

Research Integrity

Doing Political Research

Problems and Solutions

- Fraud
- Replication crises
- Publication bias and the file drawer problem
- Cherry-picking and p-hacking

- Peer review
- Data transparency and replication
- Pre-registration
- Null results reporting



Fraud

- High-profile scandals get a lot of attention eg Michael LaCour, Alice Goffman
- Causes
 - Competitive nature of academia
 - Weak safeguards?
- But such cases are thankfully rare
- More problematic are biases caused by structure of academic research involving researchers working in good faith









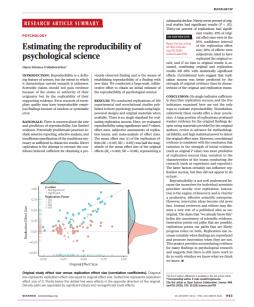
Investigating the replicability of preclinical cancer biology

Timothy M Errington1*, Maya Mathur2, Courtney K Soderber

Center for Open Science, Charlottesville, United States: 2 Quantitative Science nford University, Stanford, United States; "Science Exchange, Palo Alto, Unite tes; "University of Virginia, Charlottesville, United States

hat others can evaluate, extend, or challenge the evidence and claims. The accumulation weeds out claims that are unreliable, and supports the development of new models and the Over time, uncertainty declines and the accumulated knowledge provides useful de

Errington, 2020a; Salmon and Glymour, 1999). A scientific claim is said to be replicable w data or not. Moreover, the success or rature of an attempt to replicate rarey provises a en aniwer about the credibility of an original claim. When the replication attempt is succ infidence in the reliability of the claim increases, but that does not mean that the claim is ding can be both replicable and invalid at the same time. Repeated successful replications ca eliminate alternative explanations and potential confounding influences, and therefore in



Evaluating replicability of laboratory experiments in economics

Colin F. Camerer, 1st Anna Dreber, 2t Eskil Forsell, 2t Teck-Hua Ho, 3,4t Jürgen Huber, 5t Magnus Johannesson, 2+ Michael Kirchler, 5,6+ Johan Almenberg, 7 Adam Altmejd, 2 Taizan Chan,8 Emma Heikensten,2 Felix Holzmeister,5 Taisuke Imai,1 Siri Isaksson,2 Gideon Nave,1 Thomas Pfeiffer,9,10 Michael Razen,5 Hang Wu4

The replicability of some scientific findings has recently been called into question. To contribute data about replicability in economics, we replicated 18 studies published in the American Economic Review and the Quarterly Journal of Economics between 2011 and 2014. All of these replications followed predefined analysis plans that were made publicly available beforehand, and they all have a statistical power of at least 90% to detect the original effect size at the 5% significance level. We found a significant effect in the same direction as in the original study for 11 replications (61%); on average, the replicated effect size is 66% of the original. The replicability rate varies between 67% and 78% for four additional replicability indicators, including a prediction market measure of peer beliefs.

comes from the ability to replicate empirical findings directly and independently. Although direct replication is widely applauded (1), it is rarely carried out in empirical social science. Replication is now more important than ever, because the quality of results has been questioned in many fields, such as medicine (2-5), neuroscience (6), and genetics studies (19). (7, 8). In economics, concerns about inflated findings in empirical (9) and experimental analyses (10, 11) have also been raised. In the social sciences, psychology has been the most active in positives" and conducting direct replications

he deepest trust in scientific knowledge | (16, 17) quickly led to changes in journal publication practices (18). The recent Reproducibility Project: Psychology (RPP) replicated 100 original studies published in three top journals in psychology. The vast majority (97) of the original studies reported "positive findings," but in the effect in the same direction for 36% of these

In this report, we provide insights into the replicability of laboratory experiments in economics. Our sample consists of all 18 between-subject laboratory experimental papers published in the both self-diagnosing the forces that create "false | American Economic Review and the Quarterly Journal of Economics between 2011 and 2014. (12-15). Several high-profile replication failures
The most important statistically significant finding,

25 MARCH 2016 • VOL 351 ISSUE 6280 1433

Replication Crises

- Psychology: only 39% of findings in top journals could be replicated (Open Science Collaboration 2015)
- Economics: 61% of experiments could be replicated (Camerer et al 2016)
- Medicine: 46% of preclinical cancer experiments could be replicated (Errington et al 2021)
- Politics?

Publication Bias and the File Drawer Problem

- 10 researchers investigate the effect of religious diversity on conflict
 - 2 find that religious diversity drives conflict. These are published.
 - 8 find no relationship. These are not published.
 - Our knowledge is therefore skewed.
- What causes publication bias?
 - Journal editors, book publishers and reviewers are interested in new, exciting findings
 - Researchers lose interest if they don't find interesting results



Cherry-picking and phacking

- We investigate the effect of religious diversity on conflict and hypothesise a positive relationship between the two
 - We find no evidence that religious diversity affects most kinds of conflict, but it does increase riots
 - We publish just the finding on riots and claim that religious diversity increases conflict
- p-hacking refers to p-values, which tell us whether a finding is statistically significant
 - The traditional cut-off if 0.05, and more findings than expected fall just under 0.05
 - This implies either/both publication bias or deliberate manipulation of statistical models to achieve publishable findings

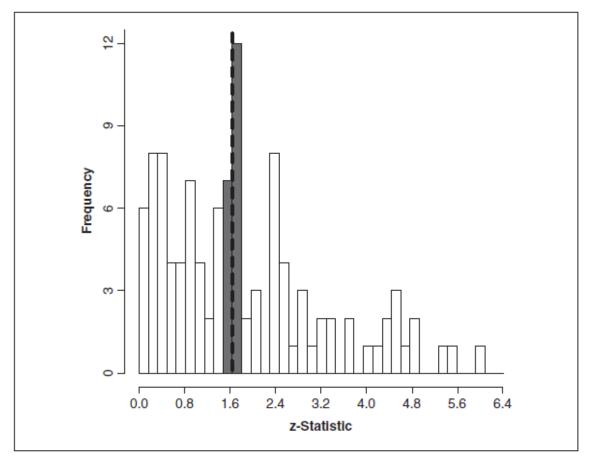


Figure 1. Histogram of z-statistics, economic voting (one-tailed) Note: Width of bars (0.16) approximately represents 10% caliper. Dotted line represents critical z-statistic (1.64) associated with p = .05 significance level for one-tailed tests.

Peer Review

- Articles and books are reviewed by experts before being published
- Origins in the 18th century but became commonplace from mid-20th century
- Usually double blind
- Possible outcomes of peer review: publish, reject, revise and resubmit
- Problems
 - Bad or slow reviewers
 - Reviewers are not neutral
 - Reviewers do not see underlying data
 - Double blind reviewing is difficult in the Internet era

Rereads

Migrant Remittances and Extralegal Responses to Crime in Latin America 3/21/20

Reviewer 1

Recommendation: Revise and Resubmit

Before sharing my feedback, I must say that I understand and appreciate that the authors have focused their efforts on offering various robustness checks, along with an extensive appendix that can persuade readers about the consistency and transparency of their findings. However, by doing so, they have substantively sacrificed the theoretical logic that sustains their argument, as well as the conceptual clarity that this study would require.

The authors expect that remittance recipients are more supportive of citizen-administered justice (H1), military involvement in policing tasks (H2), and police bending the law to capture criminals (H3). These three dimensions are part of what the authors call "extra-legal violence." However, one could divide this phenomenon into two additional concepts: taking justice in their own hands (H1) and mano dura (H2 and H3). Initially, these elements do not seem to be conceptually the same idea and may or may not be necessarily linked. Therefore, the authors must (a) clarify their concept of "extra-legal violence" and (b) connect their three different outcomes (if they indeed go together) in a theoretically consistent and logical causal chain.

Regarding (a), as noted, the authors' concept of "extra-legal violence" is a bit blurry. Also, at different points (see pages 14 and 21), the authors refer instead to "extralegal measures against crime" or, more broadly, to "violent" responses to crime, but each of these different references would imply different things. For instance, it is not clear that the military's involvement in policing tasks implies extra-legal violence or extralegality; at least in the terms in which the authors seem to present this concept. This is a point the authors would have to elaborate further. Also, if the authors' focus is on broad support for "violent responses to crime," one wonders if this would require a comparison between support for violent vs. non-violent responses to crime (e.g. support for preventive and social policies).

Regarding (b), currently, the authors build from a varied set of empirical works that indeed suggest the relationships that the authors propose here, but such correlations are sustained through diverse and not necessarily connected mechanisms. The literature and extant findings are quite broad and contradictory at times. The authors' job for this new approach to their paper, therefore, is the presentation of a coherent argument that derives from a clear set of theoretical propositions that help the reader makes sense of the broad set of findings the authors review and that present a direct argument with underlying mechanisms, even if they are unable to test them thoroughly. This task necessarily goes beyond the literature review on which they currently focus on and present as the basis of

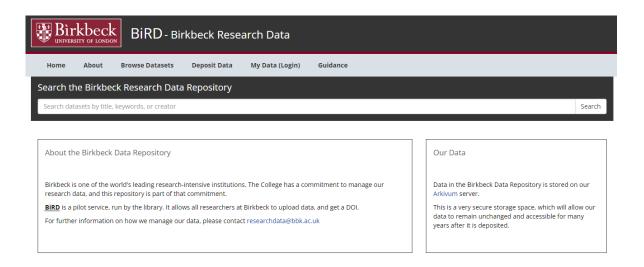
Pre-registration



- Publishing your theory and research design before conducting the research so that you cannot cherry-pick your findings
- 'Registered reports': journals commit to publishing an article without seeing the results
- Problems:
 - No common standard for pre-analysis plans (yet)
 - What about exploratory, theory-generating research?
 - Works more easily for quantitative than qualitative research
 - Genuinely interesting findings that were not initially considered might get ignored

Data Transparency and Replication

- APSA embraced a policy on 'Data Access and Research Transparency' in 2015
- The principle is that all data used in research should be made accessible so that it can be examined and replicated
- Many journals now require this, and some independently replicate analyses before publication
- Problems
 - Some data is proprietary or has been obtained at great cost to the researcher
 - Clashes with ethical principles about anonymity of research participants
 - Quantitative datasets can easily be shared, but what about qualitative field/archive notes?



Null Results Reporting

- Combatting the file drawer problem by reporting null results
- Some journals have started publishing null results reports
- Alrababa'h et al suggest doing this informally
- Problems
 - What is the incentive for researchers to write up null results?
 - Is the null result due to problems with the research design?
 - A null result is much more obvious in quantitative than qualitative research





The Null Hypothesis

Do Free Fo

Do Free Food and Beverages Bring People to the Political Meeting? The Survey Experiment of Attendance-Buying Political Studies Revie
2022, Vol. 20(4) 691-71
© The Author(s) 201
Article reuse guideline
sagepub.com/journals-permissio
DOI: 10.1177/1478929921104733
journals.sagepub.com/per

SSAGE

Milan Školník¹, Michael Haman¹
and Jan Čopík²

Abstract

In this article, we explore a phenomenon we call attendance-buying. Specifically, we investigate the case of buying the attendance of voters at political railies. Politicians and political parties promise voters certain rewards if they come to listen to them at political meetings. Often, these rewards can be food or valuable prizes. We conducted a survey experiment to discover what effect this form of attendance-buying has in Czech municipalities with up to 1000 inhabitants. We focused on small municipalities, where such a strategy can be effective because of the small number of candidates compared to large towns. We found that the effect of promising free food and drinks for attending a rally surprisingly did not manifest. With this finding, we contribute to and extend upon the theme of vote-buying. Our research also has practical implications, as local-level candidates do not have to make these food and drink expenditures because interested citizens will attend political railles regardless of these offers.

Ceywords

vote-buying, participation, survey experiment, local politics, Czech Republic

Accepted: | September 2021

Introduction

In this article, we address the issue of buying attendance at political meetings in the context of the Czech Republic. In this post-communist country in Central Europe, it is common for politicians to lure voters using various food items to have them attend

Corresponding author:

Mílan Školník, Department of Political Science, Philosophical Faculty, University of Hradec Králové, Rokitanského 62, 500 03 Hradec Králové, Czech Republic.

Email: milan.skolnik@uhk.cz

¹Department of Political Science, Philosophical Faculty, University of Hradec Králové, Hradec Kralove, Czech Republic

³Department of Humanities, Faculty of Economics and Management, Czech University of Life Sciences Prague, Praha, Czech Republic

Conclusion

- There are many threats to the integrity of political research
 - Most are caused by the structures and institutions of research and publication, rather than malfeasance
- In recent years, new strategies have been developed to tackle some of the problems
 - But these bring a new set of challenges
 - Many strategies work better for quantitative and positivist research, but can be adapted for qualitative and non-positivist work
- Reflecting on research integrity is very important for both researchers and those who read and use research