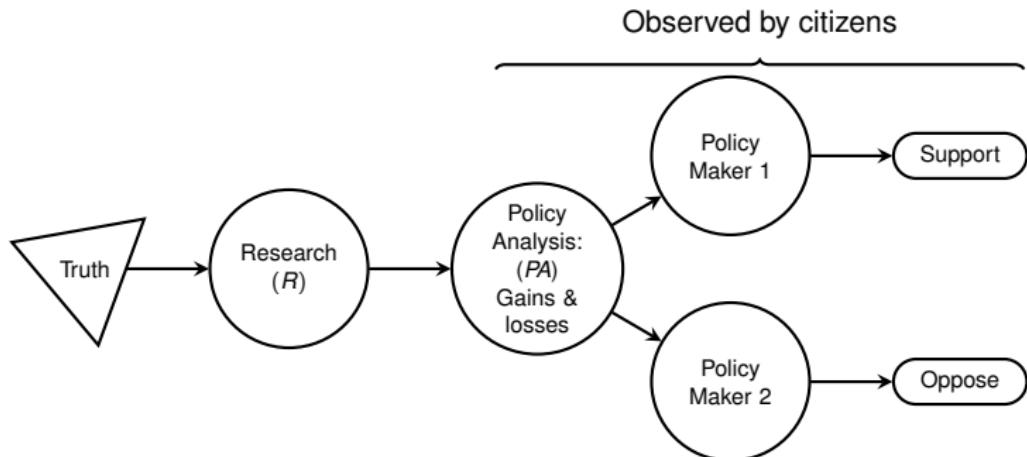


Figure: Simplified Model of Connection Between Evidence and Policy

Link Evidencia-políticas públicas con Baja Transparencia y Reproducibilidad

Transparencia en Investigación

Introducción

Ética en la Investigación Científica

Sesgo de Publicación — Registros

P-Hacking — Planes de Pre-Análisis

Replicaciones Fallidas

Lecciones Para el Análisis de Políticas Públicas

Conclusión

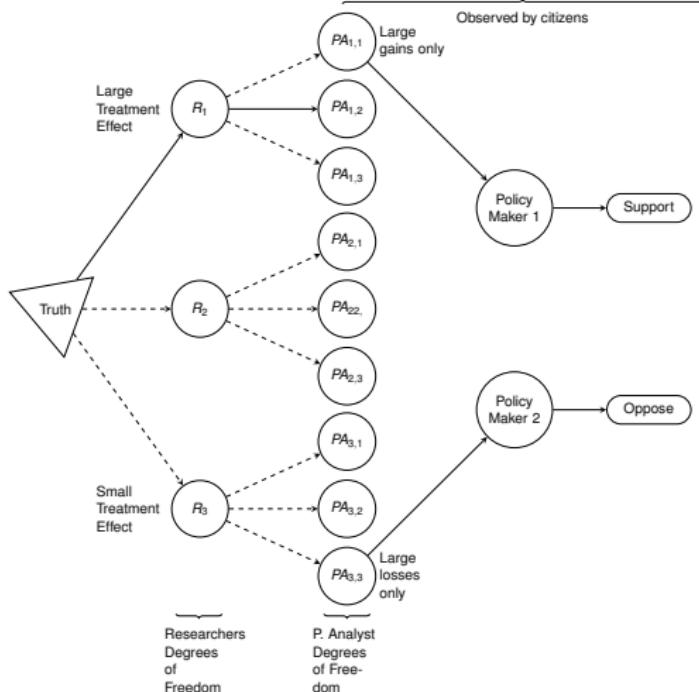


Figure: Policy-making with low TR in research and policy analysis

Nuestra Propuesta Para el Análisis de Políticas Públicas

Transparencia en Investigación

Introducción

Ética en la Investigación Científica

Sesgo de Publicación — Registros

P-Hacking — Planes de Pre-Análisis

Replicaciones Fallidas

Lecciones Para el Análisis de Políticas Públicas

Conclusión

- 1 Publicar pautas de transparencia y reproducibilidad para el análisis de políticas públicas (similar a Pautas TOP, descritas antes).
- 2 Colaborar con agencias/centro interesados en implementar estas ideas. Ejemplo aquí.
- 3 Iterar.



Outline

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

- 1 Introducción**
- 2 Ética en la Investigación Científica**
- 3 Sesgo de Publicación — Registros**
- 4 P-Hacking — Planes de Pre-Análisis**
- 5 Replicaciones Fallidas — Pautas, Protocolos y Herramientas**
- 6 Lecciones Para el Análisis de Políticas Públicas**
- 7 Conclusión**

Conclusión

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

OK, estoy convencido. Como puedo implementar esto en mi propia investigación?

- Tomen nuestro MOOC (E. Miguel). [▶ Link](#)
- Suscríbanse al blog & E-mail de BITSS [▶ Link](#)
- Aplicuen a nuestro Instituto de Verano: RT2. [▶ Link](#)
- Revisen nuestra página de recursos. [▶ Link](#)
- Interesado en incluir parte de estas prácticas en el análisis de políticas públicas? Hablen conmigo!
fhoices@berkeley.edu

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Tres días de entrenamiento.





BERKELEY INITIATIVE FOR TRANSPARENCY
IN THE SOCIAL SCIENCES

Transparencia
en
Investigación

Introducción

Ética en la
Investigación
Científica

Sesgo de
Publicación —
Registros

P-Hacking —
Planes de
Pre-Análisis

Replicaciones
Fallidas

Lecciones
Para el
Análisis de
Políticas
Públicas

Conclusión

Preguntas?

Muchas Gracias