Reproducible and Robust by Design: The Reproducibility and Robustness of Social Science Research by Method and Field

Cook, Mikola, Aparicio, Fiala, Abdul Baki ... Brodeur*

Friday 13th November, 2026

Abstract

The use of causal identification methods has increased rapidly in recent years in response to a credibility crisis. Despite their importance, little is known about the extent to which the different causal identification methods available to researchers yield credible findings in applied research. This paper documents the reproducibility and robustness of economic and political science research by method and field for 274 published articles. We find that a researcher can reproduce the results of an article using the provided data and code for only 89% of articles that are recently published in leading journals. Further, we find that 32% of articles contain a coding error or research inconsistency. Our primary finding is that 70% of originally statistically significant research remains statistically significant when subject to reasonable alternative research choices. When articles are classified into fields (e.g. labor economics, comparative political science) and their identification method (e.g. instrumental variables), we find large differences in reproducibility and robustness, particularly for the different methods. When using a many-analysts approach, we find that results from experimental methods are relatively more robust than observational methods, and that quantitative and public economics are relatively more robust than macroeconomics, economic history, and development economics. Last, 249 expert forecasters systematically underestimated the reproducibility of published research, overestimated the presence of coding errors, and forecast (correctly) that experimental methods are more robust than observational methods.

Keywords: Reproducibility, Causal Inference, Robustness, Research Design, Open Science, Economics, Political Science

JEL Codes: B41, C10, C81

*Corresponding author: Nikolai Cook (Wilfrid Laurier University), Email: ncook@wlu.ca. See Section A.1 for the full list of authors and Section A.2 for each author's contribution. Disclaimers: We acknowledge support from the Alfred P. Sloan Foundation, Coefficient Giving, and the Social Sciences and Humanities Research Council. This analysis was pre-registered on OSF at https://osf.io/zvqf8/files/kehrj. See Appendix B for a copy of the pre-analysis plan. See Appendix C for the list of included reproduction reports. All errors are the authors' responsibility.

Full author list, Section A.1.
Authors' contribution, Section A.2.

1 Introduction

The credibility revolution in the social sciences has been driven by the widespread adoption of causal identification methods (Goldsmith-Pinkham, 2026). Indeed "there is no arguing with the fact that experimental and quasi-experimental research designs are increasingly at the heart of the most influential empirical studies in applied economics" (Angrist and Pischke, 2009). Despite this, the extent to which research credibility varies across identification methods (and the fields they are differentially represented in) remains an empirical question. The different methods, ranging from randomized controlled trials to the quasi-experimental identification strategies of regression discontinuity design (RDD) and instrumental variables (IV), to structural models, all have distinct assumptions, data requirements, and analytical flexibilities that likely affect the robustness of any empirical research result. Similarly, fields differ not just in the appropriateness of the different identification methods, but also in the norms of their empirical traditions. Understanding how methodological and disciplinary features of research relate to the robustness of its results is a first step in assessing the success of the credibility revolution.

This paper compares reproducibility and robustness across economics and political science subfields and methodological approaches. To do so, it leverages a large number of reports provided by skilled researchers for 274 recently well-published articles.[1] By examining a broad set of primarily economics articles ($n = 214$) across 10 pre-specified sub-fields (such as labor economics, or industrial organization), while also including a non-trivial number in political science ($n = 60$) across an additional three fields (American, comparative, and international), we aim to shed light on the reproducibility and robustness of recently published social science research. By examining the reproducibility and robustness of such a large number of articles, we are also able to speak to the credibility of the most common identification strategies used in empirical or applied research; we examine both quasi-experimental (e.g. IV, RD) as well as experimental methods including field, laboratory, and online experiments.

Previous large-scale efforts have examined the reproducibility and robustness of scientific findings, revealing substantial heterogeneity in the reliability of results across disciplines. From targeted replication projects in economics (Camerer et al., 2018), management (Davis et al., 2023), and psychology[2] (Open Science Collaboration, 2015), to broader meta-research campaigns (Aczel et al. (2026); Brodeur et al. (2026); Fišar et al. (2023); Miske et al. (2026); Pérignon et al. (2024); Tyner et al. (2026)), these efforts document important shortcomings in the reproducibility and replicability of published research. Yet existing studies generally evaluate reproducibility and robustness within particular disciplines, samples, or research designs.[3] We systematically examine whether reproducibility and robustness differ across both substantive fields and empirical methods within

[1] This paper more than doubles the data compiled for Brodeur et al. (2026), which includes reproducibility checks for 110 articles published from 2022 onward. Here, we expand upon that data with reproducibility checks for an additional 164 articles in the same high-impact journals, including 13 reproductions from Campbell et al. (2024). Our sample includes 20,233 robustness checks from 274 reproduction reports for articles published in 12 leading journals.

[2] See Nosek et al. (2022) for a review.

[3] With the notable exception of Miske et al. (2026) who examines broad fields of Business, Economics, Education, Political Science, Psychology, and Sociology.

a common sample and measurement framework. We see two reasons this distinction is important: (1) if reproducibility varies predictably with research design or field, then failures to reproduce may reflect features of how evidence is generated rather than a uniform credibility crisis of all empirical research and (2) users of research, such as policymakers and the public, can better weigh the applicable extant literature in their own determinations.

Our large scale effort leverages reports written by staff members of the Institute for Replication (I4R) and over one thousand researchers. Many of these reports were written during and after events called "Replication Games" (Brodeur et al., 2023, 2024) – events held in partnership between host universities and I4R where small teams of replicators (typically mixtures of graduate students, university faculty members, and government researchers) check the credibility of an article's main claims. The replicators check whether the data and code provided by the article's original authors compute the published results (and whether those results are actually what the article claims them to be), detect and correct major coding errors, and conduct sensitivity analysis. The data for this project is the result of 274 reproduction reports on published articles which satisfy our pre-registered field and journal restrictions. These 274 reports, containing $20,233$ robustness checks, form the primary dataset we use.

Our results are as follows: Data is fully available, in either raw or analysis form, for 81% of articles. Recently and well-published research in economics and political science is computationally reproducible (using the data and code provided by the authors) in 89% of cases.[4] One caveat is that 32% of articles have a coding error major enough to materially affect the interpretation of the results. Our headline robustness result is that 70% of articles are robust - specifically that the article-level mean of the proportion of statistically significant results remaining statistically significant under re-analysis is 70% (following the definition in Dreber and Johannesson (2025)). There is substantial variation in both field and method for many of these markers; Data availability varies from 60% to 93% by field, and 72% to 90% by method. Coding errors vary from a minimum of 10% to 45% by field, and 14% to 39% by method. Robustness (re-analysis leads to an estimate of the same sign and still statistically significant) varies from a low of 52% to 84% by field and 66% to 89% by method. Even alternative measures of robustness vary, such as relative statistical significance (68% to 127% by field) and relative effect size (87% to 102% by method).

Our descriptive results provide an initial estimate of robustness rates across methods and subfields, but these estimates should be interpreted cautiously and not as "all else equal" differences. Because evaluating robustness checks entails subjective judgment and depends on replicator experience (Brodeur et al., 2026), we emphasize relative patterns rather than absolute levels. To assess the stability of these patterns and account for (potentially subjective) analytical choices, we also employ a many-analysts approach (Botvinik-Nezer et al. (2020); Breznau et al. (2022); Huntington-Klein et al. (2021); Menkveld et al. (2024); Silberzahn et al. (2018)). 13 independent many analyst teams are given our database of robustness checks and the same research questions. They each apply their

[4]The high data availability and computational reproducibility rates are likely due in part to the selective nature of our sample. Indeed, our focus is on articles published recently in leading journals with data and code availability policies. Furthermore, some of these journals conduct internal reproducibility checks prior to publication (Vilhuber et al. (2020)).

own set of analytical choices, such as variable definitions, model specifications, or estimation techniques, but all rely on the same dependent variable (if the original estimate remains statistically significant at the 5% level with the same sign). This approach allows us to capture the range of plausible outcomes that can emerge from a single dataset and study design, and to systematically compare how robustness varies across disciplines and methods.

The results from the many-analysts exercise reveal a high degree of consistency in the qualitative patterns across methods and fields. Despite substantial variation in model specifications and analytical choices, most analysts find that experimental methods exhibit higher robustness than observational methods. The direction of these differences is remarkably stable across analysts, even though the magnitude varies. Similarly, across sub-fields, analysts consistently identify lower robustness in areas such as macroeconomics and economic history, and higher robustness in sub-fields such as public economics and industrial organization. While some dispersion remains in estimated effect sizes, the overall ranking of methods and fields is largely consistent, reinforcing the reliability of the descriptive patterns and suggesting that our main conclusions are not driven by a particular sequence of analytical choices.

We then benchmark our descriptive results to predictions made by expert forecasters. Prediction surveys are increasingly used in meta-science research as they provide a valuable benchmark by eliciting *ex ante* expectations against which realized outcomes can be systematically compared (DellaVigna et al. (2019)). To do so, we administered a pre-registered survey on the Social Science Prediction Platform (SSPP), targeting social science researchers with varying degrees of replication experience. Respondents were asked to predict computational reproducibility rates, the prevalence of coding errors, and robustness rates for both the full sample of articles and relevant subsamples. The survey yielded 249 complete responses, and the results reveal systematic patterns: on average, forecasters underestimated computational reproducibility, overestimated the prevalence of coding errors, and underestimated robustness. Respondents consistently expected higher robustness from experimental designs (particularly field experiments) relative to observational research (especially that which relied on instrumental variables).

Our paper makes at least three contributions to the literature: (i) we provide the first large-scale comparison of robustness across identification methods and fields, (ii) we introduce a many-analysts framework to assess the stability of these methods and fields patterns to researcher degrees of freedom, and (iii) we compare observed reproducibility measures against expert forecasts. By documenting variation in robustness across methods and sub-fields, our analysis advances the goal of cumulative science by helping researchers and policymakers better calibrate their expectations (Vivalt and Coville (2023); Vivalt et al. (2025)), strengthening incentives for methodological rigor, and identifying areas most in need of reform.

The remainder of this paper is organized as follows. Section 2 describes our data collection methodology and provides descriptive statistics. Section 3 exposits the materials needed to determine the robustness of research including data availability, computational reproducibility, and the prevalence of coding errors. Section 4 documents the robustness of economics and political science research, including by method and field. Section 5 provides the results of a many-analysts approach

to answer whether reproducibility differs by method and field, and whether those findings are robust to their own alternative research choices. Section 6 details and compares the results of our pre-results expert forecast survey to the observed reproducibility. Section 7 provides a discussion of how statistical significance shifts under reproduction, also by method and field, extending the earlier work of Brodeur et al. (2026). Section 8 concludes.

1.1 Definitions

Before continuing, it is useful to define reproducibility, robustness and replication. The concepts of reproducibility and replicability have been defined in various ways across the literature with no universally accepted standard.[5] We follow the most recent taxonomy proposed in Dreber and Johannesson (2025) and applied in Brodeur et al. (2026).

We define reproducibility as the ability to duplicate published results using the same data and code. A subcategory of reproducibility, computational reproducibility refers to the ability to obtain the same results as the original article using the original article's data and code. Another subcategory, robustness reproducibility refers to the stability an original article's results to alternative, yet reasonable, analytical choices, such as different control variable horizons, key variable definitions and operationalization, or even simple sample restrictions. We refer to this last category as simply 'robustness' throughout.

In contrast, we define replicability as testing results with new or independently collected data. That is, replication involves re-estimating the same empirical question using new data to assess whether the original findings generalize. Together, these concepts provide complementary lenses to evaluate the credibility of empirical research.

2 Data

In this section, we describe our data and provide summary statistics. Our focus is to examine the reproducibility and robustness of research articles by method and sub-field. To be able to examine these possibly thin slices of data, we require a large number of reproduced articles with an even larger number of robustness checks.

The primary data used in this project is an extension in both size and scope to the data used in Brodeur et al. (2026) which examined 110 articles published in 2022 and 2023 and to several *American Economic Review* articles reproduced in Campbell et al. (2024). Our project more than doubles the previous number of articles reproduced, examining the results of reproduction reports written for 274 unique articles published primarily in 2022, 2023, and 2024.[6] Indeed, the reproducibility and

[5] An early definition from Hamermesh (2007) is "pure replication – checking on others' published papers using their data" and "scientific replication – using data representing different populations." Distinguishing between replication and reproducibility at the article and test level, Clemens (2017) defines the replication test (same sampling distribution), verification test (exact statistical analysis in paper produces the paper), reproduction test (resampling from population using identical methods), robustness test (different sampling distribution), re-analysis test (materially altering the code from the study), and extension test (a new sample by characteristic or time).

[6] In total 28 articles (10% of the sample) were published prior to 2022.

robustness checks which serve as the data used in this paper are derived, in most part, from 50 Replication Games, held in 44 cities, across 15 countries, all co-hosted by the Institute for Replication and a local organizer.[7]

From the available reproduction reports, we restrict our sample to articles published in our pre-specified list of journals and assigned to at least one sub-field included in our pre-registered pre-analysis plan. These restrictions yield a final sample of 274 unique articles.[8]

An important and idiosyncratic feature (to I4R) is the deliberate cultivation of a spirit of scientific collegiality that is explicitly fostered and demanded of all participants. For instance, before and during each Game, replicators are clearly and repeatedly told that the goal is **NOT** to show that the results of their chosen article are not reproducible. That is, I4R emphasizes to replicators that the goal is testing whether the main claims of published research are reproducible and robust. This is key as in the absence of this culture (and the process I4R has adopted) some replicators might engage in reverse specification searching (i.e. selective reporting of insignificant results, similar to the discussion in Dreber et al. (2024)).

2.1 Replication Games and Board of Directors

I4R has two streams to generate reproductions, robustness checks, and replications. The primary stream is the Replication Games. Once replicators have indicated interest and signed up to participate in one of the Replication Games (which are held as standalone events, or often integrated and popularized by the organizers of major economics conferences such as the Canadian Economics Association Meetings, the European Economics Association Meetings, the Society for Political Methodology, the Southern Economics Association Meetings, and the Western Economics Association International Meetings), they are assigned to a small team of about 3 to 5 researchers, ideally with similar fields of interest. In most cases, replicators then choose an article to reproduce from a list of about five articles that are from high impact journals with data and code availability policies, and which passed preliminary checks for data and code availability for analysis during replication games. When surveying the replicators, the most common reason given for their choice of article was the methods used (58%) followed by the journal of publication, length of time to compute results, and size of the replication package. Only 4% of teams indicated that they chose their paper because of *ex ante* beliefs that main results are (not) robust/replicable.

During the Games, replicators, who consist of university faculty, postdoctoral scholars, research scientists, as well as graduate students, provide I4R with results that are later turned into a formal report with the help of I4R staff.[9] In each, the replicators identify the core claims of an article, then attempt to computationally reproduce the original results using the available data and code.

[7] Participants are offered the possibility to virtually attend replication games; about two-thirds of participants attend replication games in-person.

[8] Appendix C lists the reports, of which 284 are represented - 10 articles had multiple teams assigned - we *analyze* reports from only first assigned team.

[9] I4R also facilitates hosting these reports, including acting as liaison between the original authors and the replicators should the replicators wish to remain anonymous. This structure and its benefits are detailed at length in Brodeur et al. (2026).

Replicators are then encouraged to explore robustness checks or modify coding choices as they see fit.[10] Replicators are then asked to provide a report detailing their complete robustness and reproducibility checks, which can continue to be worked in after the games and sent to I4R at a later date.

A secondary stream generated a minority of our reports ($n = 42$). In this stream, I4R directly recruits replicators recommended by its board of directors[11] often because these replicators are known to produce replications of particularly high-quality.

2.2 Reports

The open-ended structure of I4R's templated report is critical because research articles vary in their content (topic, methods, novelty), vary in their provided data, code, and documentation. Each of these factors individually affects which robustness checks are scientifically appropriate,[12] informative, or even feasible, and so letting researchers - replicators - do what they do best with minimal oversight is a feasible, if not possibly one of the only, approaches to reproduction and robustness at scale. To account for this variation and benefit from the topic-and-methodology expertise of our coauthors, when we provide replicators with a list of about five potential articles to choose from, we ensure they are at least closely-aligned to their areas of expertise, allowing for more meaningful and competent evaluations of robustness and reproducibility than random assignment would achieve.

While this *laissez-faire* approach offers several strengths (e.g. enabling replicators to rigorously probe key results and act as detailed reviewers on topics and methods with which they are arguably subject matter experts), it also introduces potential variability in the content and/or quality of the resulting reproduction and robustness reports. Some replicators may choose to prioritize analysis of the code and methods of their article. Other replicators may focus more on the scope of feasible reanalyses given the contents of the replication folder, tractability of the different robustness checks, data availability, script clarity, computational reproducibility, coding errors, and computation time (of the original article and the proposed robustness check). There are also individual constraints and team dynamics which can either promote or hinder the quality of a report. In other words, reproduction quality varies across teams; this is addressed by the fact that I4R staff (researchers themselves) worked with each replicator team and their results to clarify the inputs to each of the formal reports for I4R. It is also at this stage that a rich set of variables are included in our data, such as author and replicator characteristics.

[10] Pre-Games meetings provide replicators with examples of commonly implemented robustness checks. That pre-Games meeting is a necessary component of the Games, where all game participants take part in a forty-five minute pre-game presentation which details the protocol to be followed and logistics. Participants viewed the presentation live or via recording. Also, during the Games themselves, lead authors and representatives of I4R (A.B., D.M., J.P.A, L.F.) provide guidance to teams of replicators. Further, before and during each of the Games, I4R stresses the importance of *reasonable* robustness checks (Simonsohn et al. (2020)). Re-analyses are sensible tests of the research question and expected to be statistically valid and theoretically informed. This is another reason why replicators are given a choice of studies in their own field and using methods they are familiar with.

[11] All board members are nominated by the last author, A.B., who invites them to suggest researchers whose expertise is well suited to reproducing studies published in the targeted journals.

[12] For example, sensitivity analysis for an economics article leveraging a difference-in-differences analysis may require a different set of robustness checks than a political science article with a regression discontinuity design.

Once a report is completed, it undergoes editorial review by I4R. Reports are not peer-reviewed in the conventional sense, but are expected to follow both professional norms and are checked for tone by one of the lead authors (A.B., D.M., J.P.A, L.F.) or an I4R board member. A.B. (or a co-director in the event of a conflict of interest) then reaches out to the original article authors to offer them an opportunity to comment or respond. Reproducers can revise their reports based on the original authors' feedback, and original authors may also adjust their responses. Final reports and authors' responses are published openly on I4R's webpage (https://i4replication.org/). In Appendix Section C.1 we provide the entire list of articles with links to all reports.

Once replicators completed their reports and an additional survey about their experience and the replication package contents, the details are coded into a database by lead authors (N.C., A.B., D.M., and J.P.A.) in the same manner as in Brodeur et al. (2026). Each coded report-information on article-level computational reproducibility and coding errors, combining it with publication information such as year, journal, Google Scholar citations, research field, and the number of original authors, and survey based information such as data availability.

At the robustness check level, only a subset meeting specific criteria are included: excluded are extensions of the original research, heterogeneity effects, and mediation analyses, as they lack direct "original" estimates for comparison. Those that meet the criteria are then categorized into one of eight groups: (i) sample, (ii) control variables, (iii) dependent variable, (iv) main independent variable, (v) estimation method or model, (vi) inference method, (vii) weighting scheme, and (viii) replication using new data.[13] The most common robustness check categories are changing the sample (33%), controls (27%), changing the estimation method (11%), and changing the dependent variable (7%). In practice, coding decisions for robustness checks often span multiple categories. For example, changing the dependent variable may also alter the sample size due to differences in missing or available data.[14]

2.3 Summary Statistics

Our final analysis data includes the results of $20,233$ robustness checks conducted on results in 274 unique articles that were recently published in prominent economics and political science journals (all journals pre-specified). In descending order, the reproduced articles are from the *American Economic Review* (53), *The Economic Journal* (45), *American Economic Journal: Applied Economics* (28), *American Economic Journal: Economic Policy* (27), *American Political Science Review* (23), *American Journal of Political Science* (19), *The Journal of Politics* (18), *Journal of Political Economy* (16), *The Quarterly Journal of Economics* (15), *American Economic Journal: Macroeconomics* (13), *American Economic Review: Insights* (9), and *The Review of Economic Studies* (8). Our sampling of journals and articles deliberately favors high-impact journals and recent articles, allowing for the rapid reassessment of findings in eco-

[13] For example, replicators change sampling by adding or removing years, geographical units or individuals. Control variables can be added or removed. Dependent variables and independent variables can be rescaling or redefinition. An example of changing the estimation method would be to estimate a probit model instead of an ordinary least squares when the outcome was a binary variable. A change in the inference method is most typically modifying the level of clustering. Replication requires the inclusion of new data (Dreber and Johannesson, 2025).

[14] The relative proportion of robustness check types by method and field are provided in Appendix Figure A7.

nomics and political science. Further, focusing on recent articles allows our results to be interpreted as what can be achieved with existing academic norms and currently feasible journal policies.

The average number of robustness checks per article is approximately 74 (standard deviation 107).[15] The high number of checks is due to replicators testing multiple original claims within an article (average of 16 with standard deviation 32) multiple times (average of 4.6 tests per claim from 74 / 16). This is unsurprising as, for example, if an original article features 2 main outcome variables and employs 7 main specifications, reproducers performing 5 robustness checks for each outcome variable and specification would generate around 70 robustness checks. The large standard variation in the number of robustness checks *across* reports is notable but unsurprising as well for several reasons. First, some teams focused heavily on conducting robustness checks for several central hypotheses, while others were more focused on computational reproducibility or dealing with coding errors. Second, running the analysis code for some articles takes a few seconds to minutes, while others can take hours, days, or weeks.

Economics and political science sub-fields are well represented in our large sample. In descending order for economics, 113 articles are microeconomics (JEL D, 41%), 84 articles are economic development (JEL O, 31%), 70 are labor and demographic economics (JEL J, 26%), 56 are health and education (JEL I, 20%), 49 are mathematical and quantitative methods (JEL C, 18%), 41 are industrial organization (JEL L, 15%), 34 are public economics (JEL H, 12%), 29 are macroeconomics (JEL E, 11%), 29 are economic history (JEL N, 11%), 28 are financial economics (JEL G, 10%). In descending order for political science, 35 are comparative political science (13%) 20 are American political science (7%), and 5 are international relations (2%). Please note that articles in economics tend to have more than one JEL code (sub-field) assigned to them, whereas articles in political science are assigned to three mutually exclusive sub-fields.[16]

Methods are similarly well represented in our sample. In descending order, 51 articles use difference-in-differences or two-way-fixed effects (DID, 19%), 49 use field experiments (18%), 28 use instrumental variables (IV, 10%), 27 use online experiments (10%), 18 use regression discontinuity designs (RD, 7%), 14 use laboratory experiments (5%), and 157 use other methods including simulations, structural, theory, time series, synthetic control, and more (including non-causal methods)(57%). Note that articles may employ more than one empirical method and are therefore not assigned to mutually exclusive categories.

2.4 Replicator Characteristics

The replicators who provided the reports are not inexperienced. 80 reports (29%) are written by teams in which at least one replicator has published in one of the journals replicated in this project. Graduate students constitute, on average, 46% of replication team members (median 50%), with

[15] Also see Table A1 for additional statistics.

[16] Appendix Table 5 of Card and DellaVigna (2013) provides a useful, though limited, comparison. Relative to the 2010–2012 distribution of articles in the Top 5 economics journals, our sample has greater representation in microeconomics (41% versus 27%), labor (26% versus 20%), health-related fields (20% versus 8%), industrial organization (15% versus 14%), and public economics (12% versus 10%), while it has lower representation in macroeconomics (11% versus 17%) and finance (10% versus 14%)

remaining members consisting of researchers with PhDs or university faculty. The average number of years since completion of PhD among team members, while assigning zero to graduate students, is 3.3 (median 2.2), while replicator teams average 984 citations (median 115) when assigning zero to researchers without publications. These characteristics indicate that replication teams possess substantial research experience despite a relatively junior composition, reducing concerns that observed failures of reproducibility or robustness might reflect only a lack of expertise among the replicators. Indeed, while our replicators tend to have less academic experience than original authors (average years post-PhD of 11.9), their level of programming expertise is quite advanced; 12%, 49% and 32% of replicators report that their level of expertise is "Expert", "Proficient" and "Competent," respectively.

3 Reproducibility Materials

This section provides details on the raw materials available to replicators by method and field, and using those raw materials, diagnostics of the reproducibility of published articles.

First, we examine the availability of data and codes - necessary materials to achieve the most basic level of reproducibility. We differentiate between (for example) replication packages with raw data against those with analysis data only, as a replication package that provides raw data increases the scope of available robustness checks; in Brodeur et al. (2026), replicators indicated that the primary limiting factor to not conducting more robustness exercises was a lack of raw data.

After establishing the availability of data and codes, we then provide our replicators' evaluation of whether the article's results can be generated from the author-supplied materials (computational reproducibility) by method and field.

Last for this section, we provide details on our detection of (substantive) coding errors. We restrict this analysis to coding errors that, in the opinion of the replicators, have the ability to change the qualitative findings of the article, not minor ones such as those easily fixed in computation, directory, dependency, and the like.

3.1 Data and Code Availability

Our headline data availability finding is that for 81% of articles, replicators indicated they had access to either full raw or full analysis data. If only considering full access to raw data, the figure is substantially smaller, at 49%. Behind these overall numbers there is significant nuance to be found.

Table 1 provides all the details on variation regarding data and code availability in the 274 articles. Some notable findings are that 36% of replication packages do not include *any* raw data, 14% provide partial raw data, while the remaining 49% provide their full raw data. In contrast, 19% of replication packages do not include *any* analysis data, 3% provide partial analysis data, while the remaining 78% provide full analysis data. This gap between raw and analysis data availability suggests that current research transparency norms, at least in economics and political science, continue to prioritize reproducibility from analysis data over full disclosure of the research-data

pipeline. It is unsurprising to not reach 100% in any of these full access measures, as many recent and well-published papers leverage restricted or confidential data which cannot be publicly provided for privacy and other concerns (particularly in raw form). However, in these situations it may be possible to provide anonymized, aggregated, or similarly transformed analysis data which removes personally identifying information while retaining enough statistical information for computational reproduction. Similarly, it is unsurprising that data availability rates in our sample are quite high as our focus was mostly on articles with publicly available data.

In terms of code, full cleaning code (which conducts data cleaning, reshaping, variable definition, assembly of datasets) is provided markedly less often (54%) than full analysis code (which calculates statistics, produces tables and figures) which is almost always (95%). This lower provision of cleaning code remains true even conditional on providing any raw data, where cleaning code would be necessary, and is then provided at only a slightly higher rate (78%). These (relatively) high proportions of code provision, and in particular, higher code provision than data provision, is unsurprising given recent efforts towards computational reproducibility at many of the journals replicators' articles are chosen from.

Finding that less than half of replication folders provide raw data and cleaning codes, or about 4 out of 5 replication folders provide analysis data, provides context for the next section's examination of computational reproducibility. A simple linear regression does not detect a relationship between the number of robustness checks conducted with the availability of any full data ($\beta = 5.934, p = 0.688, n = 274$, also see Table 2). Notably, the estimated effect size is small compared to the mean number of robustness checks ($\mu = 73.84, n = 274$) and statistically insignificant. Similar null effects on the number of robustness checks are found for full raw data ($p = 0.358, n = 274$), full analysis data ($p = 0.447, n = 274$), any raw data ($p = 0.947, n = 274$), any analysis data ($p = 0.225, n = 274$), any raw data with cleaning code ($p = 0.879, n = 274$), and any analysis data with any analysis code ($p = 0.192, n = 274$). Of course, it is possible that the *character* of the robustness checks, rather than the number, may change with better data and code availability, however their number seems unaffected.

3.1.1 By Method and Field

We present data availability by method and field in Figure 1.

The top left panel presents data availability by field. The headline result is presented in black, where 81% of $n = 274$ articles provide either full raw or full analysis data. There is variation by field. For example, of the 49 reproduced articles that have a "C" JEL code (indicating they contribute to quantitative methods including econometric modeling, game theory, and design of experiments), 90% offer either full raw or full analysis data. As articles typically have more than one JEL code, an article can be included more than once in the figure, with the exception of the "all fields" proportion where every article enters only once. Articles in political science are assigned a unique and mutually exclusive category from American Politics, Comparative Politics, and International Politics. The estimate that 81% of articles provide full access to raw or analysis data can be interpreted exactly

as that - if a reader were to examine, at random, one of the articles in our sample, 4 out of 5 times they would have *full* access to either the raw or analysis data. Articles in economics tend to provide data at higher rates than political science, though variation in economics is evident; quantitative economics (JEL C) and financial economics (JEL G) provide high rates of data availability (90% and 93% respectively), which can be contrasted against the much lower 71% provided by the 34 articles contributing to public economics (JEL H).

The top right panel presents data availability by causal identification method. Just as with fields (JEL codes), some articles apply more than one identification method, so an article can be included more than once in the figure with the exception of the "all fields" proportion. We find this unit of analysis useful for reader interpretation; if an article has two distinct causal identification methods such as a regression discontinuity design (for example, distance to a border) and an instrumental variable (for example, a temperature inversion or rainfall), the reader can extrapolate that when an article uses an IV, they can expect full access to raw or analysis data 75% of the time. In addition, we find that articles leveraging a field experiment provide the second-best full access to data *in any form* at 88%, trailing only behind difference-in-differences (and two-way fixed-effects) which provide full access to data 90% of the time.

In the bottom panel, we present the intersection of field (columns) and method (rows) for every field $\times$ method combination. The first thing to notice is the breadth of field-method coverage afforded by the $n = 274$ articles reproduced by our replicators; we observe data availability for almost all subfields for DiD, IV, RD, and field experiments. Experiments, including lab and online, are more limited in subfield representation, which is likely explained by the fact that some subfields are not amenable to randomization (for example, economic history has no field, online, nor lab randomization in our sample). Each of these cells can be interpreted as, for example in the fifth column, 82% of articles that use DiD in public economics provide full access to some form of their data. In contrast, only 50% of articles that use a field experiment in public economics provide access to data, however please note that the dagger next to the proportion indicates that fewer than 5 distinct articles contribute to that proportion. The general pattern that emerges is that full access to some form of data, be it raw or analysis, is the norm in our sample of recent, well published, and selected towards publicly available data, articles. Darker cells indicate higher proportions of data access, while lighter cells indicate a visible pattern of less-than-full access in one of the political science fields (though we caution over interpretation as those cells have few observations which can be determined from the top left panel).

As previous research has found that access to raw data was the primary limiting factor to replicators' ability to probe the robustness of published results (Brodeur et al., 2026), we examine access to full raw data (irrespective of analysis data being provided) in the similarly constructed Figure 2. Unsurprisingly, full access to raw data is rarer, and particularly rare in articles using IV (39%). Standout fields for raw data include quantitative articles (59%), financial economics (61%), and industrial organization (56%). Each of the bottom panel cells can be interpreted as, for example in the sixth column, 62% of articles that use DiD in health & education economics provide access to raw data. In contrast, only 38% of articles that use a field experiment in the same field provide access to

raw data - likely due to the aforementioned privacy reasons.

3.2 Computational Reproducibility

Computational reproducibility is the ability of a replicator to reproduce the findings of an article from that article's replication package, including its data (either raw or analysis) and code (either cleaning or analysis).

Our headline computational reproducibility result is that 89% of articles are considered to be fully reproducible from either raw or analysis data by our replicators (see also Table 1). Of course, this overall proportion is affected by availability of data, and belies significant variation by method and field.

The proportion of articles considered fully reproducible from raw data (45%) is fairly close to the proportion of articles that provide complete raw data (49%), while remaining less than half that of those considered fully reproducible from analysis data (89%).

3.2.1 By Method and Field

We present computational reproducibility by method and field in Figure 3.

The top left panel presents the proportion of articles that replicators considered to be computationally reproducible on average (89% for the 274 articles in the pre-registered fields). Computational reproducibility is high in economics, with the worst-performing fields, macroeconomics and public economics, still achieving a computational reproducibility rate of 76% and 82%, respectively. This means that one in four macroeconomics and one in five public economics articles, published recently in top journals, are not computationally reproducible. Industrial organization, which also was a leader in raw data availability, is the most computationally reproducible field in economics. Political science sees the greatest variation. 94% of comparative political science is indicated as being computationally reproducible, but only 60% (3 of 5) international political science articles are computationally reproducible.

The top right panel presents computational reproducibility by causal identification method. Once again, computational reproducibility is high; on average at 89%. The 'worst' performing method, with a computational reproducibility of 83%, is regression discontinuity. The standout method, which achieved a computational reproducibility of 100%, is lab experiments. To be specific, for the $n = 14$ distinct articles that used lab experiments, when independent replicators ran their data and code, every team of replicators was able to reproduce the original article's published results. Second is online experiments at 96%. Third is field experiments at 94%. The quasi-experimental methods all trail the experimental methods - DID, IV and RD, all achieve relatively worse (though still well performing) rates of 92%, 89%, and 83%, respectively.

In the bottom panel, we present the intersection of field (columns) and method (rows) for every field $\times$ method combination. The general pattern that emerges is that computational reproducibility is high (darker colors) for most, if not all, combinations. Indeed, for many of the field $\times$ method

combinations, computational reproducibility is at 100%, and not just for cells with less than 5 reproduced articles (indicated by the superscript daggers), whereas the lightest colors, representing international politics using online experiments or public economics using a field experiment, have fewer than 5 articles, making it difficult to draw conclusions about them.

3.3 Coding Errors

We next examine the incidence of coding errors (importantly, and in light of the previous section, the fact that an article is computationally reproducible does not preclude the presence of even major coding errors, as our replicators have proven to be sophisticated researchers capable of diagnosing and repairing original code). This analysis excludes minor issues—such as software version conflicts or missing dependencies. Instead, we focus on errors with the potential to affect the original article's empirical results or substantive conclusions.

Our headline result is that we find that 32% of articles contain a coding error in their replication package which has the potential to change the substantive take-away from the original article.

Table 3 provides details on the types of coding errors detected by the replicators, the most common types are related to control variables (10%)(i.e. a mismatch between article and code in the control variables included) and changing the sample (11%). These are followed by changing estimation procedures (7%). The least common coding errors and inconsistencies are those related to the dependent variable (5%), main independent variable (4%) and inference (4%). The table also provides tests of the differences of the coding errors. When comparing economics and political science, there is no general trend of the coding error types being more likely to be detected in one or the other (one p-value is 0.055 for the difference between the proportion of 'control variables' coding errors). When comparing experimental methods (field, online, lab experiments) to quasi-experimental methods (DID, IV, and RD), there is again no statistically significant difference at that high level (p-values range from 0.063 through 0.752).

3.3.1 By Method and Field

We present coding error rates by method and field in Figure 4.

The top left panel presents the proportion of articles that replicators considered to have a coding error by field. On average, 32% of articles contained a coding error. The best-performing field in economics, i.e. the JEL code with the lowest coding error rate, is macroeconomics with 28% of articles. A similarly honorable mention is JEL H, public economics. In political science, coding error rates are relatively lower with American politics (10%), Comparative politics (26%), and International politics (40%), two of which perform better than the total sample average.

The top right panel presents coding error rate by causal identification method. The best performing method is the lab experiment (14%), while the remainder of methods cluster around the 33% average.

In the bottom panel, we present the intersection of field (columns) and method (rows) for every field $\times$ method combination. We highlight the fact that, since a lower number is desirable for

coding errors as compared to higher numbers for data availability or computational reproducibility, we now present the field and method combinations in red rather than blue cells. In this way, more red cells, such as quantitative articles using DID, financial articles using RD, or labor articles using IV, with high coding error rates are easy to spot. The general theme is that coding errors are concentrated within a narrow range of cells, with many zeros and low numbers. Indeed, the darkest cells, and many of the ones indicating medium levels of coding errors, are indicated to be from combinations of field and method with fewer than the requisite 5 articles for any form of confidence.

Overall, these findings suggest that coding errors can be found in any field or method, with high error rates in some well-populated fields (e.g. microeconomics). Further, these results underscore that data availability or computational reproducibility are not by themselves sufficient for correctness of published empirical analyses. Measuring correctness seems to require exactly what this project conducted - a deep engagement with each component article by teams of researchers familiar with the methods, literatures, and programming languages in which the original articles were written.

4 Robustness by Method and Field

In this section, we detail our main results for robustness by method and field.

Two features are required for robustness - the re-analysis test statistic must also be statistically significant at the 5% level, and must also be of the same sign.[17] We provide additional measures of robustness in the next subsection, including those that are appropriate for originally statistically insignificant results.

We find that research in economics and political science is 70% robust. Our 70% rate of robustness is relatively high in comparison to previously published replications.[18] For example, when evaluating the replicability of 18 laboratory experiments in economics, Camerer et al. (2016) found a significant effect in the same direction as the original article for 11 of them; a replication rate of 61%. Another example, when evaluating the replicability of 21 social science experiments published in *Nature* and *Science* between 2010 and 2015, Camerer et al. (2018) found a significant effect in the same direction as the original article for 13 of them; a replication rate of 62%. Of course, the present project is different in both scope and size. We primarily conduct robustness exercises (using the provided data and codes of original articles and testing whether the original results stand up to alternative, yet reasonable, researcher decisions) whereas the previous two examples recreated the

[17] The current robustness analysis uses results from 187 articles only. This is due to at least three reasons. First, some of the reports do not conduct robustness checks. Second, a direct comparison between the original and re-analysis is sometimes not possible. This can arise when replicators use a different estimator or rescale the dependent or main independent variable. In other cases, the methods used in the original paper do not produce coefficient estimates or p-values as the primary quantities of interest, making a direct comparison infeasible. Third, only 216 reports contain a robustness check of at least one statistically significant result at the 5% level. This restriction matters as measuring the robustness of an originally statistically *in*significant result is likely to behave very differently (especially with respect to flipping sign) than one that is originally statistically significant.

[18] When excluding replications (therefore keeping robustness checks, strictly defined) the robustness rate is an unchanged 70%.

experimental conditions of the original conditions and re-ran the experiments. The closest project to ours is indeed already part of our project; when examining the robustness of 110 recently and well published articles published in economics and political science journals from 2022 onward, Brodeur et al. (2026) found a robustness rate of 72%. We provide more comparisons to that first project in a later dedicated section. Nonetheless, we take as a positive sign for both the economics and political science fields that the robustness rate is as high as it is.

4.1 By Method and Field

We detail the robustness of economics and political science by field and method in Figure 5.

The top-left panel first provides an estimate of the overall robustness rate of 70% (black marker and 95% confidence interval whiskers). Successive rows provide robustness rates by JEL code.

For articles that had JEL code E (macroeconomics), the average of the article-level robustness rates is 52%, the worst performing subfield in economics. In contrast, the best performing subfields (in economics) are quantitative, micro, and health and education economics (JEL codes C,D, and I), with much higher 74% robustness rates. When considering political science, the worst performer is comparative politics at 67% while the best is American politics with a very-high 84% robustness rate.

The top-right panel first provides an identical overall robustness rate, then successive columns provide robustness rates by method. We have grouped the quasi-experimental methods (DID, IV, and RD), and then present rates by the distinguishable experimental methods with randomization (field, lab, and online experiments). All other methods are assigned the 'other' category.

Just as articles can (and tend to) have more than one JEL code, so too can articles have more than one identification method assigned to their robustness checks. This is why it is possible for an article to enter into both, say, the IV and DID estimates if, for example, in one section of the article a two-way fixed effects approach is used, and later, an instrument is used to address potential endogeneity.

The methods with the highest robustness rate are experimental. Online experiments are the most robust with a robustness rate of 89%, followed by field experiments with a rate of 79% and lab experiments with a rate of 77%. The quasi-experimental methods follow with on-average lower robustness rates, and in the order one might 'expect' from the 'folk wisdom' of the profession (Brodeur et al., 2020); RD is the best performing non-experimental method with a robustness rate of 76%, whereas IV has a rate of 70% and difference-in-differences has the lowest rate of all labeled identification methods at 66%; that is, when it comes to articles applying difference-in-differences (or two-way fixed effects) *more than one in three articles* are not robust to alternative reasonable researcher choices, on average.

The bottom panel of Figure 5 is a heat map which provides information on the individual method (rows) and field (columns) combinations. Each cell contains the percent average of the article-level average of the robustness of the individual test statistics indicated in white text. That is, the average of the articles with JEL code D (microeconomics) that use DID is 71%. Darker blues

represent higher robustness rates, while lighter colors indicate lower. We have provided, next to the text of each cell, a superscript dagger symbol if the robustness rate comes from a cell composed of an average of less than 5 unique articles (to prevent conclusions being drawn about a particular method $\times$ field combination without sufficient evidence). We do remind the reader that the cells in the heat map are not evenly distributed in terms of sample size.[19]

4.2 Other Measures of Robustness

4.2.1 Robustness of Insignificant Results

In Figure A3, we provide the robustness of originally statistically insignificant results. One difference between the definition of robustness here as compared to the previous subsection is, with results that are originally statistically insignificant, the likelihood of re-analysis estimating an opposing sign is much higher than if re-analysis is examining an originally statistically significant result. For this reason, robustness is defined as remaining insignificant ($p > 0.05$) without regard to sign here. See Figure A4 for the 10% significance level.

The robustness of statistically insignificant results is 88%. That is, when replicators probe the robustness of 'null' results to reasonable alternative choices, the null results tend to stay null. The rate is well above the robustness rate of statistically significant results, and the sample is not small (141 reports).

Most fields provide similar rates of robustness for insignificant results, though macroeconomics and economic history have lower rates (79% and 72% respectively) but also have the smallest sample size. Similarly, many of the identification methods behave similarly with high rates of robustness for originally statistically insignificant results with one exception. Instrumental variables, well below the average of all others, with 82% - particularly when compared to the next-lowest identification strategy of DID with 88%.

4.2.2 Relative t-Statistic

In Figure A5, we present the relative t-statistic of Dreber and Johannesson (2025) by method and field. A value of 1 here implies that replicators find statistical significance, measured via the t-statistic, exactly the same on average as the original article. This analysis is restricted to the same set of articles as in Figure 5 (originally statistically significant results) for a number of reasons, the least of which is that a ratio with a denominator value near zero can easily explode.

We find an overall relative t-statistic of 92%. The field with the lowest relative t-statistic is macroeconomics (68%) while the economics sub-field with the highest relative t-statistic is public

[19] Our analysis throughout is at the article-level. An alternative approach would be to examine reproducibility and robustness at the robustness check level. When doing so, we find similar results. While the overall robustness rate is 77%, the experimental methods dominate the quasi-experimental methods, with online experiments reaching a robustness rate of 92%, lab experiments reaching a robustness rate of 88%, and field experiments reaching a robustness rate of 82%. In contrast, tests derived from RDs are 70% robust, followed by DID with 69%, and the worst performing of the identified methods is IV with the lowest rate of 67%. These methods results map closely to the article-level results. See Figure A1. See Figure A2 for the 10% significance level.

economics (124%). Political science performs well under this measure. American politics achieves a relative t-statistic of 93% and comparative politics of 87%. Methods have similar variation to sub-fields. Lab, online, and field experiments (almost) dominate the quasi-experimental methods. The t-statistics of lab experiments are 95% the size of the original, online experiments are 95%, and field experiments are 91%. In contrast, while relative t-statistics of articles with an RD is 92%, the ratio falls to 84% for IV. The anomaly is DiD, with a very wide confidence interval, but with an average of 115%. The bottom panel displays the ratios for each field and method combination, in the form of a ratio so that, for example, the 0.83 of JEL G (financial) economics using DID represents a relative t-statistic of 83%.

4.2.3 Relative Effect Size of Significant Results

In Figure A6, we present the relative effect size by method and field. A value of 1 here implies that replicators find effect sizes exactly the same on average.[20] In order to ensure comparability between this exercise and the ones previous, we restrict to the same set of articles as in Figure 5 and Figure A5 (originally statistically significant results).

We find an overall relative effect size of 90%. The field with the lowest relative effect size is macroeconomics (72%) while the economics sub-field with the highest relative effect size under robustness is quantitative (93%). As with the relative t-statistic, political science performs well under the relative effect size measure. American politics achieves a relative effect size of 90%, the middling is comparative politics of 92%, and international politics finds the best relative effect size of 95%. Methods have seemingly more variation than fields. Most methods cluster around a relative effect size of 100%, such as IV (102%) and RD (94%), while DiD relative effect sizes are sizably smaller under reproduction at 89%. The largest and similar to IV is the lab experiment at 102%. The field experiment is at 97% and the online experiment achieves 96%. The bottom panel displays the ratios for each field and method combination, in the form of a ratio so that, for example, the high IV measure can be seen as a consequence of micro and development economics using IV with a relative effect ratio of 1.13 and 1.12, which is the result of more than 5 distinct original articles (no daggers).

5 Determinants of Robustness: A Many-Analysts Approach

5.1 The Many-Analysts Approach

To complement our analyses, in this section, we adopt a many-analysts approach. The goal of this exercise is to assess whether independent analysts, working from the same data, arrive at consistent conclusions when analyzing the relationship between robustness by method and field, with freedom to account for a plethora of potential modeling decisions including many article, author,

[20] To put the following magnitudes into context, Camerer et al. (2016) found a replication relative effect size of 66%, Camerer et al. (2018) found a replication relative effect size of 50%, while Brodeur et al. (2026) found a reproduction effect size of around 100% (though the latter reported the median).

and replicator characteristics. Individual analysts were recruited from our pool of coauthors. Each participating analyst was compensated with 250 USD for timely submission of results.

Analysts were free to choose their own estimation techniques, weighting scheme, statistics package, model specifications, and inference methods used, but they were not permitted to alter the dependent variables or the key independent variables (methods and fields) to ensure comparability across submissions. Indeed, they were provided with a common dependent variable: an indicator for whether a re-analysis remains statistically significant at the 5% level with the same sign, restricted to original estimates significant at the 5% level.[21] As the goal was to continue investigating the *relative* rankings of robustness, the analysts were asked to use JEL code D (microeconomics, behavioral and experimental economics) as a baseline for sub-fields and the field experiment as baseline method. Analysts were also free to construct their own derived variables from the data given to them or use new data. The database includes detailed information on replicators, original authors, article, and replication package characteristics. Each analyst was instructed to report marginal effects for their specifications. To accommodate multiple approaches, the submission system allowed analysts to report multiple models, with each specification treated as a distinct entry. Of note, analysts were not made aware of any of the present paper's findings, including the expert forecast results. Any findings or forecasts were reported to analysts and other coauthors by the lead authors only once the many-analysts had already reported their marginal effects.

5.2 Many-Analysts Results

We received completed submissions from 13 many analyst teams. The heterogeneity in analytical choices provides a useful measure of the robustness of our conclusions: if findings converge across analysts despite specification differences, this strengthens the confidence we can have in the relative rankings presented in the earlier sections. The results are shown in Figure 6.

In the top two panels, each marker represents a many-analyst's t-statistic for that coefficient. For example, 12 analysts provided comparable coefficients for JEL code C. The t-statistics ranged between 0 to 4, with an average of 1.37 (black diamond, 95% confidence intervals in whiskers, $p = 0.001$ against the null hypothesis the mean is equal to zero). Similarly, 12 analysts provided estimates for differences-in-differences, with coefficients' associated t-statistics mostly negative, with an average of -5.18. For this analysis, we present t-statistics as independent analyst heterogeneity is more likely to influence the scale of the coefficients (making it difficult but not impossible to draw conclusions from naive comparisons of the coefficients but much less so the signal given by the associated t-statistic).

Across methods, a clear consensus emerged. Analysts consistently found that robustness rates were highest for experimental designs, particularly for online experiments, but also for field experiments (the baseline or comparison category) and laboratory experiments, as compared to statistically significantly lower rates for the quasi-experimental approaches (DID, IV, and RD), each with

[21] The many analysts also prepared results for three other supplementary outcomes. The second dependent variable is defined analogously for the 10% significance level. The third and fourth dependent variables capture whether re-analyses of originally insignificant results remain insignificant at the 5% and at the 10% level, respectively.

a negative statistically significant average t-statistic. In most independent analyses, the baseline category (field experiment) outperformed non-experimental designs by a wide margin, and the direction of the effect was highly consistent across analysts. Although some variation was observed in the precise magnitude, very few analysts reported estimates that contradicted this overall pattern.

Across sub-fields, the results also showed broad agreement. Economics sub-fields such as Public Economics (JEL H), Mathematical and Quantitative Methods (JEL C), Industrial Organization (JEL L) have relatively high robustness rates, while Macroeconomics (JEL E), Labor and Demographic Economics (JEL J), Economic History (JEL N) and Economic Development, Innovation, Technological Change, and Growth (JEL O) had relatively low robustness rates. Health, Education and Welfare (JEL I) and Financial Economics (JEL G) were in-between the lowest and highest robustness rates, and not significantly different from the baseline. Within political science, if robustness rates differed from economics, they were relatively higher.

The bottom panels present a similar pattern using the effect sizes of the associated coefficients reported by the many analysts. As expected, effect size estimates tended to have wider relative variance than the t-statistics of the top panel. For this reason, the bottom panels display the median effect size with a hollow diamond (to account for the potential outsize influence of outliers) while also providing a visualization (line and bar) of the inter-quartile range. In this manner, we can visually detect outliers well outside the inter-quartile range, and note the following. For methods, the inter-quartile range of effect sizes for the quasi-experimental methods are negative, while the lab experiment, field experiment (baseline), and online experiments (in particular) perform better. For fields, JEL C, H, L and American political science seem to perform quite well.

Overall, the many-analysts exercise demonstrates that the conclusions of this study are not an artifact of a single modeling approach. Despite employing a diverse set of specifications, many independent analysts found similar results: robustness rates do indeed systematically vary by method and field, with experimental designs exhibiting the highest levels of robustness. Where variation exists, it largely concerns the size rather than the direction of the estimate, further adding confidence of the stability of our main conclusions.[22]

6 Predicting Reproducibility: Expert Forecasts

In this section, we examine the difference between the previous sections' *observed* reproducibility and robustness of research by field and method to the *expected* reproducibility and robustness by the research community itself. To do so, we administered a pre-registered Qualtrics-based survey to experts in social science research on the Social Science Prediction Platform (SSPP). There are at least two benefits to expert forecasts in social science research - including the fact that this project gains more nuanced interpretation of its results through the direct comparison between observed and

[22]While this section's many-analyst exercise precludes the necessity for a single regression (and its potentially idiosyncratic choices, Table A3 column 4 conducts such a linear probability model with field and method regressors to find similar a result; that the robustness rate for field, lab, and online experiments are statistically significantly higher than the other (excluded) category. The prior columns display results for the provision of full data, computational reproducibility, and coding error prevalence in the same manner as in Sections 3 and 4.

expected reproducibility rates. Forecasts have also been championed for their ability to mitigate bias against null results (by facilitating comparisons between observation and expectation, rather than against the null hypothesis) and, less so for this project, for improving experimental design (by, for example, providing guidance on which treatment arms should be applied in a power-constrained setting) (DellaVigna et al., 2019).[23]

For our purposes, expert forecasts provide a valuable window into how well experts can anticipate the reproducibility and robustness of empirical findings within their discipline (Dreber (2025); Dreber et al. (2015)). By asking researchers to predict which sub-fields and methodological approaches are most likely to yield robust results, we can test whether expert expectations align with empirical observation. Accurate forecasts would suggest that researchers have internalized reliable cues about reproducibility and robustness, including through experience with data quality, research designs, and publication practices. Conversely, systematic forecasting errors would indicate that researchers struggle to identify robust research ex ante.

Our survey (https://socialscienceprediction.org/s/xeb2c2) first provides definitions and a description of our data and its generating process, then asks forecasters to provide their predictions and judgments concerning computational reproducibility, coding errors, and robustness (defined in numerous ways, including the same-sign and preserved statistical significance of earlier, along with the re-analysis t-statistic and proportion of effect sizes that would be smaller under re-analysis). Most questions ask about the entire sample, while the robustness reproducibility questions were asked by field and method.[24] All survey questions were preregistered. The survey ends with questions on the field, affiliation, and position of the forecaster.

For computational reproducibility, we ask forecasters: "What percent of studies do you expect to be fully computationally reproducible?" They were provided the following definition: "Results/tests which are not computationally reproducible are those that fail to be reproduced with the replication folder provided. Common reasons include: code (cleaning or analytic) failing to run, results deviating from what was presented in the published paper, and data and code only being partially provided."

For coding errors, we ask forecasters: "Excluding minor issues like missing packages, broken pathways, and versioning issues, what percent of studies do you expect to have at least one coding error?" Where they were provided the following additional information: "Note that not all coding errors necessarily result in a failure to reproduce the claims presented in a published paper. A study could be fully reproducible, yet its programs contain coding errors (e.g. observations erroneously duplicated), or coding which doesn't match the description in the paper but keeps main results unaffected (e.g. clustering at different levels)."

For robustness, we ask forecasters: "What do you expect the robustness reproducibility rate

[23] Recent work on the first 100 projects on the platform has found that forecasts of treatment effects (the platform is often used prior to experimental research) tend to be quite predictive of actual treatment effects (DellaVigna and Vivalt, 2025). Earlier work notes that expert forecasts were correlated with, but outperformed, literature (meta-analytic) based estimates of treatment effects (DellaVigna and Pope, 2018).

[24] This is a deviation from our pre-analysis plan in that we only asked the robustness reproducibility question by method and field. The decision to limit the remaining questions by subfield and method was taken after considering the length of the survey and potential forecaster fatigue.

across claims to be for the following methods and subfields?" They were provided with the following definition: "robustness reproducibility is the extent to which a result in an original study remains robust to alternative plausible analytical decisions, while using the same data as in the original study. In our case, we say a result/claim is robustly reproducible if the re-analysis is statistically significant with the same sign (positive or negative) as the original study's statistic. If you answer 50, this means you believe that 50% of the original analyses are robust to various types of re-analyses recoding (redefining the sample, dependent variable, etc.)"

Rates were elicited for nine methods, two fields (economics and political science in general), 13 sub-fields for economics and three sub-fields for political science.[25]

6.1 Recruitment and Participation

The survey was open on SSPP from September 30, 2024 to October 31, 2024, closing with 249 completed responses (out of 303 responses).[26] Invitations were also sent to coauthors who had completed a reproduction report as of Fall 2024. The survey was also advertised on one of SSPP's social media accounts and reposted by I4R.

The survey was conducted in *Qualtrics* with a forced response for all questions about reproducibility and robustness. An opt-out option "not familiar" was always offered, which accounts for the differing response rates by field and method. For instance, 16 forecasters did not provide a forecast of the robustness of instrumental variables. The median time to complete the survey is 8 minutes and 12 seconds. We do not make any sample restrictions based on completion time in our analyses.

The sample consists of 44% graduate students, 11% post-docs, 30% faculty members or researchers at a university and 15% other (e.g. researcher in a think tank or government agency). Forecasters are mostly economists (70%) and political scientists (11%). 73% of forecasters attended one or more of I4R's replication games (Brodeur et al. (2023, 2026)), suggesting that most our forecasters have experience reproducing studies. 35% of forecasters also report having heard about some of the preliminary results from Brodeur et al. (2026) combining 110 reproduction reports, even if only informally.

[25]The methods are (i) difference-in-differences, (ii) instrumental variables, (iii) laboratory experiment, (iv) matching and synthetic control, (v) online experiment, (vi) field experiment, (vii) regression discontinuity design, (viii) time series, and (ix) structural modeling. For economics sub-fields we used (i) microeconomics, behavioral and experimental economics, (ii) macroeconomics and monetary economics, (iii) international economics, (iv) financial economics, (v) public economics, (vi) health, education and welfare economics, (vii) labor and demographic economics, (viii) law and economics, (ix) industrial organization, (x) economic history, (xi) economic development, innovation, technological change, and growth, (xii) political economy and comparative economic systems, (xiii) environmental and urban economics. The three sub-fields for political science are (i) American politics, (ii) comparative politics and (iii) international relations. We restrict presentation in this section to the pre-registered sub-fields and methods. Additional results, for example for law and economics, international economics, and environmental economics, are available upon request.

[26]Staff at SSPP mentioned that our survey was "one of our most highly responded surveys this year". Only forecasters who completed the entire survey are included in our analysis.

6.2 Forecast Results

Our pre-registered hypothesis is that researchers accurately predict robustness rates by method and sub-field.[27] We present our overall forecast results in Figure 7. The distribution of forecasts for each of these reproducibility measures is presented in the figure as grayscale kernel densities, to visualize the variation of forecasts, for example, we note none are uni-modal and all are characterized by high variance. For each measure, the average of the 249 predictions (and its 95% confidence interval) and the observed rate are presented, along with the results of a statistical test of their difference.

We find that forecasters are, on average, overly pessimistic. Forecasters underestimated computational reproducibility. They predicted that only 68% of articles would be fully computationally reproducible, while our observed proportion is 89% (Wilcoxon signed-rank test of the difference $p < 0.001$). In contrast, forecasters overestimated the rate of coding errors, expecting 51% of articles to have a major coding error, as opposed to the observed proportion of 32% (Wilcoxon signed-rank test of the difference $p < 0.001$).[28] Forecasters underestimated the robustness of initially statistically significant effects, expecting that only 59% would be robust whereas we observed a 70% robustness rate (Wilcoxon signed-rank test of the difference $p < 0.001$). In terms of additional measures of robustness, forecasters underestimated the relative t-statistic under reproduction, expecting only 76% as compared to the observed 92% (Wilcoxon signed-rank test of the difference $p < 0.001$). Also, forecasters slightly overestimated the proportion of effect sizes that decreased in absolute magnitude under re-analysis, predicting 58% compared to the observed 54%, a difference of four percentage points.[29]

Figure 8 shows that forecasters systematically underestimate the robustness of research across both methods and fields.

The figure is composed of a series of densities from 0 to 100, indicating forecasters' forecasts to what they expect the robustness rate of, for example, studies classified as JEL D (Microeconomics) (64%) or those using a field experiment as their identification method (73%). Should a forecaster select, say, 61, they are indicating they expected 61% of studies using an RD somewhere in their analyses to be robust. The distributions of these forecasts for both methods and fields are as reporting only the mean forecast (indicated by a black hollow circle with 95% confidence interval) hides much of the variation that is evident within each distribution. For example, the mean forecasts of RD and Online Experiments are the same at 61% - which belies the fact that the variance for RD is smaller - which hints at varying beliefs of the robustness of online experiments. In addition to the number of forecasters who offered their predictions for each method and field, we present the statis-

[27] Forecasters are reasonably confident in their predictions. The median forecaster is "somewhat" confident. The remaining responses are distributed as follows: 32 report not being confident at all, 87 report being slightly confident, 102 report being somewhat confident, 21 report being very confident, and 7 report being extremely confident. We consider this to be a good thing since recent research has shown that confidence is inversely related to accuracy (DellaVigna and Vivalt, 2025).

[28] These two differences in expectations may arise from the top-tier sample of journals we draw our reproduction articles from, as many of those journals conduct thorough internal computational reproducibility checks prior to publication.

[29] For smaller under re-analysis effect sizes, we calculate the proportion of comparable re-analysis estimates that are smaller in absolute magnitude than the corresponding original estimates. We first calculate this proportion within each article, restricting attention to originally statistically significant estimates, and then average across articles.

tical significance of a Wilcoxon signed-rank test comparing forecasts to the observed reproducibility rate (black diamonds with 95% confidence intervals).

In the left panel and center panels of Figure 8, we see that forecasters consistently under-estimate the robustness of each field except for macroeconomics, where they overestimate the robustness of the field in their forecasts. Some fields (JEL I for example) show larger differences between expected and observed robustness than others. We find that the difference between forecast and observed robustness rate is always statistically significant (with the exception of JEL J Labor).

In the right panel of Figure 8, we see that the average robustness rate forecast falls below the realized values for all methods, confirming the overall pessimism among researchers. While the magnitude of these errors varies– field experiments tend to be better predicted than for example, online experiments – the difference between robustness forecast and observed robustness is always statistically significant well below any traditional inveterate statistical significance threshold (with the exception of field experiments where $p = 0.054$).

In order to understand this phenomenon better, we sort and present these average forecast errors in Figure 9. The left panel provides results using the absolute mean error (averaging the absolute value of each forecaster's error against the observed rate), with lowest error at the top and increasing error toward the bottom of the figure. The figure contains both methods (red) and fields (green and blue).

Reading from top to bottom, an earlier observation becomes clearer; the best predicted method is the field experiment (which was expected to perform the best of all the methods) followed by the quasi-experimental method DID, then laboratory experiments and RDs. The best predicted fields are those in economics, with the notable caveat that it is unclear whether this reflects greater pessimism by the forecasters (the predictions for political science fields are lower than economics) or greater robustness (the observed reproducibility of the political science subsample is higher than economics) or, more likely, both. These rankings are relatively robust to an additional, alternative, but sensible construction using mean (not absolute) error for the forecasters, presented in the right panel. There, experimental methods take the top spots (with the field experiment placing first, lab experiments are second) and most economics sub-fields but JEL I (health, education) are better predicted than political science.

7 Statistical Significance after Reproduction: Extension of Brodeur et al. (2026)

This section begins as a replication and extension of Brodeur et al. (2026). That project, among other results, showed how the distribution of statistical significance changes under replication and reproduction (in particular, by the *type* of re-analysis) for our first 110 articles. We then provide results of how statistical significance changes using a dataset more than double the size. We then examine these changes by method and field.

The top panels of Figure 10 display the distribution of original absolute t-statistics if they are

sourced from the first project (left, panels detailed in Brodeur et al. (2026)) or from sample of new articles, which more than doubled the number of original articles we examine (right panels). We have included a kernel density, with 95% confidence intervals in orange, for the distribution of the associated t-statistics after re-analysis (either reproduction or replication).

Comparing the distributions of original statistical significance, they look similar. They both have (at least a local) maximum at-or-above a t-statistic of 2, consistent with statistical significance at the 5% level. Further, they both have a local minimum just after a t-statistic of value 1. In the left panel, re-analysis seemed to 'recover' an excess mass of test statistics above the statistical significance threshold and re-distribute them fairly uniformly between $t = 0$ and $t = 1.96$. A reduction in that local maximum can be seen, and the local minima is also fairly smoothed. In the right panel, a similar pattern occurs, though the excess mass seems to come more from larger t-statistics clustered around 3 than before, and the local maxima and local minima remain, although their difference is significantly reduced under re-analysis.

The bottom panels present the associated p-values, with grayscale representing all fields and methods for the 110 original articles in the left panel, and newly added articles on the right. While the p-curve (distribution of p-values) is a mechanical counterpart of the t-curve (distribution of t-statistics), certain characteristics consistent with both the publication bias and p-hacking phenomena are apparent in this form. For example, there is a large 'spike' or increase in mass just below the $p = 0.05$ in the original p-value distribution. Under re-analyses, the p-curve is instead, monotonically decreasing. For new articles, re-analyses less dramatically altered the distribution around the 5% statistical significance threshold, but did reduce the large amount of very small p-values represented by the first bin.

These visual patterns of shifting statistical significance will be presented by field and method at the end of the next section.[30]

7.1 By Method and Field

This project represents a significant increase in sample size over an earlier study, Brodeur et al. (2026), which analyzed the re-analyses of 110 published articles. As this project's aim is to investigate the reproducibility of research by field and method, a significant expansion was necessary to ensure sufficient evidence for each field and method is available. To do so, we extend that original sample of 110 to now include the results of re-analyses of 274 articles.

Figure 11 presents the distribution of t-statistics first for the entire sample of 274 pre-registered pre-analysis reproductions (in gray). The solid area represents the original t-statistics while a black Epanechnikov kernel (with 95% confidence interval) represents the distribution of re-analysis t-statistics. We see that, under reproduction, the local minima located near a t-stat of 1 is greatly reduced.

[30] Formal tests of p-hacking and publication bias using the caliper method, are presented in Table A2. The threshold is the inveterate 5% statistical significance threshold, and the test assumes that there should be equal 'mass' of test statistics above and below the threshold.

We then present the same two distributions by field and method in the succeeding rows and columns. Following the convention of the rest of this paper, economics sub-fields are presented in green, political science sub-fields are presented in blue, while methods (drawing from articles in either field) are presented in the rightmost column in red.

What is obvious is that this shift in statistical significance differs, sometimes strikingly, by method and field. First, the characteristics of the original article significance differ. For example, economic history has a large local maximum around $t = 2.5$ with very little mass to the left of that maxima. In contrast, health and education economics has a large mass (indeed a global maximum) around $t = 0$, that is, statistically insignificant results. Of course, this can be driven by a number of factors, not the least of which could be the underlying distribution of unknowable 'true' effects that each of the articles was attempting to estimate. Nevertheless, under re-analysis, the economics sub-fields see a general smoothing of the t-distribution with smaller local maxima around $t = 2$ with much of the mass under re-analysis recovered and distributed between $t = 0$ and $t = 2$. Re-analysis for the three political science sub-fields behaves slightly differently (which was also the case in Brodeur et al. (2026)). Under re-analysis, the test statistics do not seem to respond as much as the economics sub-fields.

Concerning methods, originally published test statistics for the quasi-experimental methods of DiD, IV, and RD have very large, almost unimodal peaks just above $t = 2$. Under re-analysis, these three methods show large amounts of mass recovered below the $t = 2$ threshold. In comparison, the three forms of experiments (field, lab, online) have very close (if not perfectly overlapping) distributions of test statistics after re-analysis.

Figure 12 distills and presents how statistical significance changes from the originally published results under re-analysis at the test-statistic level. For each, the vertical axis represents whether the original test statistic was statistically significant or not. The horizontal axis represents whether that test statistic, under re-analysis was found to be statistically significant at the 5% level, not statistically significant, or if the sign of the re-analysis estimated effect was of the opposite sign to the original. The numbers in each represent the proportion of tests by row. For example, in the top left grayscale panel (indicating all fields and methods) for those test statistics that were originally significant at the 5% level, 72% remained statistically significant (and did not switch sign). In contrast, 6% flipped sign (regardless of statistical significance) while the remaining 22% were no longer statistically significant. In the second row of the panel, the same transitions of the rows are noted, with 77% of originally statistically insignificant results remaining statistically insignificant under re-analysis (and without flipping sign). Unsurprisingly, a larger proportion of these results flip sign than statistically significant ones (though this is likely mechanical due to increased proximity that a statistically insignificant result has to $t = 0$) and a smaller proportion (8%) become statistically significant. For these proportions, we take as evidence of robustness the relative proportions on each matrix's diagonal - the proportion of statistically significant results that remain statistically significant and the proportion of statistically insignificant results that remain the same.

The successive rows and columns present the same transition shifts by field and method. There is large variation in the proportions of interest. For example, macroeconomics (JEL E) has the lowest

proportion of robust statistically significant test statistics at 48%. In contrast, the highest is political science's American politics at 82% (the best-performing sub-field of economics is 79%). In order, the best performing methods are online experiments (88%), lab experiments (79%), field experiments (78%), then RD (77%), IV (70%), and finally DID (67%) (the other category, due to its broadness, we are careful not to over-interpret).

8 Conclusion

Understanding which methods and fields produce robust findings is critical for both scientific progress and its policy application. Robust research shapes the credibility of entire disciplines in the eyes of scholars, policymakers, and the public. If certain identification methods systematically yield more robust and reproducible results, this can inform the direction of future researchers, the prioritization of methodological training, and policymakers' allocation of attention to research toward the most robust approaches. Similarly, identifying sub-fields with lower robustness rates highlights areas where stronger data practices, improved transparency, or more stringent editorial standards may be needed. Overall, our findings suggest that the credibility revolution is not uniform, but is better characterized as a heterogeneous process, with substantial variation in robustness across methods and sub-fields. This has important implications for how empirical research is evaluated in the future: rather than applying uniform optimism following the credibility revolution, expectations regarding robustness and reproducibility can, following the implications of this project, be better calibrated to the underlying research method and field.

Our expert forecast survey highlights a consistent gap between researchers' expectations and actual reproducibility outcomes. While forecasters correctly anticipated relative patterns such as higher robustness for experimental methods than for quasi-experiment methods, they systematically underestimated computational reproducibility, overestimated the prevalence of coding errors, and underestimated robustness overall. The gap between expert forecasts and observed outcomes suggests that researchers may hold systematically pessimistic beliefs about reproducibility, highlighting a disconnect between perceived and actual credibility. These discrepancies suggest that researchers, especially those in fields other than economics and politics, may hold somewhat pessimistic priors about the credibility of published work, on which we hope the results contained herein will provide a welcome update (with the caveat that our sample represents well-published and recent empirical work).

More broadly, we believe that this large-scale reproducibility initiative has the potential to reshape research practices and publication norms in meaningful ways. By providing systematic evidence on which methods and sub-fields yield more or less robust findings, it can guide the academic community of journals, editors, and referees toward a more structured evaluation process. For example, the traditional referee system could be complemented by (some) reviewers tasked specifically with assessing replication packages, scrutinizing code, and ensuring computational reproducibility (Vilhuber, 2020). Another key implication is that empirical research should not be evaluated using uniform standards across methods. Instead, expectations regarding robustness

could be calibrated to the underlying research method and field. Editors and reviewers could thus rely on our evidence to determine which robustness checks matter in different empirical contexts, tailoring expectations by method and sub-field rather than applying uniform criteria.

Beyond its implications for the review process, our project also encourages researchers to adopt more transparent and careful practices, serving both as a model for improving standards and as a deterrent to questionable research practices. In doing so, it highlights the inherently collaborative nature of science, where reproduction and robustness checks are embraced as integral to scholarly progress. Over time, such efforts may foster a cultural and institutional shift in which reliability is not merely an afterthought but a central criterion for publication and recognition.

References

Aczel, Balazs, Barnabas Szaszi, Harry T Clelland, Marton Kovacs, Felix Holzmeister, Don Van Ravenzwaaij, Hannah Schulz-Kümpel, Sabine Hoffmann, Gustav Nilsonne, Livia Kosa et al., "Investigating the analytical robustness of the social and behavioural sciences," *Nature*, 2026, *652* (8108), 135–142.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2009.

Botvinik-Nezer, Rotem, Felix Holzmeister, Colin F Camerer, Anna Dreber, Juergen Huber, Magnus Johannesson, Michael Kirchler, Roni Iwanir, Jeanette A Mumford et al., "Variability in the Analysis of a Single Neuroimaging Dataset by Many Teams," *Nature*, 2020, *582* (7810), 84–88.

Breznau, Nate, Eike Mark Rinke, Alexander Wuttke, Hung HV Nguyen, Muna Adem, Jule Adriaans et al., "Observing Many Researchers Using the Same Data and Hypothesis Reveals a Hidden Universe of Uncertainty," *Proceedings of the National Academy of Sciences*, 2022, *119* (44), e2203150119.

Brodeur, Abel, Anna Dreber, Fernando Hoces de la Guardia, and Edward Miguel, "Replication Games: How to Make Reproducibility Research More Systematic," *Nature*, 2023, *621* (7980), 684–686.

_ , _ , _ , and _ , "Reproduction and replication at scale," *Nature Human Behaviour*, 2024, *8* (1), 2–3.

_ , Derek Mikola, Nikolai Cook, Lenka Fiala, Thomas Brailey, Ryan Briggs, Alexandra De Gendre, Yannick Dupraz, Jacopo Gabani, Romain Gauriot et al., "Reproducibility and Robustness of Economics and Political Science Research," *Nature*, 2026, *652* (8108), 151–156.

_ , Nikolai Cook, and Anthony Heyes, "Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics," *American Economic Review*, 2020, *110* (11), 3634–3660.

Camerer, Colin F, Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan et al., "Evaluating Replicability of Laboratory Experiments in Economics," *Science*, 2016, *351* (6280), 1433–1436.

_ , _ , Felix Holzmeister, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Gideon Nave, Brian A Nosek et al., "Evaluating the Replicability of Social Science Experiments in Nature and Science Between 2010 and 2015," *Nature Human Behaviour*, 2018, *2* (9), 637–644.

Campbell, Douglas, Abel Brodeur, Anna Dreber, Magnus Johannesson, Joseph Kopecky, Lester Lusher, and Nikita Tsoy, "The Robustness Reproducibility of the American Economic Review," 2024. I4R Discussion Paper Series 124.

Card, David and Stefano DellaVigna, "Nine facts about top journals in economics," *Journal of Economic literature*, 2013, *51* (1), 144–161.

Clemens, Michael A, "The Meaning of Failed Replications: A Review and Proposal," *Journal of Economic Surveys*, 2017, *31* (1), 326–342.

Davis, Andrew M, Blair Flicker, Kyle Hyndman, Elena Katok, Samantha Keppler, Stephen Leider, Xiaoyang Long, and Jordan D Tong, "A Replication Study of Operations Management Experiments in Management Science," *Management Science*, 2023, *69* (9), 4977–4991.

DellaVigna, Stefano and Devin Pope, "What Motivates Effort? Evidence and Expert Forecasts," *Review of Economic Studies*, 2018, *85* (2), 1029–1069.

_ and Eva Vivalt, "Forecasting social science: evidence from 100 projects," Technical Report, National Bureau of Economic Research 2025.

_ , Devin Pope, and Eva Vivalt, "Predict Science to Improve Science," *Science*, 2019, *366* (6464), 428–429.

Dreber, Anna, "Predicting Replication Outcomes and Generalizability of Results," *European Review of Agricultural Economics*, 2025, p. jbaf035.

_ **and Magnus Johannesson**, "A Framework for Evaluating Reproducibility and Replicability in Economics," *Economic Inquiry*, 2025, *63* (2), 338–356.

_ **,** _ **, and Yifan Yang**, "Selective reporting of placebo tests in top economics journals," *Economic Inquiry*, 2024, *62* (3), 921–932.

_ **, Thomas Pfeiffer, Johan Almenberg, Siri Isaksson, Brad Wilson, Yiling Chen, Brian A Nosek, and Magnus Johannesson**, "Using Prediction Markets to Estimate the Reproducibility of Scientific Research," *Proceedings of the National Academy of Sciences*, 2015, *112* (50), 15343–15347.

Fišar, Miloš, Ben Greiner, Christoph Huber, Elena Katok, Ali I Ozkes, and Management Science Reproducibility Collaboration, "Reproducibility in Management Science," *Management Science*, 2023.

Goldsmith-Pinkham, Paul, "Tracking the credibility revolution across fields," Technical Report, National Bureau of Economic Research 2026.

Hamermesh, Daniel S, "Replication in Economics," *Canadian Journal of Economics/Revue canadienne d'économique*, 2007, *40* (3), 715–733.

Huntington-Klein, Nick, Andreu Arenas, Emily Beam, Marco Bertoni, Jeffrey R Bloem, Pralhad Burli, Naibin Chen, Paul Grieco, Godwin Ekpe, Todd Pugatch et al., "The Influence of Hidden Researcher Decisions in Applied Microeconomics," *Economic Inquiry*, 2021, *59* (3), 944–960.

Menkveld, Albert J, Anna Dreber, Felix Holzmeister, Juergen Huber, Magnus Johannesson, Michael Kirchler, Sebastian Neusüss, Michael Razen, Utz Weitzel, David Abad-Díaz et al., "Nonstandard Errors," *Journal of Finance*, 2024, *79* (3), 2339–2390.

Miske, Olivia, Anna Lou Abatayo, Mason Daley, Mirka Dirzo, Nicholas Fox, Noah Haber, Krystal M Hahn, Melissa Kline Struhl, Brinna Mawhinney, Priya Silverstein et al., "Investigating the Reproducibility of the Social and Behavioural Sciences," *Nature*, 2026, *652* (8108), 126–134.

Nosek, Brian A, Tom E Hardwicke, Hannah Moshontz, Aurélien Allard, Katherine S Corker, Anna Dreber, Fiona Fidler, Joe Hilgard, Melissa Kline Struhl, Michèle B Nuijten et al., "Replicability, Robustness, and Reproducibility in Psychological Science," *Annual Review of Psychology*, 2022, *73*, 719–748.

Open Science Collaboration, "Estimating the Reproducibility of Psychological Science," *Science*, 2015, *349* (6251), aac4716.

Pérignon, Christophe, Olivier Akmansoy, Christophe Hurlin, Anna Dreber, Felix Holzmeister, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Albert J Menkveld, Michael Razen et al., "Computational Reproducibility in Finance: Evidence from 1,000 Tests," *The Review of Financial Studies*, 2024, *37* (11), 3558–3593.

Silberzahn, Raphael, Eric L Uhlmann, Daniel P Martin, Pasquale Anselmi, Frederik Aust, Eli Awtrey, Štěpán Bahník et al., "Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results," *Advances in Methods and Practices in Psychological Science*, 2018, *1* (3), 337–356.

Simonsohn, Uri, Joseph P Simmons, and Leif D Nelson, "Specification Curve Analysis," *Nature Human Behaviour*, 2020, *4* (11), 1208–1214.

Tyner, Andrew H, Anna Lou Abatayo, Mason Daley, Samuel Field, Nicholas Fox, Noah A Haber, Krystal M Hahn, Melissa Kline Struhl, Brinna Mawhinney, Olivia Miske et al., "Investigating the Replicability of the Social and Behavioural Sciences," *Nature*, 2026, *652* (8108), 143–150.

Vilhuber, Lars, "Reproducibility and Replicability in Economics," *Harvard Data Science Review*, 2020, *2*

(4).

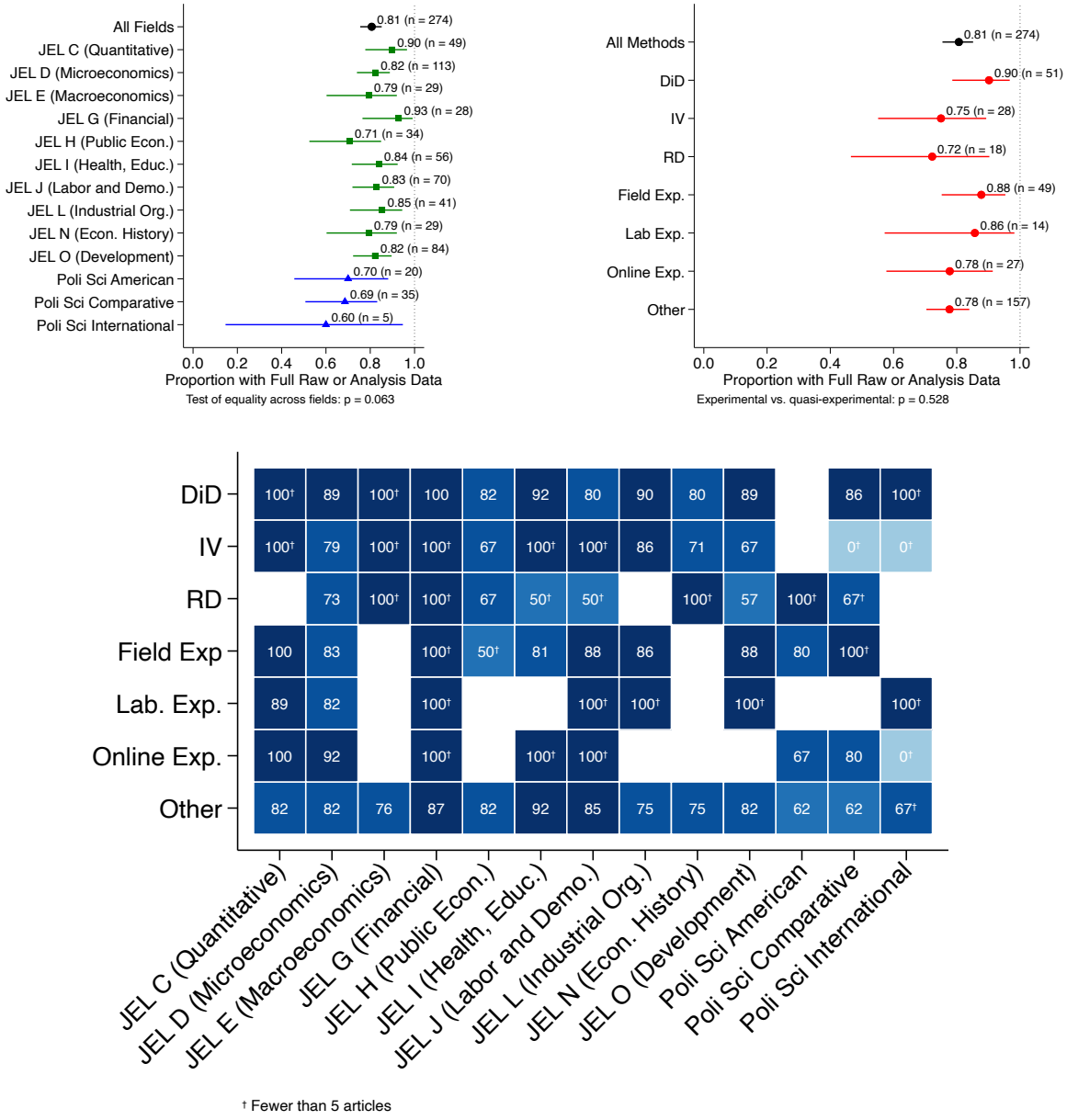
_ , **James Turrito, and Keesler Welch**, "Report by the AEA Data Editor," *AEA Papers and Proceedings*, May 2020, *110*, 764–75.

Vivalt, Eva, Aidan Coville et al., "Local Knowledge, Formal Evidence, and Policy Decisions," *Journal of Development Economics*, 2025, *174*, 103425.

_ **and** _ , "How do Policymakers Update their Beliefs?," *Journal of Development Economics*, 2023, *165*, 103121.

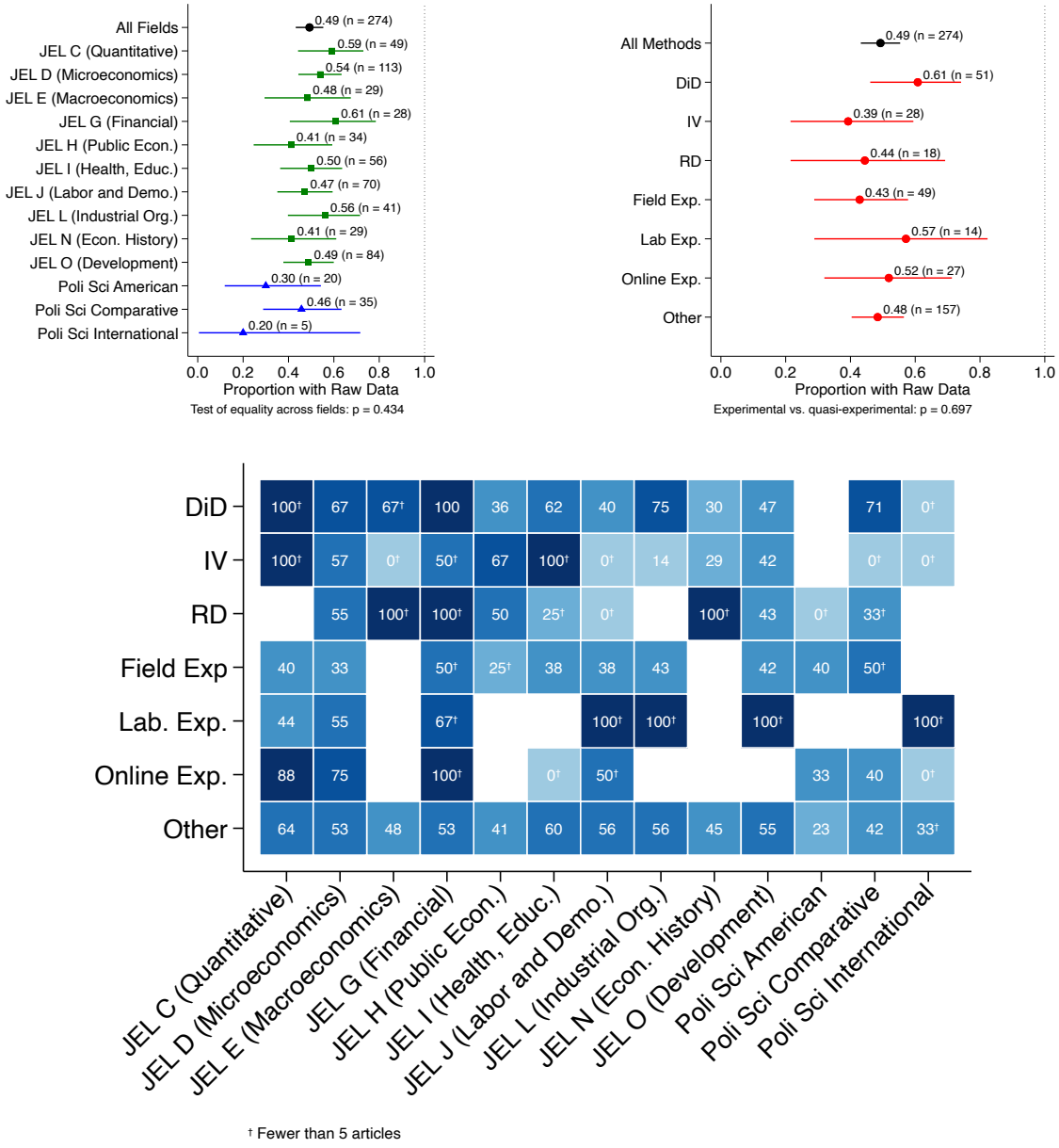
Figures

Figure 1: Data Availability



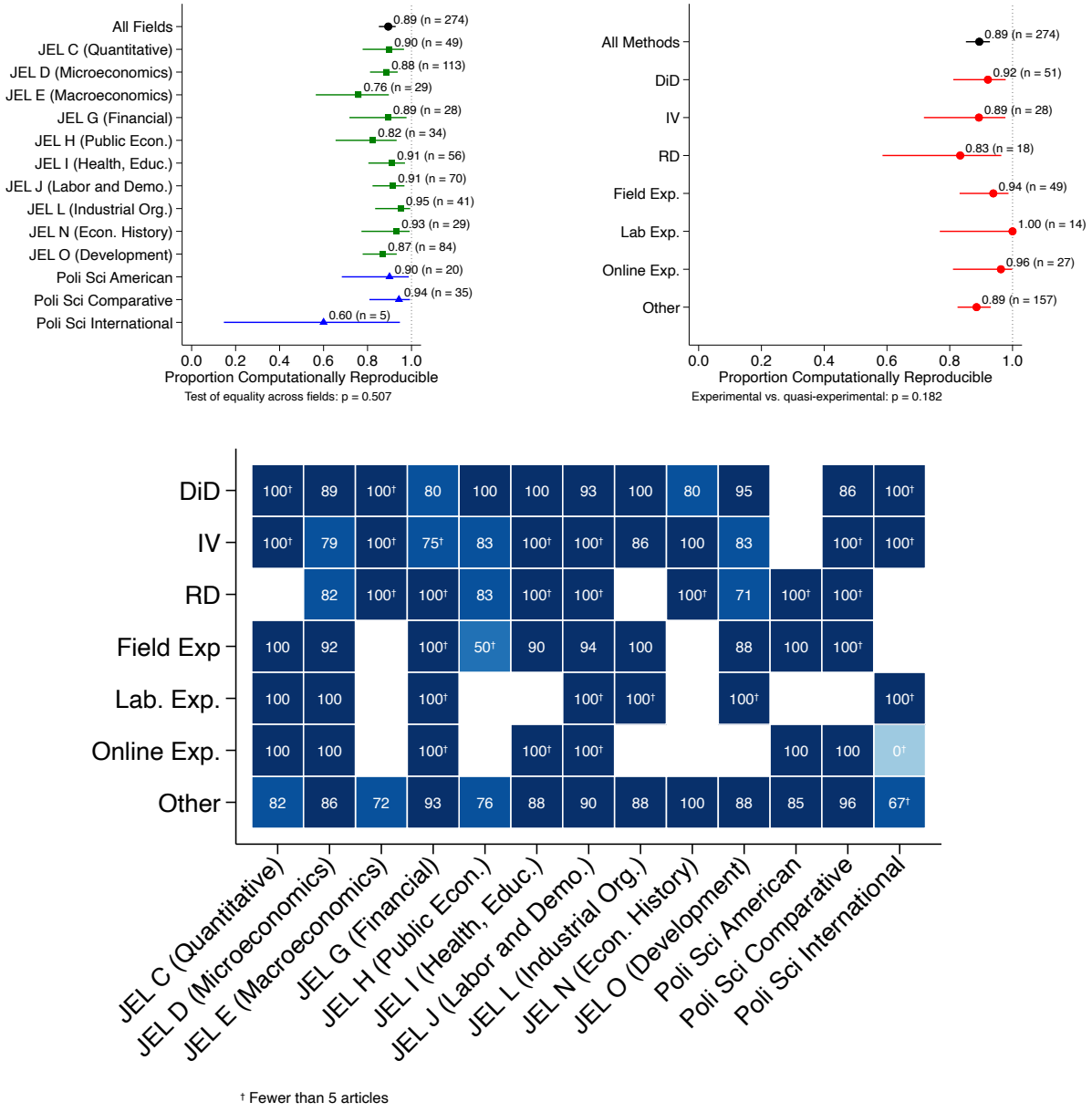
Notes: Replicators indicated whether the article's raw or analysis data was fully available. Restricted to PAP-specified sub-fields. An observation is an article. Articles are equally weighted. **Top left:** Proportion of articles that provide access to full raw or full analysis data by sub-field. Economics sub-fields (JEL codes, green squares) are not mutually exclusive, an article with JEL codes D and J appears in both D and J estimates. Political science sub-fields (blue triangles) are mutually exclusive. 95% confidence intervals provided. **Top right:** Proportion of articles that provide access to full raw or full analysis data by method. An article with an IV and DID appears in both IV and DID estimates. 'DID' includes difference-in-differences and two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Proportion of articles that provide access to full raw or full analysis data by sub-field and method. Darker cells indicate greater full raw or full analysis data availability.

Figure 2: Raw Data Availability



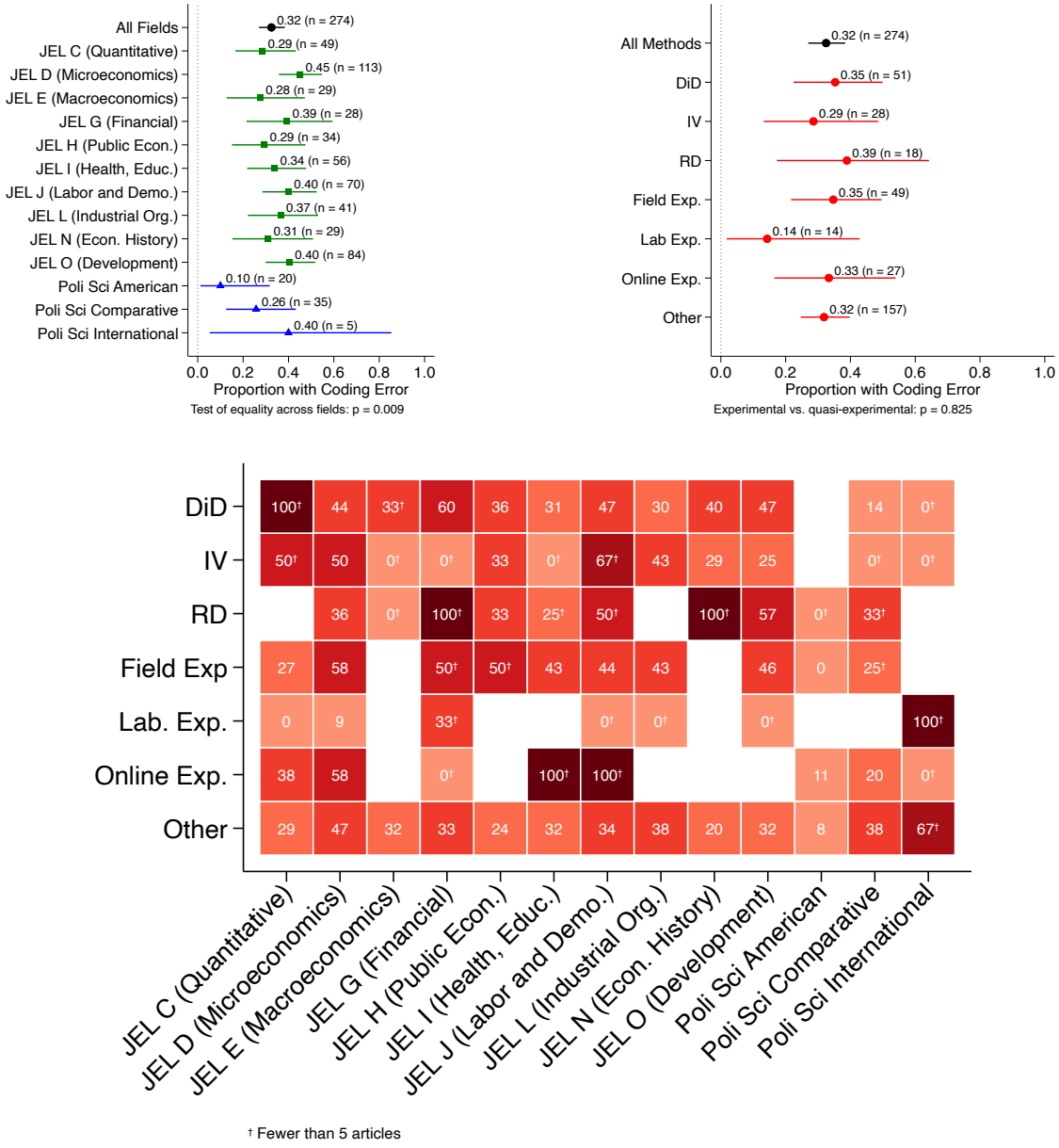
Notes: Replicators indicated whether the article's raw data was fully available. Restricted to PAP-specified sub-fields. An observation is an article. Articles are equally weighted. **Top left:** Proportion of articles that provide access to full raw data by sub-field. Economics sub-fields (JEL codes, green squares) are not mutually exclusive, an article with JEL codes D and J appears in both D and J estimates. Political science sub-fields (blue triangles) are mutually exclusive. 95% confidence intervals provided. **Top right:** Proportion of articles that provide access to full raw data by method. An article with an IV and DID appears in both IV and DID estimates. 'DID' includes difference-in-differences and two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Proportion of articles that provide access to full raw data by sub-field and method. Darker cells indicate greater full raw data availability.

Figure 3: Computational Reproducibility



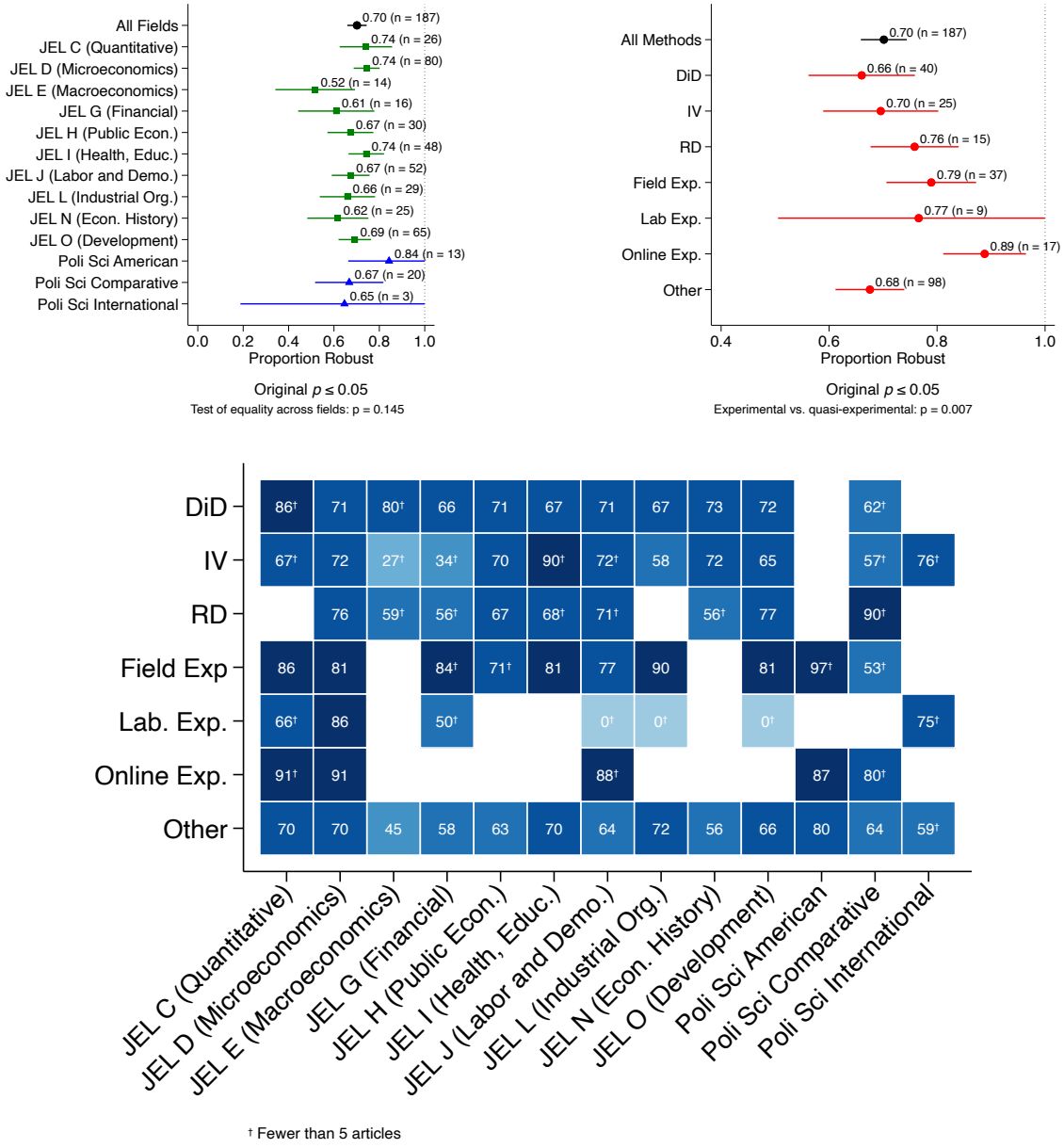
Notes: *Computational reproducibility* is whether an article's original results can be reproduced based on data and codes provided by the original authors. Replicators indicated whether the article was computationally reproducible from the article's replication package. Restricted to PAP-specified sub-fields. **Top left:** Proportion of articles that are computationally reproducible by sub-field. An observation is an article. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Proportion of article-methods that are computationally reproducible by method. An observation is an article-method (an article with an IV and DID appears in both IV and DID estimates). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Proportion of article-methods that are computationally reproducible by sub-field and method. Article-methods are equally weighted. Darker cells indicate greater computational reproducibility.

Figure 4: Coding Error Rate



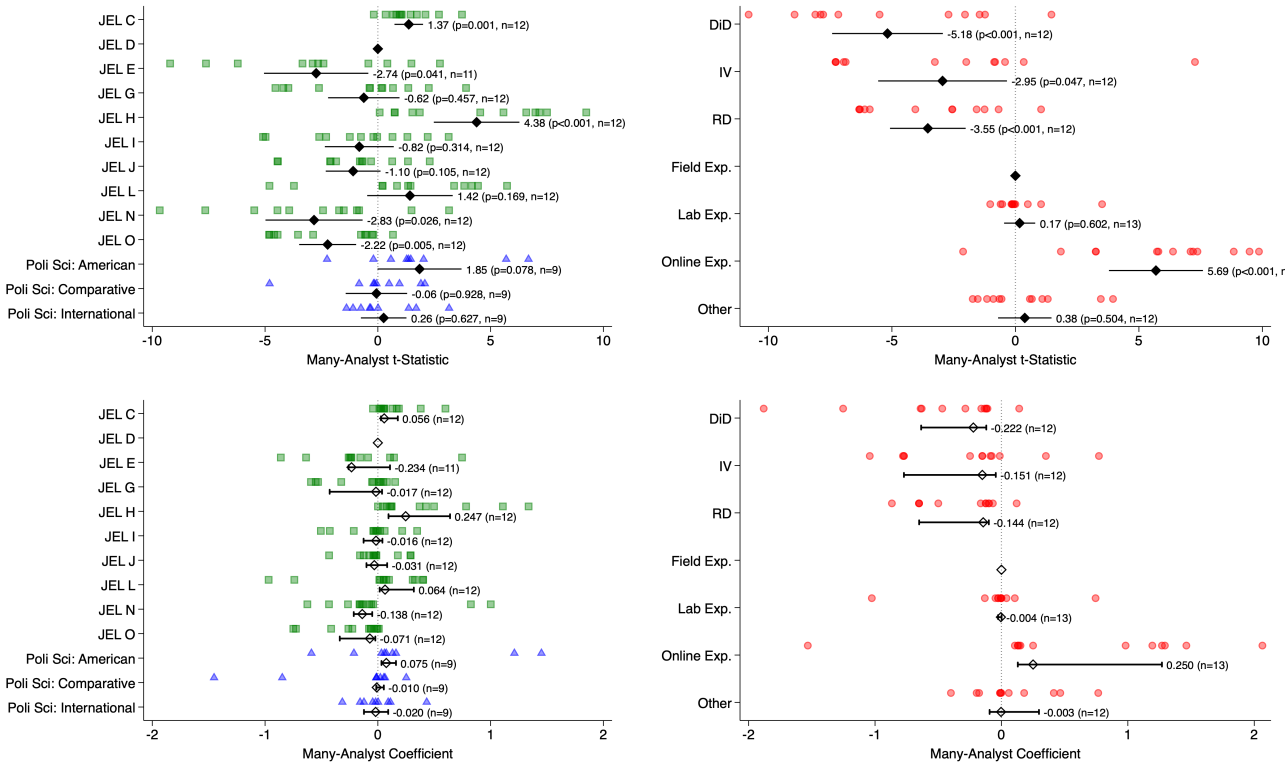
Notes: A *coding error* is one with the potential to affect empirical results or the substantive conclusions of an article. Replicators indicated in their reports whether the article contained a coding error. Restricted to PAP-specified sub-fields. **Top left:** Proportion of articles that have a coding error by sub-field. An observation is an article. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Proportion of article-methods that have a coding error by method. An observation is an article-method (an article with an IV and DID appears in both IV and DID estimates). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Proportion of article-methods that have a coding error by sub-field and method. Article-methods are equally weighted. Darker cells indicate greater coding error rates.

Figure 5: Robustness if $p \leq 0.05$



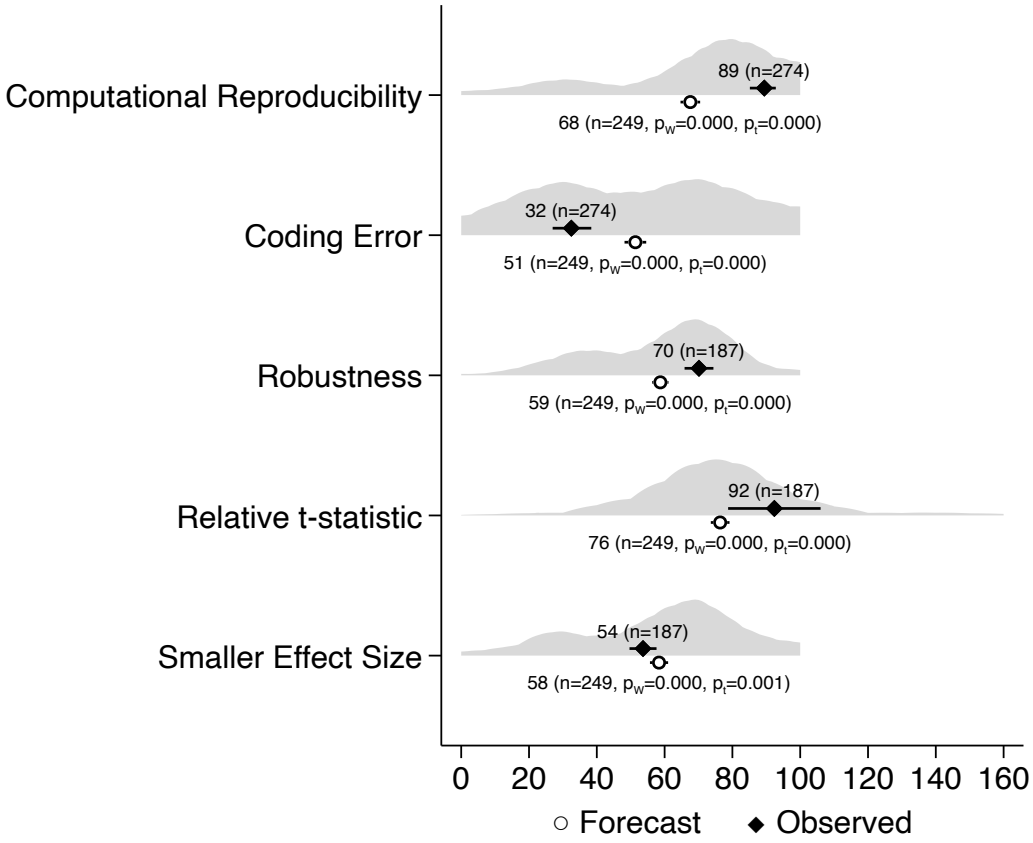
Notes: *Robustness* is whether an article's original results are robust to alternative reasonable analytical decisions on the same data. Replicators indicated whether the article's results are robust (replicators conducted multiple robustness checks per article, we aggregate that mean to the article level). Restricted to originally statistically significant results ($p \leq 0.05$) and PAP-specified sub-fields. **Top left:** Article-level average of robustness by sub-field. An observation is an article. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Article-method average of reproducibility-robustness by method. An observation is an article-method (an article with an IV and DID appears in both IV and DID estimates). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Article-method average of reproducibility-robustness by sub-field and method. Article-methods are equally weighted. Darker cells indicate greater reproducibility-robustness. **Related Figures:** For other measures of reproducibility see Figure A1 for the robustness check level, Figure A2 for statistically significant results ($p < 0.10$), Figure A3 for statistically insignificant results ($p > 0.05$), Figure A4 for statistically insignificant results ($p > 0.10$), Figure A5 for relative t-statistics, and Figure A6 for relative effect sizes.

Figure 6: Many Analysts Results



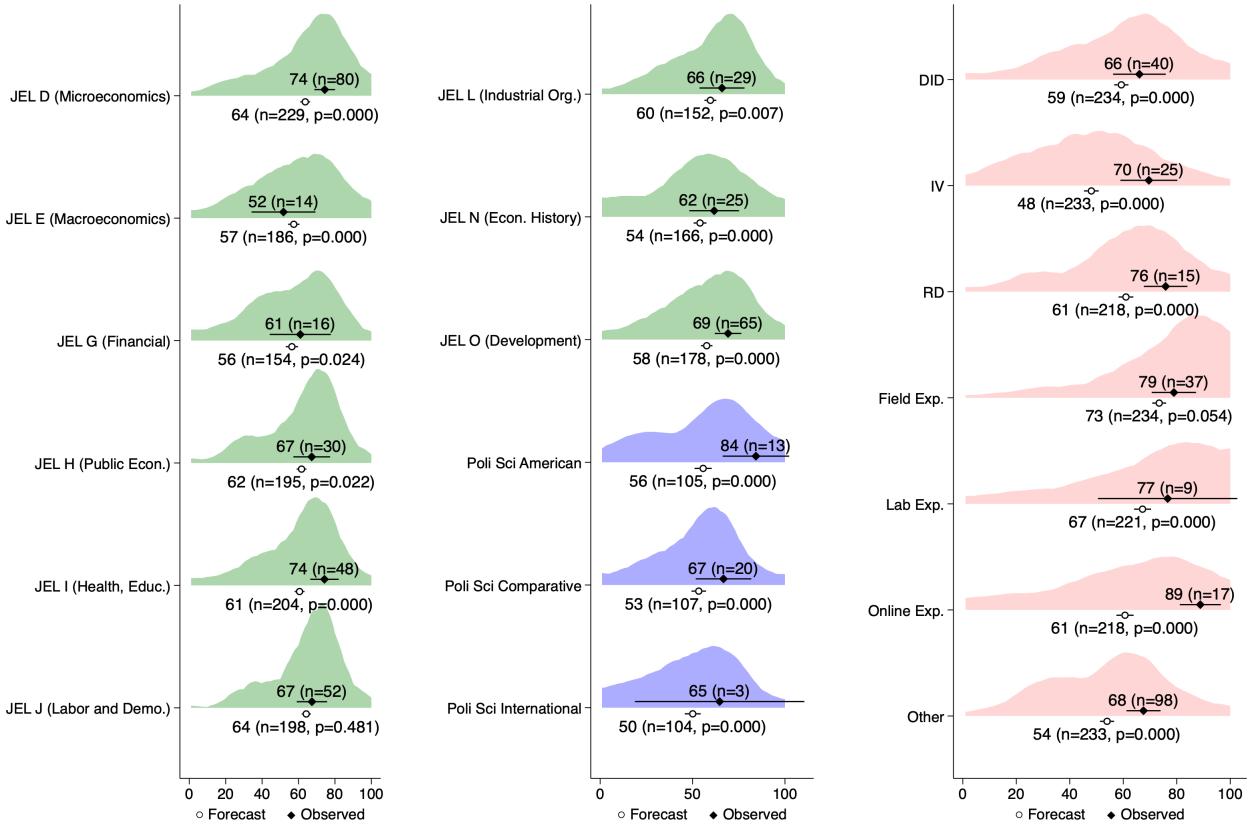
All Panels: A total of 13 independent many-analyst teams provided estimates of the robustness by field and method (not all teams provided estimates for every field and method). Teams were provided with the data used throughout this paper. Green markers represent coefficients for individual JEL codes. Blue triangles represent coefficients for mutually exclusive political science fields. Red circles represent coefficients for methods. All coefficients are comparisons to a baseline field (JEL D) and method (field experiment). **Top Panels:** Each marker represents the t-statistic associated with the field or method labeled by the vertical axis of a regression where the dependent variable is an indicator variable that takes a value one if an originally statistically significant result remained statistically significant under re-analysis. The horizontal axis indicates the sign and magnitude of the t-statistic. Black diamonds represent the average with 95% confidence intervals included. **Bottom Panels:** Each marker represents the estimated effect size associated with the field or method labeled by the vertical axis of a regression where the dependent variable is an indicator variable that takes a value one if an originally statistically significant result remained statistically significant under re-analysis. The horizontal axis indicates the sign and magnitude of the estimated effect size. Black hollow diamonds represent the median with the inter-quartile range of estimates included.

Figure 7: Forecast Results



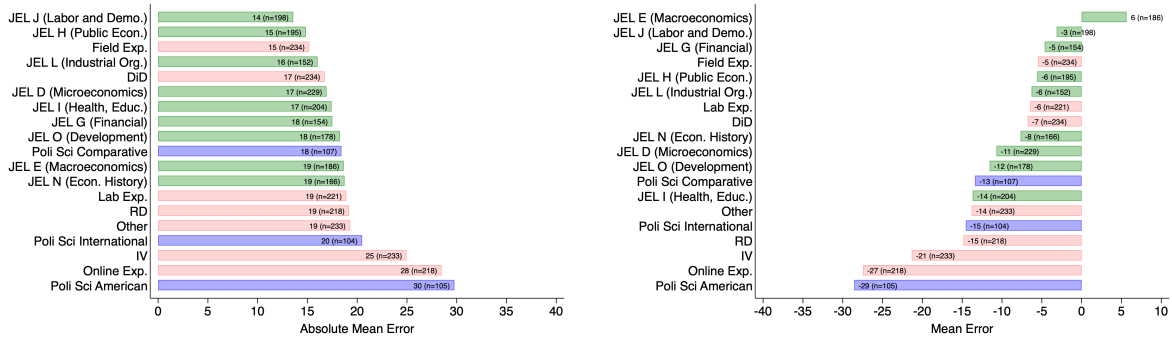
Notes: 249 experts completed forecasts on a pre-results prediction survey hosted by the *Social Science Prediction Platform* between September 30th to October 31st 2024. In that survey, experts were asked to forecast the effect of reproduction on computational reproducibility, coding errors, robustness, relative t-statistic, and proportion of effect sizes that would be smaller after re-analysis. Kernel densities display the distribution of forecasts, from 0-100, of the respondents (relative t-statistic was allowed to vary from 0-200). Black hollow circles represent the forecast mean, with 95% confidence intervals represented in bars. Black diamonds represent the reproduction mean, with 95% confidence intervals represented in bars. P-values (p_w) are from two-sided, one-sample Wilcoxon signed-rank tests of whether individual forecast errors (each respondent's forecast minus the observed value) are centered at zero. The test ranks the absolute forecast errors (within question) and accounts for their signs, providing a nonparametric assessment of systematic over- or underprediction. P-values (p_t) are from a two-sided one-sample t-test of whether the mean forecast equals the observed value.

Figure 8: Forecast Results By Field and Method



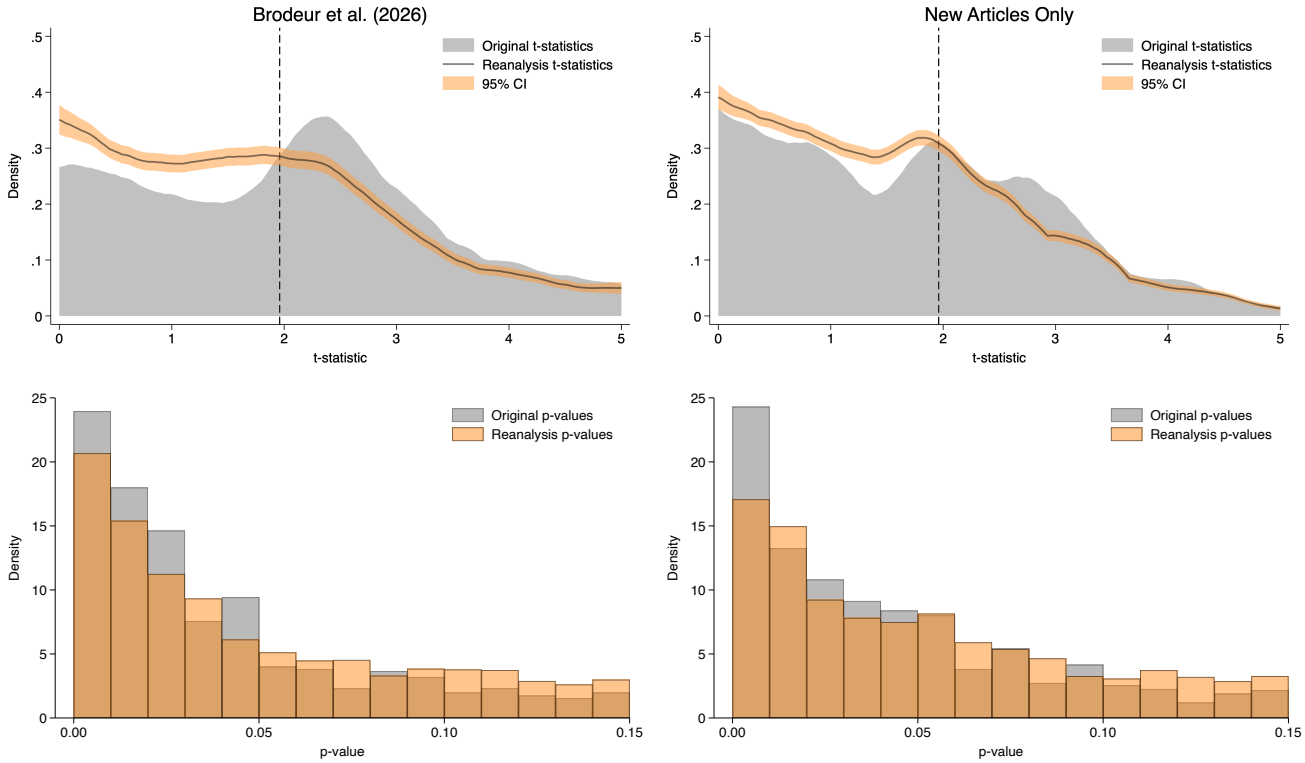
All Panels: 249 experts completed forecasts on a pre-results prediction survey hosted by the *Social Science Prediction Platform* between September 30th to October 31st 2024. In that survey, experts were asked to predict robustness rates by sub-field (in Economics, at the JEL code level, in Political Science, for the three fields of American Politics, Comparative Politics, and International Relations) and method. Kernel densities display the distribution of scores, from 0-100, of how robust the respondents believed each field and method would be. Green densities represent economics. Blue densities represent political science. Red densities represent methods. Black circles represent the forecast mean, with 95% confidence intervals represented in bars. Black diamonds represent the replication mean, with 95% confidence intervals represented in bars. P-values are from one-sample Wilcoxon signed-rank tests of individual field- or method-specific forecast errors (forecast minus observed value) against zero.

Figure 9: Prediction Errors



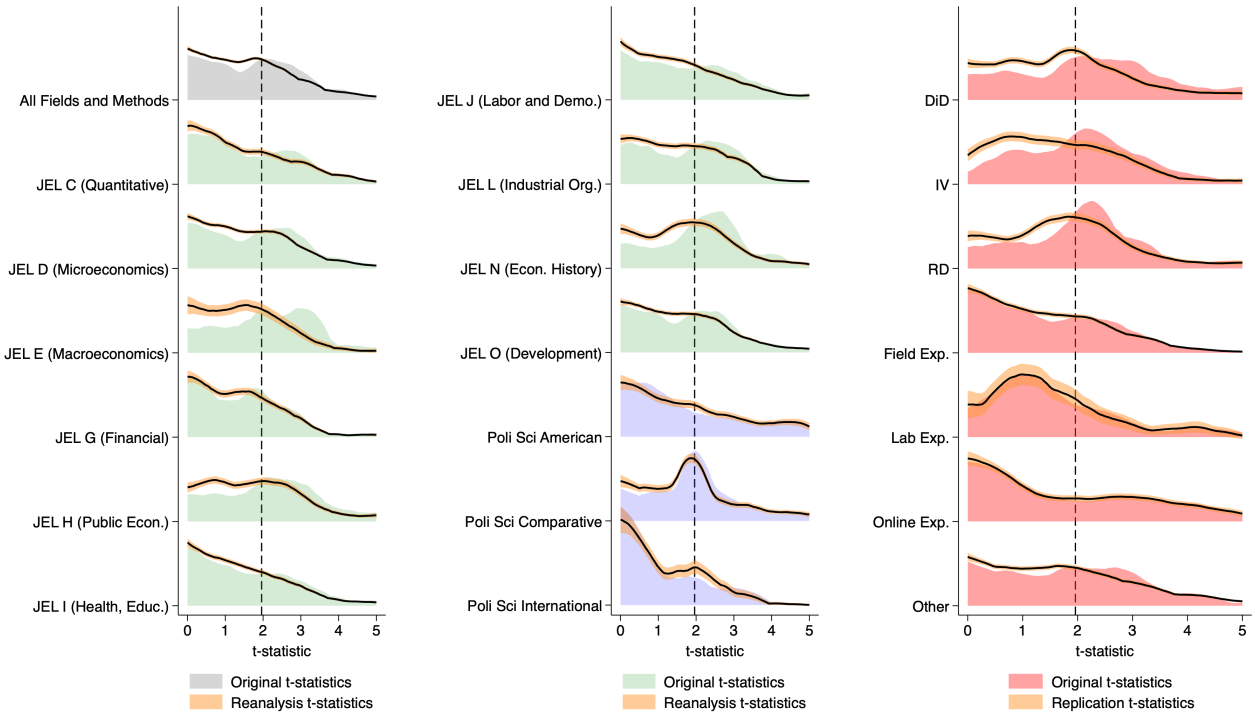
All Panels: Green bars represent fields in economics. Blue bars represent fields in political science. Red bars represent methods (not exclusive to the fields). **Left panel:** Sorted by the methods and fields the expert forecasters were best able to predict. Average of the absolute value of (forecast - observed) by method and field, sorted from best-well predicted to least-well predicted. **Right panel:** Average values of the mean signed forecast error of (forecast - observed) by method and field, sorted from least under-estimated to most under-estimated.

Figure 10: Comparison to Brodeur et al. (2026)



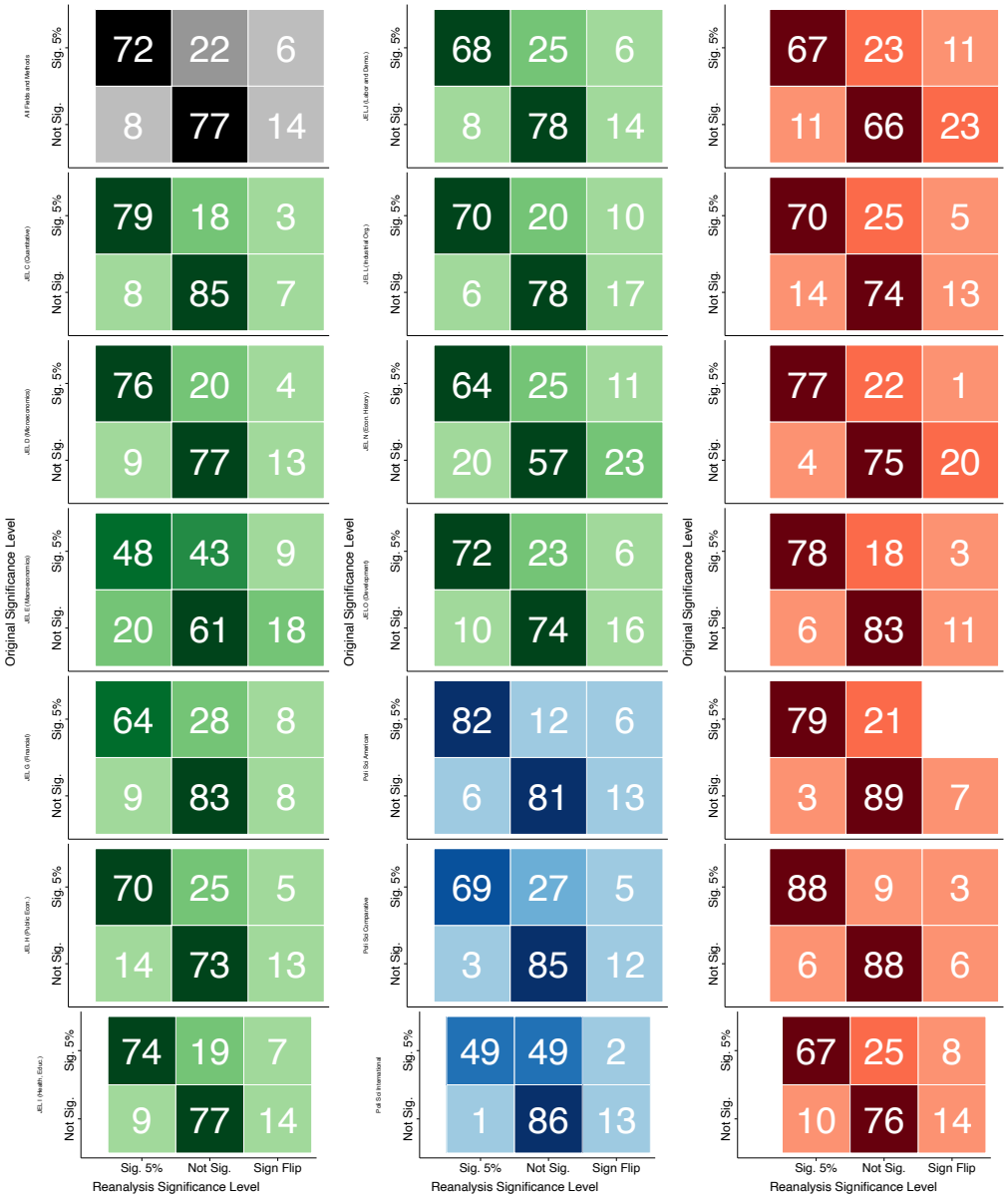
Top left: Gray density represents the t-statistics ($t \in [0, 5]$) from 110 articles used in Brodeur et al. (2026) for all fields and methods. A black Epanechnikov kernel line, with 95% confidence interval, represents the reproduction and replication t-statistics for the same 110 articles. Vertical reference lines are displayed at the conventional two-tailed significance level of 5%. **Top right:** t-statistics from new articles only. **Bottom left:** p-values from 110 articles used in Brodeur et al. (2026) for all fields and methods, ($p \in [0.0025, 0.1500]$ for illustrative purposes). **Bottom right:** p-values from new articles only.

Figure 11: Patterns of Statistical Significance by Field and Method



All Panels: All panels display filled densities of originally published test statistics for $t \in [0, 5]$. Imposed on top are Epanechnikov kernel lines, with 95% confidence intervals, of the associated reproduction test statistics. Vertical reference lines are displayed at conventional two-tailed significance levels. Green curves represent fields in economics. Blue curves represent fields in political science. Red curves represent methods (not exclusive to the fields).

Figure 12: Shifts of Statistical Significance



All Panels: Shifts across statistical significance ($p \leq 0.05$) and insignificance regions ($p > 0.05$). Each row is the original statistical significance. Each column is the re-analysis statistical significance. Cells represent the proportion of originally statistically (in)significant test statistics that stayed the same (diagonals) or changed to either statistical significance, insignificance, or flipped sign. All article-methods equally weighted. Green boxes represent fields in economics. Blue boxes represent fields in political science. Red boxes represent methods (not exclusive to the fields).

Tables

Table 1: Data Availability and Computational Reproducibility

	Fully	Partially	None	All
Any data	221	5	48	274
	(81%)	(2%)	(18%)	(100%)
Raw data	135	39	100	274
	(49%)	(14%)	(36%)	(100%)
Analysis data	214	7	53	274
	(78%)	(3%)	(19%)	(100%)
Cleaning code	149	47	78	274
	(54%)	(17%)	(28%)	(100%)
Analysis code	261	12	1	274
	(95%)	(4%)	(0%)	(100%)
Cleaning if raw data	135	29	10	174
	(78%)	(17%)	(6%)	(100%)
Analysis code if any data	217	8	1	226
	(96%)	(4%)	(0%)	(100%)
Reproducible from any data	245	16	13	274
	(89%)	(6%)	(5%)	(100%)
Reproducible from raw data	123	31	120	274
	(45%)	(11%)	(44%)	(100%)
Reproducible from analysis data	244	17	13	274
	(89%)	(6%)	(5%)	(100%)

Notes: Sample restricted to 274 articles with originally pre-registered JEL codes and political science fields. "Any data" refers to the full provision of either raw data or analysis data. "Raw data" and "analysis data" refer to the full provision of the indicated type. "Cleaning code" refers to the scripts and codes that prepare data, and do not (but can include) analysis code. "Analysis code" refers to scripts and codes that calculate and present the results of the article. "Reproducible from..." refers to the fact that replicators' considered the article reproducible from the indicated data level.

Table 2: Number of Robustness Checks by Data and Code Availability

	(1) Any Full Data	(2) Full Raw	(3) Full Analysis	(4) Any Raw	(5) Any Analysis	(6) Any Raw and Cleaning	(7) Any Analysis and A. Code
Available	5.934	11.888	10.448	0.871	16.157	-2.009	17.218
	(14.786)	(12.917)	(13.732)	(13.069)	(13.284)	(13.165)	(13.158)
Articles	274	274	274	274	274	274	274
Avg Checks	73.84	73.84	73.84	73.84	73.84	73.84	73.84

Notes: The unit of observation is an article. Each column is a separate regression. The dependent variable in each is the number of robustness checks conducted for that article by replicators. The primary independent variable is indicated by the column title. For example, the primary independent variable in the first column is a binary indicator that takes the value one if there is full provision of either raw or analysis data. Standard errors in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

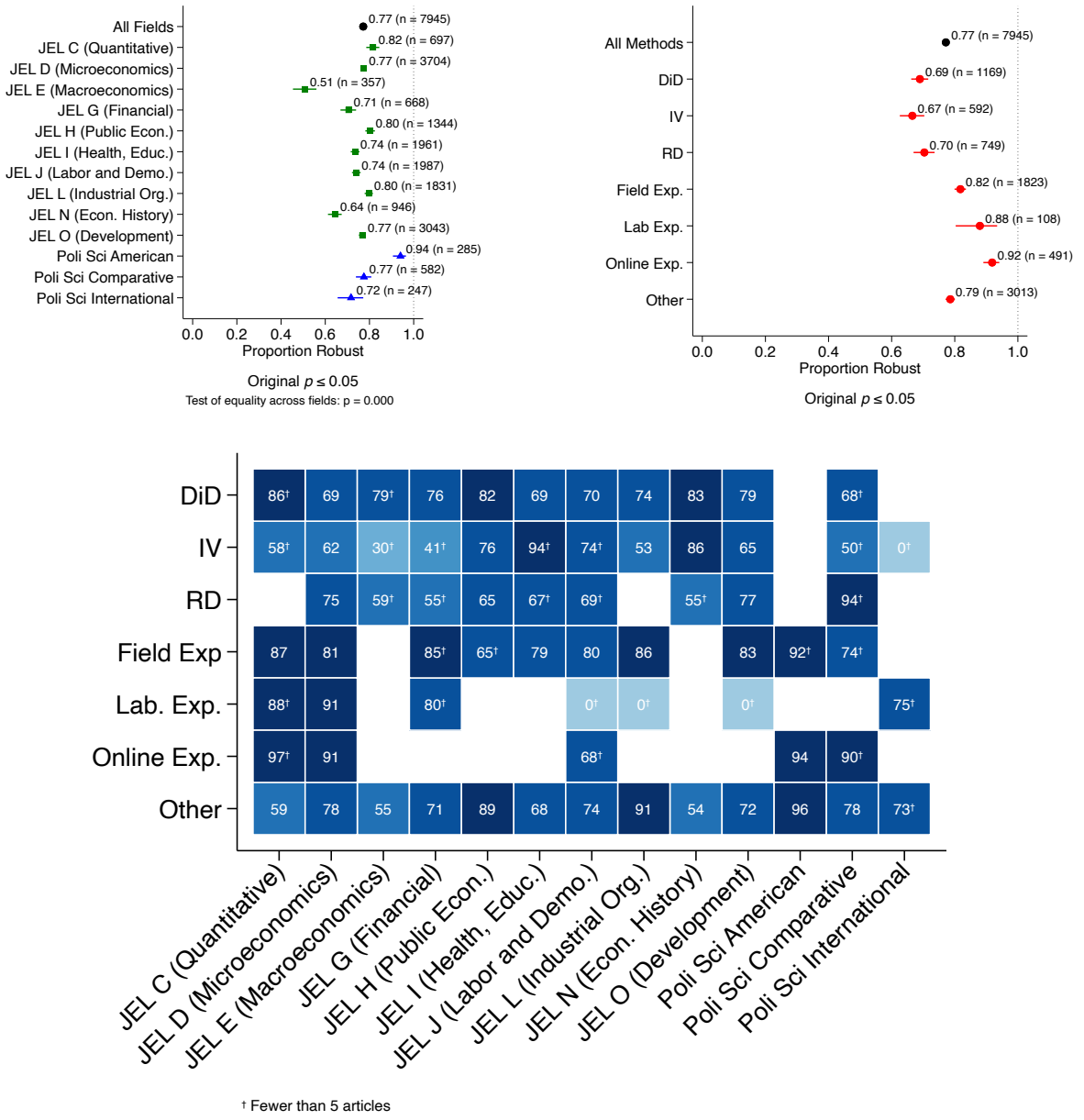
Table 3: Coding Error Types

	Total	Econ	PolSci	Diff	p-value	Exp	Quasi-Exp	Diff	p-value
Dependent var.	0.047	0.056	0.017	0.039	0.204	0.069	0.043	0.025	0.458
Main ind. var.	0.044	0.047	0.033	0.013	0.654	0.080	0.043	0.037	0.303
Control var.	0.099	0.117	0.033	0.083	0.055	0.115	0.130	-0.015	0.752
Estimator	0.069	0.084	0.017	0.067	0.069	0.023	0.087	-0.064	0.063
Inference	0.040	0.042	0.033	0.009	0.761	0.023	0.033	-0.010	0.696
Sample	0.106	0.103	0.117	-0.014	0.758	0.092	0.130	-0.038	0.414
Other	0.106	0.107	0.100	0.007	0.868	0.069	0.109	-0.040	0.352
N	274.000	214.000	60.000	.	.	87.000	92.000	.	.

Notes: The sample is 274 articles with pre-registered JEL codes and political science fields. Each row represents a different type of coding error; types are not mutually exclusive. "Econ" refers to articles published in an economics journal. The difference between fields in percentage points is provided in column 4, and a p-value of a test of equality of proportions is presented in column 5. Succeeding columns repeat the exercise between experimental methods (field, online, and lab experiments) against quasi-experimental methods (DiD, IV, and RD). 'Other' methods are excluded in those comparisons.

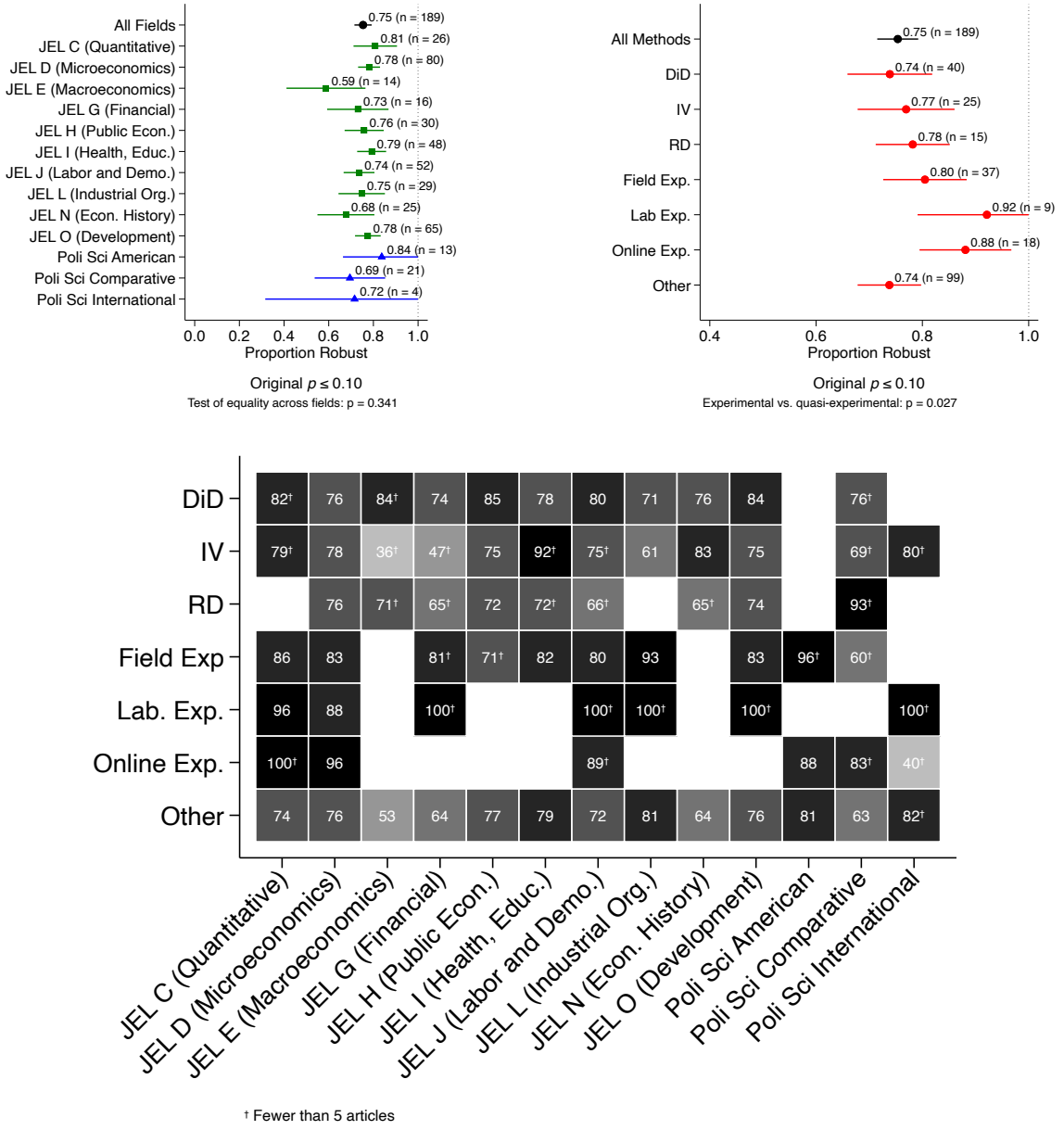
Appendix Figures

Figure A1: Reproducibility if $p \leq 0.05$ at the robustness check level



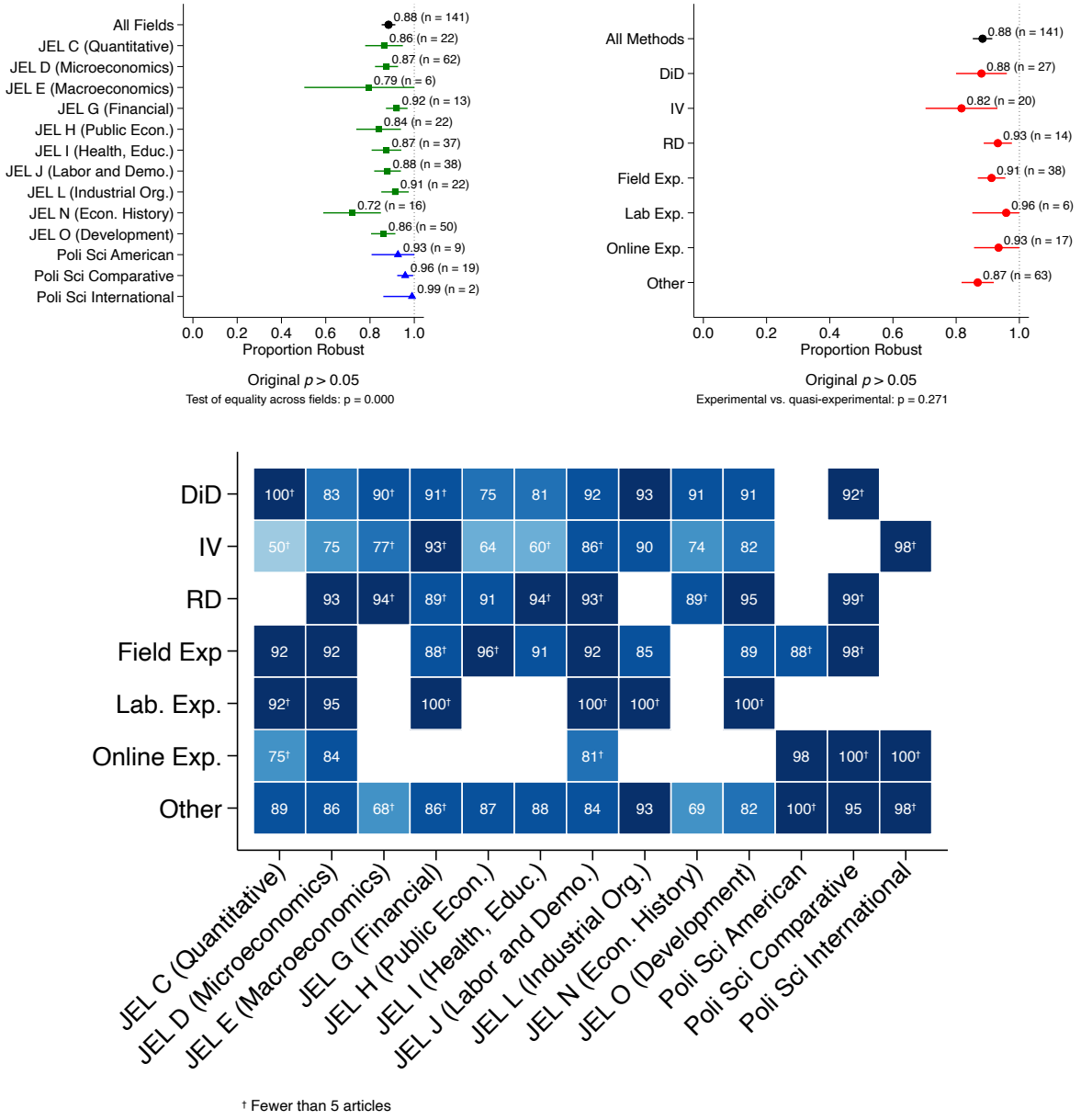
Notes: *Robustness* is whether original results are robust to alternative reasonable analytical decisions on the same data. Replicators indicated whether each result was robust (replicators conducted multiple robustness checks per article, this figure does not aggregate them to the article level). Restricted to originally statistically significant results ($p \leq 0.05$) and PAP-specified sub-fields. **Top left:** Test-level average of reproducibility-robustness by sub-field. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Test-level average of reproducibility-robustness by method. 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Test-level average of reproducibility-robustness by sub-field and method. Tests are equally weighted. Darker cells indicate greater reproducibility-robustness.

Figure A2: Robustness if $p \leq 0.10$



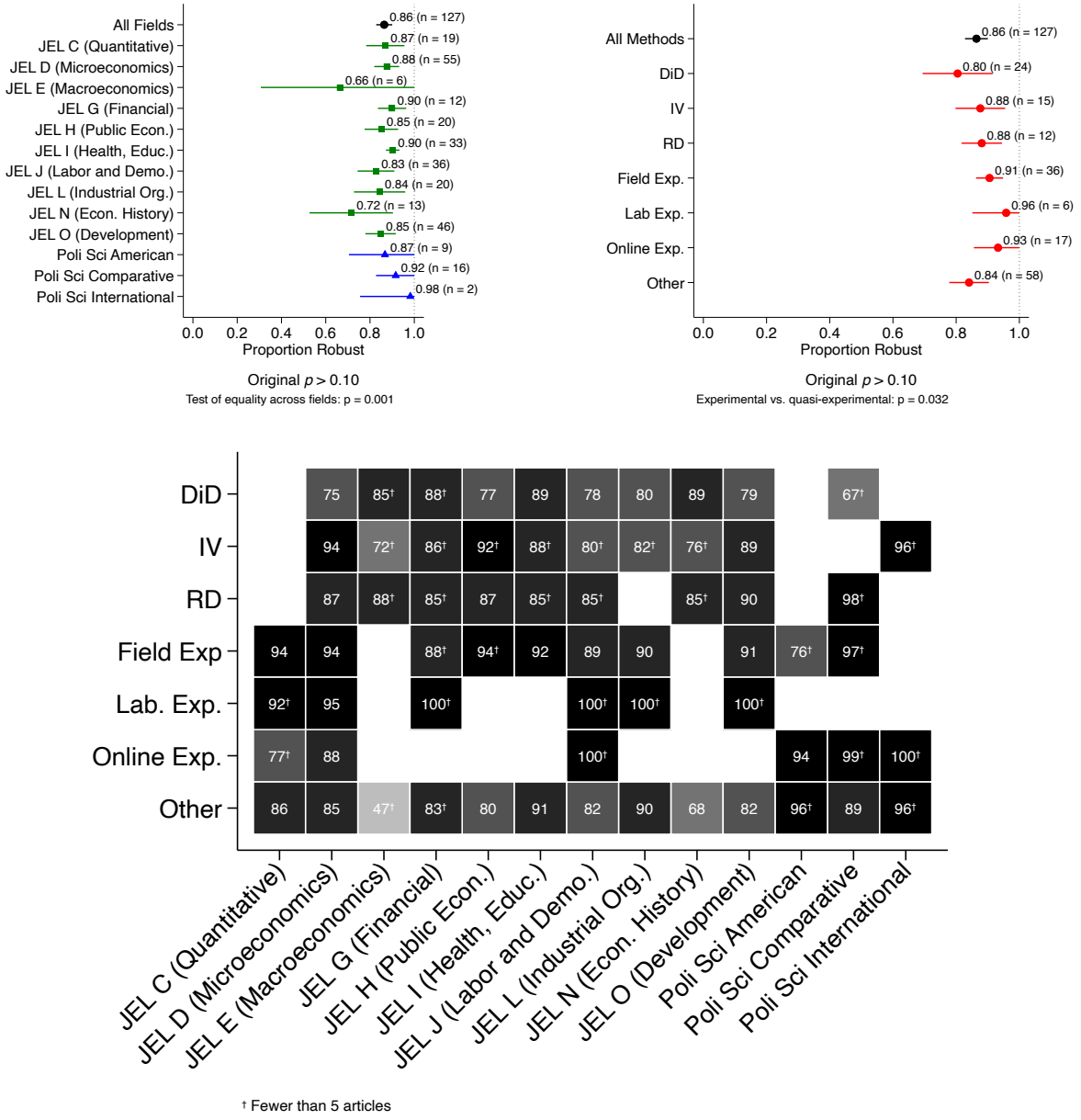
Notes: *Robustness* is whether an article's original results are robust to alternative reasonable analytical decisions on the same data. Replicators indicated whether the article's results are robust (replicators conducted multiple robustness checks per article, we aggregate that mean to the article level). Restricted to originally statistically significant results ($p \leq 0.10$) and PAP-specified sub-fields. **Top left:** Article-level average of robustness by sub-field. An observation is an article. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Article-method average of reproducibility-robustness by method. An observation is an article-method (an article with an IV and DID appears in both IV and DID estimates). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Article-method average of reproducibility-robustness by sub-field and method. Article-methods are equally weighted. Darker cells indicate greater reproducibility-robustness.

Figure A3: Reproducibility if $p > 0.05$



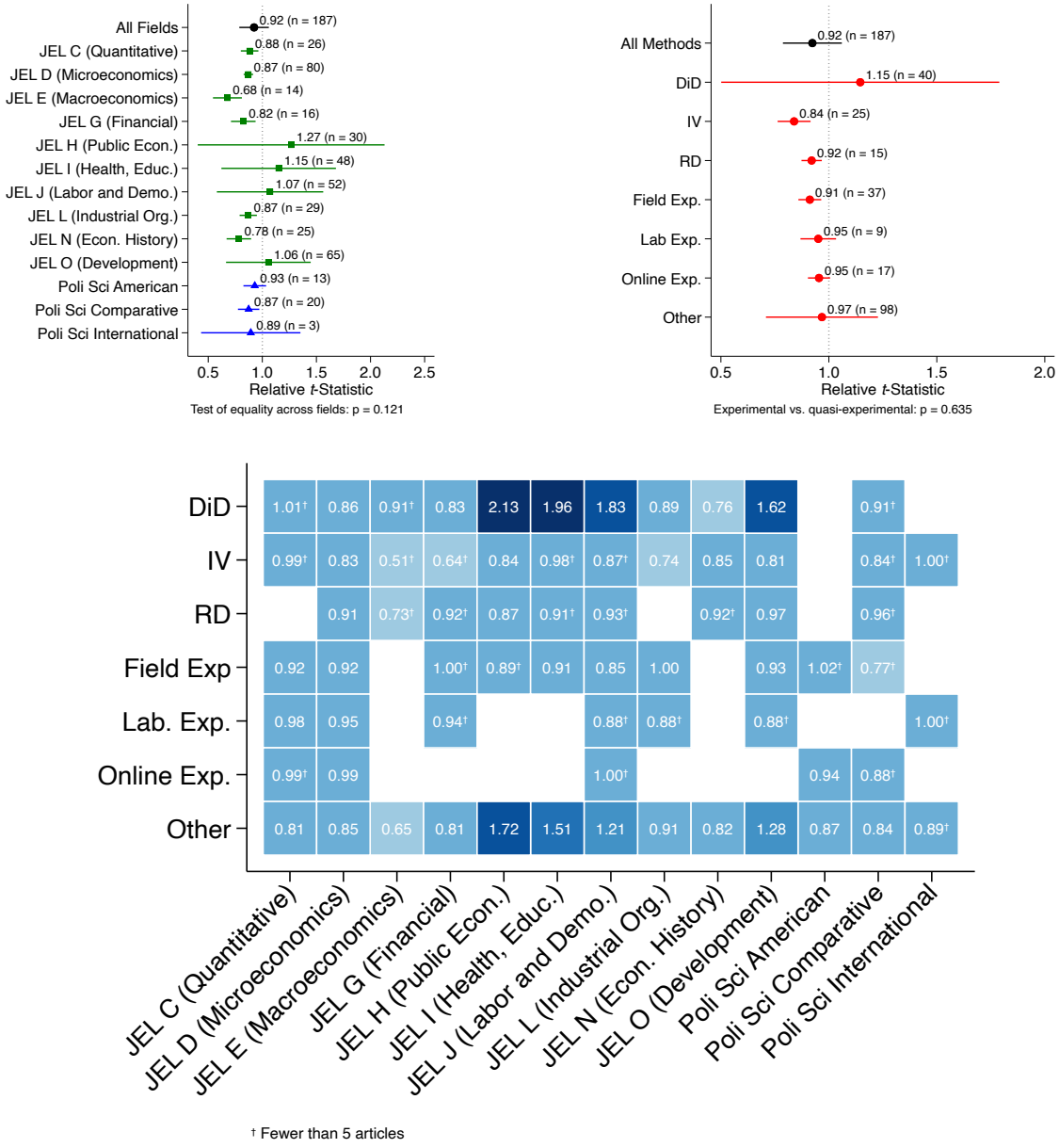
Notes: *Robustness* is whether an article's original results are robust to alternative reasonable analytical decisions on the same data. Replicators indicated whether the article's results are robust (replicators conducted multiple robustness checks per article, we aggregate that mean to the article level). Restricted to originally statistically insignificant results ($p > 0.05$) and PAP-specified sub-fields. **Top left:** Article-level average of reproducibility-robustness by sub-field. An observation is an article. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Article-method average of reproducibility-robustness by method. An observation is an article-method (an article with an IV and DID appears in both IV and DID estimates). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Article-method average of reproducibility-robustness by sub-field and method. Article-methods are equally weighted. Darker cells indicate greater reproducibility-robustness.

Figure A4: Reproducibility if $p > 0.10$



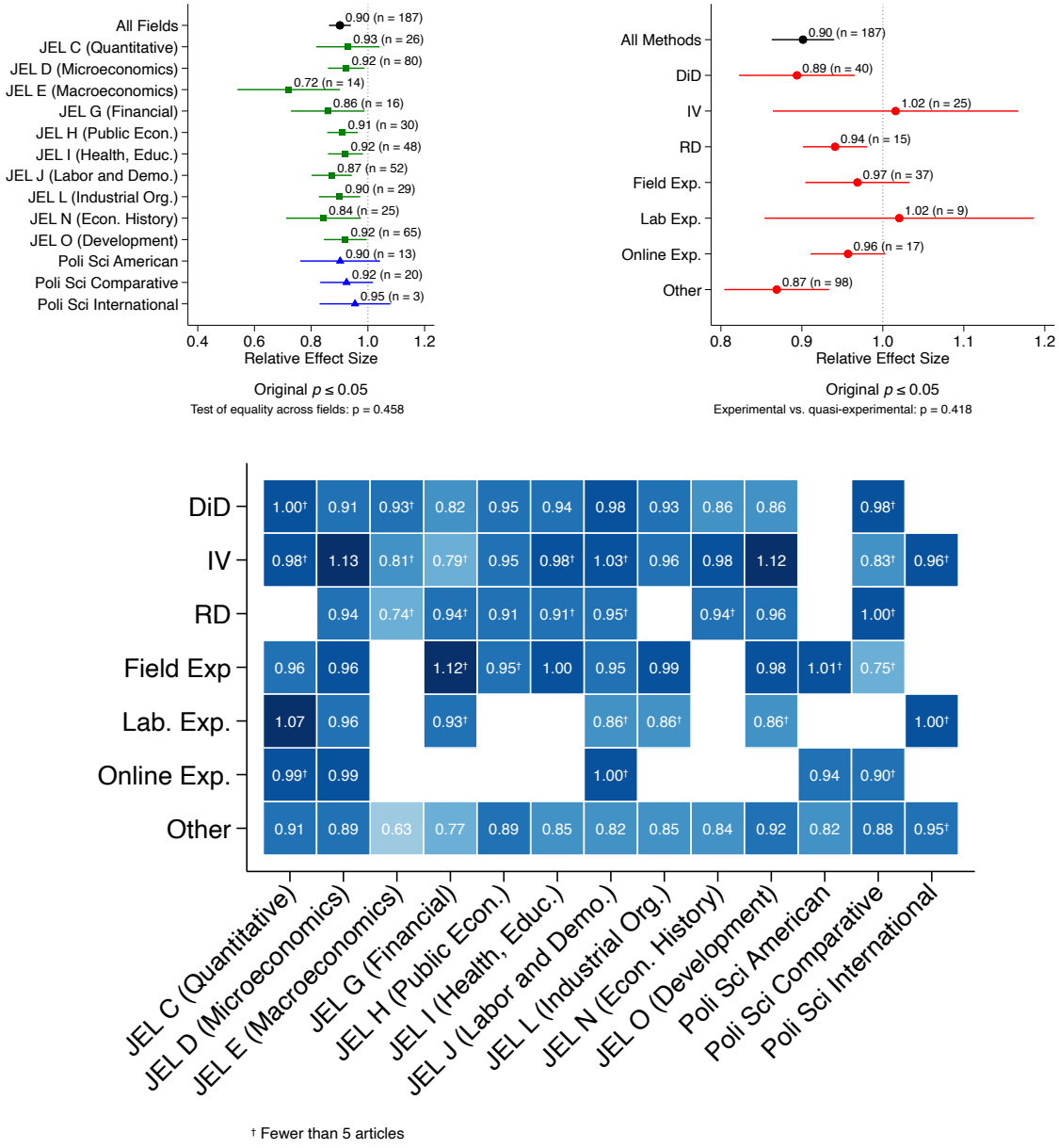
Notes: *Robustness* is whether an article's original results are robust to alternative reasonable analytical decisions on the same data. Replicators indicated whether the article's results are robust (replicators conducted multiple robustness checks per article, we aggregate that mean to the article level). Restricted to originally statistically insignificant results ($p > 0.10$) and PAP-specified sub-fields. **Top left:** Article-level average of reproducibility-robustness by sub-field. An observation is an article. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Article-method average of reproducibility-robustness by method. An observation is an article-method (an article with an IV and DID appears in both IV and DID estimates). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Article-method average of reproducibility-robustness by sub-field and method. Article-methods are equally weighted. Darker cells indicate greater reproducibility-robustness.

Figure A5: Relative t-Statistic



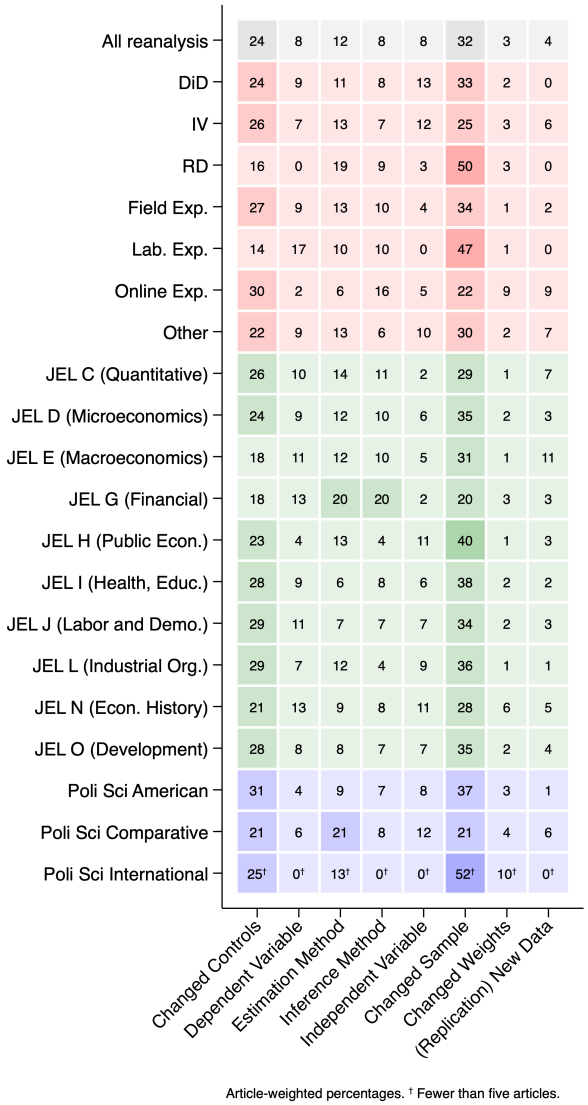
Notes: A *relative t-statistic* is the ratio of a robustness (re-analysis) t-statistic to the associated article t-statistic. Replicators provided multiple robustness t-statistics, we aggregate to the article level. Restricted to PAP-specified sub-fields. Restricted to originally statistically significant results ($p \leq 0.05$) and PAP-specified sub-fields. **Top left:** Article-level average of relative t-statistics by sub-field. An observation is an article. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Article-method average of relative t-statistics by method. An observation is an article-method (an article with an IV and DID appears in both IV and DID estimates). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Article-method average of relative t-statistics by sub-field and method. Article-methods are equally weighted. Darker cells indicate greater relative t-statistics.

Figure A6: Relative Effect Size if $p \leq 0.05$



Notes: A *relative effect size* is the ratio of a robustness (re-analysis) effect size (e.g. regression coefficient) to the associated original effect size (all in absolute value). Replicators provided multiple robustness effects sizes, we aggregate to the article level. Restricted to PAP-specified sub-fields and originally statistically significant results. **Top left:** Article-level average of relative effect size by sub-field. An observation is an article. Economics sub-fields (JEL codes, green squares) are assigned at article-level (an article with JEL codes D and J appears in both D and J estimates). Political science sub-fields (blue triangles) are assigned at article-level and are mutually exclusive. 95% confidence intervals provided. **Top right:** Article-method average of relative effect size by method. An observation is an article-method (an article with an IV and DID appears in both IV and DID estimates). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal). 95% confidence intervals provided. **Heatmap:** Article-method average of relative effect size by sub-field and method. Article-methods are equally weighted. Darker cells indicate greater relative effect sizes.

Figure A7: Robustness Check Type by Method and Field



Notes: Each cell presents the percentage of robustness checks, articles-equally-weighted, within row (row-cell sum to 100, restricted to only these 8 most common types). Each column represents a robustness check type. Robustness checks which span multiple categories given one over the number of categories in each row-cell. Restricted to PAP-specified sub-fields. Economics sub-fields (JEL codes, green squares) are assigned at article-level. Political science sub-fields (blue squares) are assigned at article-level and are mutually exclusive. Methods are mutually exclusive (red squares). 'DID' includes difference-in-differences or two-way fixed effects. 'Other' includes simulations, structural, theory, time series, synthetic control, and more (including non-causal).

Appendix Tables

Table A1: Summary Statistics

	mean	min	max	count
Number of (numeric) Robustness Checks	73.84	1	616	274
Data Availability	0.81	0	1	274
Raw Data Availability	0.49	0	1	274
Computationally Reproducible	0.89	0	1	274
Coding Error	0.32	0	1	274
Reproducible if Orig. Stat. Sig. 5%	0.70	0	1	187
Method				
Difference-in-Differences	0.19	0	1	274
Instrumental Variables	0.10	0	1	274
Regression Discontinuity	0.07	0	1	274
Field Experiment	0.18	0	1	274
Laboratory Experiment	0.05	0	1	274
Online Experiment	0.10	0	1	274
Other	0.57	0	1	274
Field				
Quantitative Methods (JEL C)	0.18	0	1	274
Microeconomics (JEL D)	0.41	0	1	274
Macroeconomics (JEL E)	0.11	0	1	274
Financial Economics (JEL G)	0.10	0	1	274
Public Economics (JEL H)	0.12	0	1	274
Health, Education, and Welfare (JEL I)	0.20	0	1	274
Labor and Demographic Economics (JEL J)	0.26	0	1	274
Industrial Organization (JEL L)	0.15	0	1	274
Economic History (JEL N)	0.11	0	1	274
Economic Development and Growth (JEL O)	0.31	0	1	274
American Politics	0.07	0	1	274
Comparative Politics	0.13	0	1	274
International Relations	0.02	0	1	274
Journal				
American Economic Journal: Applied Economics	0.10	0	1	274
American Economic Journal: Economic Policy	0.10	0	1	274
American Economic Journal: Macroeconomics	0.05	0	1	274
American Economic Review	0.19	0	1	274
American Economic Review: Insights	0.03	0	1	274
American Journal of Political Science	0.07	0	1	274
American Political Science Review	0.08	0	1	274
Journal of Political Economy	0.06	0	1	274
The Economic Journal	0.16	0	1	274
The Journal of Politics	0.07	0	1	274
The Quarterly Journal of Economics	0.05	0	1	274
The Review of Economic Studies	0.03	0	1	274

Table A2: Caliper Test for Publication Bias and p-Hacking

	(1)	(2)	(3)	(4)
Difference-in-Differences	0.176**	0.129	0.160	0.189
	(0.087)	(0.088)	(0.118)	(0.160)
Instrumental Variables	0.198**	0.197**	0.251**	0.274**
	(0.085)	(0.095)	(0.110)	(0.127)
Regression Discontinuity	0.100	0.032	-0.001	-0.029
	(0.093)	(0.091)	(0.136)	(0.136)
Laboratory Experiment	0.119	0.265**	0.060	0.044
	(0.171)	(0.114)	(0.166)	(0.185)
Online Experiment	-0.095	-0.054	0.146	0.301**
	(0.129)	(0.163)	(0.146)	(0.145)
Other	0.090	0.087	0.009	0.051
	(0.082)	(0.091)	(0.119)	(0.131)
Quantitative Methods (JEL C)	-0.027	-0.003	0.085	0.141
	(0.067)	(0.064)	(0.091)	(0.098)
Microeconomics (JEL D)	-0.024	-0.048	-0.072	0.078
	(0.063)	(0.068)	(0.088)	(0.086)
Macroeconomics (JEL E)	-0.010	-0.062	0.252	0.719***
	(0.126)	(0.177)	(0.226)	(0.273)
Financial Economics (JEL G)	-0.164**	-0.194**	-0.089	-0.098
	(0.073)	(0.076)	(0.087)	(0.094)
Public Economics (JEL H)	-0.038	-0.036	-0.092	-0.074
	(0.071)	(0.083)	(0.104)	(0.113)
Health, Education, and Welfare (JEL I)	-0.008	-0.055	-0.012	0.210**
	(0.068)	(0.075)	(0.095)	(0.102)
Labor and Demographic Economics (JEL J)	0.026	0.003	-0.034	-0.028
	(0.065)	(0.070)	(0.087)	(0.092)
Industrial Organization (JEL L)	-0.030	-0.072	-0.150	-0.079
	(0.073)	(0.080)	(0.099)	(0.127)
Economic History (JEL N)	0.053	0.070	0.076	0.170*
	(0.092)	(0.099)	(0.113)	(0.101)
Economic Development and Growth (JEL O)	0.015	0.036	0.150*	0.185**
	(0.060)	(0.064)	(0.078)	(0.082)
American Politics	-0.106	-0.284***	-0.301	0.119
	(0.138)	(0.110)	(0.197)	(0.241)
Comparative Politics	-0.054	-0.130	-0.099	0.168
	(0.123)	(0.130)	(0.161)	(0.175)
International Relations	-0.051	-0.278		
	(0.255)	(0.247)		
Window	±.50	±.35	±.20	±.10
Test Statistics	1410	1009	597	330
Articles	162	147	127	101

Notes: An observation is a unique published test statistic (duplicates that exactly match all three of p-value, coefficient, and standard error dropped). The dependent variable is an indicator that takes the value one if the published test statistic is significant at the 5% level, and zero otherwise (a positive coefficient is indicative of publication bias or p-hacking, jointly). Methods (DiD, IV, etc.) are mutually exclusive. The reference group is field experiment. Fields (JEL codes) are not exclusive. Weighted by the inverse of the number of unique tests in an article. Standard errors clustered at the article level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A3: Reproducibility and Robustness by Method and Field

	Full Data	Comp. Rep.	Coding Error	Sig. (5%)	Sig. (10%)	Insig. (5%)	Insig. (10%)
Difference-in-Differences	0.180***	0.041	0.002	0.043	0.054***	0.039	-0.006
	(0.036)	(0.022)	(0.033)	(0.029)	(0.013)	(0.085)	(0.076)
Instrumental Variables	-0.030	0.036	-0.103	0.063*	0.081*	-0.007	0.086
	(0.045)	(0.044)	(0.062)	(0.031)	(0.037)	(0.087)	(0.072)
Regression Discontinuity	0.011	-0.001	0.002	0.111***	0.077**	0.119***	0.084
	(0.092)	(0.022)	(0.089)	(0.027)	(0.022)	(0.031)	(0.045)
Field Experiment	0.108***	0.092**	-0.004	0.108*	0.061	0.075**	0.080*
	(0.029)	(0.031)	(0.035)	(0.047)	(0.052)	(0.023)	(0.033)
Laboratory Experiment	-0.001	0.176**	-0.273**	0.133**	0.245***	0.118**	0.152**
	(0.072)	(0.049)	(0.084)	(0.048)	(0.048)	(0.048)	(0.055)
Online Experiment	0.033	0.114**	0.068**	0.198***	0.184***	0.082*	0.150**
	(0.055)	(0.039)	(0.020)	(0.039)	(0.035)	(0.039)	(0.055)
Quantitative Methods (JEL C)	0.039	-0.042	-0.125	-0.016	0.013	-0.044	-0.026
	(0.097)	(0.097)	(0.068)	(0.083)	(0.091)	(0.061)	(0.044)
Microeconomics (JEL D)	-0.072**	-0.037	0.244**	0.052	0.008	-0.031	-0.018
	(0.020)	(0.035)	(0.078)	(0.067)	(0.054)	(0.030)	(0.041)
Macroeconomics (JEL E)	-0.077	-0.121	-0.045	-0.134	-0.124	-0.046	-0.183
	(0.074)	(0.088)	(0.061)	(0.129)	(0.134)	(0.039)	(0.113)
Financial Economics (JEL G)	0.087	0.019	0.081	-0.068	-0.014	0.069*	0.086
	(0.055)	(0.064)	(0.086)	(0.070)	(0.050)	(0.032)	(0.051)
Public Economics (JEL H)	-0.170***	-0.063	-0.092*	-0.041	0.009	-0.041	-0.004
	(0.035)	(0.087)	(0.041)	(0.048)	(0.061)	(0.057)	(0.042)
Health, Educ., and Welfare (JEL I)	-0.023	0.018	-0.081**	0.070**	0.060**	-0.005	0.079
	(0.039)	(0.062)	(0.032)	(0.023)	(0.017)	(0.047)	(0.052)
Labor and Demo. Econ. (JEL J)	-0.043	0.030	0.106*	-0.025	-0.008	0.035	-0.011
	(0.046)	(0.026)	(0.048)	(0.063)	(0.067)	(0.019)	(0.041)
Industrial Organization (JEL L)	-0.069*	0.064	0.002	-0.031	-0.013	0.050	0.001
	(0.035)	(0.054)	(0.099)	(0.072)	(0.056)	(0.027)	(0.047)
Econ. History (JEL N)	-0.077	0.070	-0.041	-0.033	-0.064	-0.156*	-0.093
	(0.059)	(0.083)	(0.129)	(0.169)	(0.146)	(0.075)	(0.119)
Econ. Dev. and Growth (JEL O)	-0.050	-0.076	0.070	-0.019	0.042	-0.020	-0.030
	(0.047)	(0.050)	(0.047)	(0.054)	(0.062)	(0.022)	(0.051)
American Politics	-0.247	-0.067	-0.185***	0.051	0.032	-0.025	-0.098
	(0.144)	(0.136)	(0.021)	(0.138)	(0.155)	(0.071)	(0.125)
Comparative Politics	-0.272**	0.014	-0.005	-0.028	-0.027	0.057	0.050
	(0.109)	(0.071)	(0.071)	(0.123)	(0.166)	(0.060)	(0.049)
International Relations	-0.341*	-0.366	0.206	-0.055	-0.086	0.098*	0.051
	(0.140)	(0.195)	(0.266)	(0.137)	(0.154)	(0.041)	(0.032)
Articles	274	274	274	187	189	141	127

Notes: An observation is an article. The dependent variable is (1) an indicator that the replication package has either full raw or analysis data (2) an indicator that the article is computationally reproducible (3) an indicator that the replication package had a coding error (4) the article-level statistically significant reproducibility rate at 5% (5) 10% (6) insignificant at 5% (7) 10% under reanalysis. At this article-level, methods (DiD, IV, etc.) are not mutually exclusive. The reference group is other identification methods. Fields (JEL codes) are not mutually exclusive. Standard errors two-way clustered at the method and article level in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A ONLINE APPENDIX A

A.1 Author list

Nikolai Cook (Wilfrid Laurier University), Derek Mikola (University of Ottawa and Institute for Replication), Juan P. Aparicio (University of Ottawa and Institute for Replication), Lenka Fiala (University of Ottawa and Institute for Replication), Ghina Abdul Baki (University of Ottawa and Institute for Replication), Alexander Cassiel Abajian (Yale School of the Environment), Kseniya Abanokova (World Bank), Mamdouh Abdelkader (Department of Economics, University of Ottawa), Sahar Abi-Hassan (Department of Political Science, University of Tulsa), Mohamed Alhadi A Abubraidaa (Simon Fraser University), Joop Adema (University of Innsbruck), George Agyeah (Walmart Inc), Farida Akhtar (Macquarie University), Shumi Akhtar (The University of Sydney), Dinara Alpysbayeva (Faculty of Economics and Administration, Masaryk University), Micah Altman (Center for Research in Equitable and Open Scholarship, Massachusetts Institute of Technology), Flavia Alves (Carleton University), Satyam Anand (Georgetown University), Mirinda Lee (Martin) Anderson (Brigham Young University-Idaho), Zubaria Andlib (Lancaster University, the UK, Federal Urdu University of Arts, Science and Technology, Islamabad, Pakistan), David Angenendt (SKEMA Business School, Université Côte d'Azur (GREDEG),), Konstantin Ash (StrategEast), Thomas Ash (Anderson School of Management, University of California, Los Angeles), Tobias Auer (London School of Economics), Antonio Avila-Uribe (Institution and university), Idrissa Aladji AYA (Université Clermont Auvergne, CNRS, IRD, CERDI, F-63000 CLERMONT FERRAND, France.), Abdelkarim Azami idrissi (University of lorraine, Bureau d'Économie Théorique et Appliquée (BETA)), Marc Bachmeier (Freie Universität Berlin), Lior Bakalo (The Faculty of Data and Decision Sciences, Technion – Israel Institute of Technology, Haifa, Israel), E. Grant Baldwin (University of California, Los Angeles), Praveena Bandara (Net Zero Industrial Policy Lab), Amanda Baptista Denham (University- UTS), Söhnke M. Bartram (University of Warwick and CEPR), Gabriel Bayle (GATE, CNRS, Université Lumière Lyon 2, Université Jean-Monnet Saint-Etienne, emlyon business school), Dana Bazarkulova (Nazarbayev University), Clayton Becker (ULCA), Louis-Philippe Beland (Carleton University, Department of Economics), Kendon Bell (Scarlatti), David Benatia (HEC Montréal), Guillaume Bérard (Luxembourg Institute of Socio-Economic Research (LISER)), Marta Bernardi (Research Associate and PhD Candidate at the Dresden University of Technology), Tim Bersak (Wofford College), Simone Bertoli (Université Clermont Auvergne, CNRS, IRD, CERDI), Konrad Bierl (Resource Economics Group, Humboldt University Berlin; Berlin School of Economics), Paul Binder (LMU Munich), Boris Blagov (RWI - Leibniz Institute for Economic Research), Lisa Bogler (Department of Economics and Centre for Modern Indian Studies, Georg-August-Universität Göttingen), Carl Bonander (Karlstad University), Arianna Bondi (McGill University), Stephanie Bonds (RAND Corporation), Adil Boufarsi (Université Grenoble Alpes, INRAE, UMR GAEL, 38000 Grenoble, France.), Adam Bouyamourn (Center for Data Science, NYU), Thomas Brailey (University of Oxford), Michelle Bronsard (Stanford University), Roberto Brunetti (LEMMA, Université Paris Panthéon-Assas), Lukas Brunninger (Vienna University of Economics and Business), Jens Christian Brunstad (University of Bergen), Calvin Bryan (Washington and Lee University), Nino Buliskeria (Nazarbayev University), Emil Bustos (Independent (previously: Research Institute of Industrial Economics)), Nienke Buters (Technical University Munich (TUM)), Anand Butler (University of Colorado, Boulder), Alistair Cameron (Laterite Ltd), Gewei Cao (Max Planck In-

stitute for the Study of Crime, Security and Law), Michelle Escobar Carias (Newcastle University Business School, University of Melbourne (Department of Economics)), Florian Caro (Yale University), Alice Carraro (University of Padova), Fortuna Casoria (Université Bourgogne Europe, Burgundy School of Business, CEREN EA 7477, 21000 Dijon, France.), Antoine Castet (University of Rouen Normandy, LERN), Brian Castro-Aguirre (University of Houston), Simonas Čepėnas (ISM University of Management and Economics), Luca Cermak (Institute of Economics, Ulm University), Marek Chadim (Yale University), Cédric Chambru (École normale supérieure de Lyon, CERGIC, EMR CNRS EqCo 5001), Patrick Chester (Stevens Institute of Technology), Erik Chi (Texas A&M University), Hongbi Choi (Texas A&M University (Department of Political Science)), Taesoo Choi (Department of Agricultural and Applied Economics, University of Wisconsin-Madison), Shih-Hsien Chuang (Northwest Missouri State University), Costin Ciobanu (Aarhus University), Elías Cisneros (The University of Texas at Dallas), Melchior Clerc (Université Clermont Auvergne, CNRS, IRD and CERDI), Guila Cohen (McGill University), Adrien Coiffard (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France), Jason Collins (University of Technology Sydney), Matthew Collins (University of Galway), Kenneth Colombe (University of Bonn, Department of Economics), Mariarosaria Comunale (International Monetary Fund), Marie Connolly (Université du Québec à Montréal), Mattia Coppo (Multilateral Investment Guarantee Agency (MIGA)), Guillaume Coqueret (EMLYON Business School), Cesar Cornejo (London School of Hygiene & Tropical Medicine (LSHTM), University College London (UCL)), Claudio Costanzo (ECARES), Jacques Courbe (UCLA Department of Political Science), Mathieu Couttenier (Ecole Normale Superieure de Lyon), Paolo Crosetto (INRAE), Andrés Cruz (University of Texas at Austin), Efrén Cruz Cortés (Northwestern University), John Cullinan (University of Galway), Tommaso D'Amelio (ECARES, Université Libre de Bruxelles), Sarah C. Dahmann (The University of Melbourne), Shangze Dai (University of Florida, Food and Resource Economics Department), Cem Mert Dalli (Aarhus University/University of Gothenburg), Justin Dang (University of San Diego), Thi Kim Dung Dang (KDI School of Public Policy and Management), Daniel Dao (University of Oxford), Valdonė Darškuvienė (ISM University of Management and Economics), Amreeta Das (Graduate Student, University of California Merced), Cole Davis (Brown University), Piersilvio De Bortoli (University of Trento), Florian De Bundel (UCLouvain & Fonds de la Recherche Scientifique - FNRS (F.R.S.-FNRS)), Alison de Farias (University of Southern California), Alexandra de Gendre (The University of Melbourne), Marjorie De la Cruz (Department of Public Policy and Administration, Florida International University, Miami, FL), Charles de Pierpont de Burnot (Université catholique de Louvain (UCLouvain), Louvain Institute of Data Analysis and Modeling in economics and statistics (LIDAM)), Lachlan Deer (University of Melbourne, Department of Marketing and Management), Matthew DeHaven (Brown University), Léonie Delpeyrou (University of Lausanne, IDHEAP), Michael Denly (Texas A&M University), Rose Deperrois (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France; GAEL, Univ. Grenoble Alpes, CNRS, INRAE, Grenoble INP, Grenoble, France), Scott Desposato (University of California, San Diego), Elodie Djemai (Université Paris-Dauphine PSL), Jesugnon Ezechias Djima (University of Houston), Ruggero Doino (World Bank), Yuanqing Dong (Stockholm School of Economics), Pierce Donovan (University of Nevada, Reno), Flóra Drucker (Düsseldorf Institute for Competition Economics, HHU Düsseldorf and ELTE KRTK), Zhuoran Du (University of New South Wales), Dimitri DUBOIS (CEE-M, Univ. Montpellier, CNRS, INRAe, Institut Agro, Montpellier, France), Julian Duggan (University of Southern California), Erwan

Dujeancourt (Stockholm University and Linnaeus University), Thibaut Duprey (Bank of Canada), Adam Akmal Dzulkipli (Centre for Health Economics, Monash University), Nour Elashmawy (World Bank), Ali Elminejad (Nazarbayev University), Erik Engberg (Örebro University), Julia Francesca Engel (Kiel University), Joseph Enguehard (Kiel Institute), Aaron Erlich (McGill University, Centre for the Study of Democratic Citizenship (CSDC)), Antal Ertl (ELTE KRTK & University of Iceland), Francesco Fallucchi (University of Bergamo), Valeria Fanghella (EDHEC Business School), Ismail El Fassi (University of St.Gallen), Youssef El Fassi (HEC Lausanne, University of Lausanne), Joshua Ferrer (American University), Maria Elena Filippin (Central Bank of Ireland and Uppsala University), Kira Finan (University College Dublin), Olivia Finan (University College Dublin), Blaine Finstein (University of California, Los Angeles), Jack Fitzgerald (Stanford University, Vrije Universiteit Amsterdam, and Institute for Replication), Francesca Foliano (University of Southampton), Frank M. Fossen (University of Nevada-Reno), Felix Atta Fosu (Smith School of Business, Queen's University), Giotto Frean (Victoria University of Wellington), Dimitria Freitas (Humboldt University, Thuenen Institute, TU Dresden), Joris Frese (Nuffield College, University of Oxford), Johann D Gaebler (Harvard University), Sebastian Gallegos (UAI Business School), Rachel Ganly (University of Oxford), Thomas Gareau-Paquette (Cornell University), Attila Gáspár (ELTE Centre for Economic and Regional Studies; Central European University), Romain Gauriot (Deakin University), Diogo Geraldes (University College Dublin; Universidade de Coimbra), Forough Ghadamyari (University of Victoria), Michele Giannetti (University of Padua), Julia Girard (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France; GAEL, Univ. Grenoble Alpes, CNRS, INRAE, Grenoble INP, Grenoble, France), Alexandros-Christos Gkotinakos (Aristotle University of Thessaloniki), Paula E. Gobbi (ECARES, Université Libre de Bruxelles), Joshua Goetz (University of California, Los Angeles (Political Science)), Felipe Gonzalez (King's College London and PUC-Chile), Pedro Gonzalez-Fernandez (Heidelberg University), Adrien Gosselin-Pali (University of Namur, CRED, DEFIPP, Belgium), Mark Granberg (Ratio institute), Pau Grau-Vilalta (Princeton University), Míde Griffin (Trinity College Dublin), Hasmik Grigoryan (University of Galway, J.E. Cairnes School of Business and Economics), Jiancheng Gu (University for Continuing Education Krems), Xhezaira Guleksi (University of Erlangen-Nuremberg), Gaygysyz Guljanov (University of Muenster), Liwen Guo (Department of Social Medicine and Health Management, Xiangya School of Public Health, Central South University), Florian Habermann (University of Lausanne), Joanne Haddad (Georgetown University), Fabian Haggerty (LMU Munich), Timo Haller (Berlin School of Economics (BSoE)), Adam Hamdan (Brown University), Rida Hameed (American University), Marie-lin Hamel (Université Laval), Olle Hammar (Linnaeus University and Stockholm School of Economics), Barry Hashimoto (Independent Researcher), Christina Sarah Hauser (University College Dublin and UCD Geary Institute for Public Policy), Karl Hauser (Université Grenoble Alpes), Reginald Hebert (Georgia State University), Florian Heine (Vrije Universiteit Amsterdam), Benjamin Helms (Texas A&M University), Moritz Hennicke (Maastricht University, Université libre de Bruxelles), Matthew Hepplewhite (University of Oxford), Andrew Hess (UCLA), John Hicks (University of Wisconsin–Madison, Department of Political Science), Kien Hoang-Le (Queen Mary University of London), Matthew Holian (San Jose State University), Felix Holzmeister (Department of Economics, University of Innsbruck), Joonkyo Hong (National Taiwan University), Daniel Horn (Corvinus University Budapest; ELTE KRTK), Sven Hörner (University of Bayreuth), Yue Hua (International University of Japan), Jingzhou (Ron) Huang (Chicago Booth), Kathleen Hui (Department of Justice), Sara

Hutchinson Tovar (Monash University), John Iselin (The Budget Lab at Yale), Boris Ivanov (Institute for Employment Research (IAB)), Niklas Jakobsson (Karlstad Business School, Karlstad University), Bo-Yeon Jang (University College Dublin), Michael Jankowski (Independent Researcher), Joakim Jansson (Linnaeus University), Joachim Jarreau (Université Paris-Dauphine PSL), Michael Jetter (University of Western Australia), Mofei Jia (Xi'an Jiaotong-Liverpool University), Donghua Karry Jiao (University of Cambridge), Dominik Jockers (Heidelberg Institute of Global Health, Heidelberg University), Magnus Johannesson (Department of Economics, Stockholm School of Economics), Naimi Johansson (Department of clinical science and education, Karolinska Institutet), Indigo Jones (Forward College & Elinor Ostrom Research Unit, France), Jan Kabatek (The University of Melbourne), Kanishka Kacker (Indian Statistical Institute), Zhanna Kapsalyamova (Nazarbayev University), Lamis Kattan (Georgetown University), Edmund Kelly (University of Oxford), Aicha Kharazi (University of Warwick), Martin Kihlstedt (Copenhagen Business School), Do Won Kim (University of Maryland), Inae Kim (Seoul National University), Kathleen Kimmey, Sorravich Kingsuwankul (Emlyon Business School, GATE, CNRS, Université Lyon 2, Université Jean Monnet Saint-Etienne, Lyon, 69007, France), Kiss Hubert János (ELTE KRTK & Corvinus University of Budapest), Slawomir Konopa (Warsaw University of Life Sciences; Indiana University Bloomington, Ostrom Workshop), Orestis Kopsacheilis (Technical University of Munich; University of Crete), Noah Kouchekinia (University of California, Irvine), Yao Thibaut Kpegli (University of Pau et des Pays de l'Adour), Adithya Krishna (Macquarie University), Pijus Krūminas (ISM University of Management and Economics), Jona Krutaj (Center for Decision Research and Experimental Economics, University of Nottingham, UK), Essi Kujansuu (University of Innsbruck), Ezgi Kurt (QMUL), Emily Lucinda Kurtz (Carleton College), Christoph F. Kurz (Ludwig-Maximilians-Universität Munich, Germany), Matthew Kutam (Brown University), Andrea La Nauze (Deakin University, University of Queensland and CESifo), Asher Labovich (Brown University), Martial Laguerre (EMLYON Business School & Université Claude Bernard Lyon 1), Parijat Lal (World Bank Group), Gregor Lauth (Humboldt University Berlin and RFBerlin), Antoine Leblois (CEE-M (Univ Montpellier, CNRS, INRAE, Institut Agro), Montpellier, France), John Lee (university), Lukas Lehner (University College London), Emilio Lehoucq (Northwestern University), Andrea Marietta Leina (University of Bologna), Francesca Lipari (Facultad de Ciencias Económicas y Empresariales Universidad Complutense de Madrid), Sarah Listabarth (TU Dresden), Haoxuan Liu (Cornerstone Research), Jingle Liu (World Bank), Yang Liu (Boston College), Yang Liu (The University of Melbourne), Chenhui Lu (University of Florida, Department of Economics), Shijia Luo (International Monetary Fund), Davide Luparello (Florida State University), Félix Lupien (Université du Québec à Montréal (UQAM)), Alexander Lutsenko (Northwestern University), Ke Lyu (University of Nevada, Reno), Anna Mader (Department of Agricultural Economics, Humboldt University Berlin, Germany), Nicolas Mäder (University of San Diego, Knauss School of Business), Prateek Mahajan (University of Texas at Austin), Simon Maier (Ulm University - Institute of Economics), Christina Maimone (Northwestern University), Anne-Gaëlle Maltese (University of Lille, CLERSE), Andrea Manera (IMF), Chenyu Mao (World Bank), Steeve Marchand (University of Melbourne), Jan Marcus (Freie Universität Berlin), Giovanni Paolo Mariani (ECARES - Universite Libre de Bruxelles), Diego Marino-Fages (Durham Unviersity), Franco Mariuzzo (University of East Anglia), Aidas Masiliunas (University of Sheffield), Damien Mayaux (University Paris Dauphine - PSL), Nathan McCoslin (Texas A&M University (Department of Political Science)), Sean McCulloch (Brown Univer-

sity), Jack McLaughlin (Monash University), Ryan McWay (University of Minnesota), Maria Medellin Esguerra (Brown University), Muhammad Zia Mehmood (Occidental College), Kevin Mei (University of Texas at Austin), Francesca Meli (University of Bristol), Wietse Mesman (Tilburg University), Catherine Michaud-Leclerc (Université Laval), Fabian Mierisch (Independent Researcher), Cuichi Miess (LMU Munich), Joshua B. Miller (University of Melbourne), Dibya Deepta Mishra (Rice University), Belay Moges (Dilla University), Morteza Mohandes Mojarrad (Tilburg University), Victoria Mooers (Columbia University), Guillaume Morlet (ETH Zurich), Andrei Munteanu (Université du Québec à Montréal), Kapil Narwal (Ivey Business School, Western University, Canada), Nicolas Nastri (American University), Frieder Neunhoeffer (Universidade de Lisboa, ISEG Lisbon School of Economics and Management, ISEG Research), Hang Anh (Tia) Nguyen (University of Wollongong), P. Linh Nguyen (French Institut for Demographic Studies, University of Mannheim), Viet Nguyen-Tien (London School of Economics), Giorgi Nikolaishvili (Wake Forest University), Vincent Notte (PhD student at UCLouvain (belgium)), Albert Nsengumuremyi (Department of Agricultural Economics, Humboldt University Berlin, Germany), Patrick Nüß (Max Planck Institute for Human Development, WZB Berlin), Stephen Daniel O'Connell (Emory university), Kento Ohara (University of Oxford), Cedric Okou (International Monetary Fund), Miguel Olivo-Villabrille (Australian National University), Tim Ölkers (Department of Agricultural Economics, Humboldt University Berlin, Germany), Joseph T. Ornstein (University of Georgia), Christian Oswald (University of the Bundeswehr Munich), Neerbewendé Abdoul Rachid PAFADNAM (Clermont Auvergne University (CERDI)), Fatma Selcen Palut (Erasmus University Rotterdam), Scott Robert Patterson (Université de Montréal), Nathaniel Pattison (Southern Methodist University), Stefanie Pechar (Technical University of Munich), Fidel Petros (WZB Berlin), Linh Pham (Lake Forest College), Tho Pham (University of York), Florian Ploeckl (Adelaide University), Maria Polipciuc (Friedrich Schiller University), Jakub Procházka (Masaryk University, Faculty of Economics and Administration), Roger Prudon (Lancaster University), Klajdi Puka (Western University), Pengzhan Qian (Royal Holloway University of London), Prasanthi Ramakrishnan (Southern Methodist University), Alejandra Ramos (Trinity College Dublin), Ryan D. Ratcliff (University of San Diego), Romy Reggiani Theodat (University of Ottawa), David Reinstein (Director of The Unjournal), Yohan Renard (University of Orléans, Department of Economics, France), Morgane Rigaux (Université Libre de Bruxelles (ECARES)), Sabrina Rittinger (Ulm University), Lorenzo Rocco (University of Padova), Maria Clara Rodrigues (Yale University), Guido A. Romero (Emory University), Julie Rosaz (Université Bourgogne Europe, Burgundy School of Business, CEREN EA 7477, 21000 Dijon, France.), Paul Rosmer (Humboldt University of Berlin and Berlin School of Economics), Jonathan Roth (Brown University), Rebecca Chepkoech Rotich (University of Edinburgh), Christina Rott (Vrije Universiteit Amsterdam and Tinbergen Institute), Nicolas P. Rougier (Centre Inria de l'université de Bordeaux), Victor Rudakov (Freie Universität Berlin), Nicolas Rudolf (University of Lausanne), Meike Rudolph (Kiel University & Institute for Employment Research (IAB)), Alessia Russo (University of Padua), Erica Ryan (University of Maryland, College Park and Amazon), Martin Søyland (University of Oslo), Galiya Sagyndykova (Nazarbayev University), Camille Saint-Macary (IRD, UMR LEDa, PSL, Université Paris-Dauphine, France), Orkun Saka (City St George's, University of London), Nicolas Salamanca (The University of Melbourne), Zeeshan Samad (American University of Beirut), Alexandra Sandström (Uppsala University), Sudipta Sarangi (Virginia Tech), Theodoros Saroglou (Interdisciplinary Transformation University Austria), Marcello Sartarelli (Universidad Complutense de

Madrid & ICAE), Alexandre Sauquet (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France), Viraj Sawant (University of California Los Angeles), Alvise Scarabosio (UC Berkeley), Francesco Scarazzato (Vienna University of Economics and Business), Katrin Schefbänker (University Innsbruck), David Schulze (Ruhr-University Bochum, Germany), Julia Schwarz (Kiel University), Stefan Seifert (University of Göttingen, Department of Agricultural Economics and Rural Development), Mateo Servent (Centro de Investigación y Docencia Económicas (CIDE)), Alex Sharples (Scarlatti), Isaac Shon (Columbia University), William Sidemo-Holm (Lund University), Claire Siebenmorgen (Georg-August Universität Göttingen), Christoph Siemroth (University of Essex), Rui Silva (University of East Anglia), Larissa Simoes (Instituto de Pesquisa Econômica Aplicada, Brazil), Wai Yan Siu (Institute for Coastal Adaptation and Resilience and Department of Economics, Old Dominion Univeristy), Balasurya Sivakumar (Oslo Metropolitan University), Ingvild Lindgren Skarpeid (FAIR and Norwegian School of Economics), Dislène Senan Sossou (University Laval), Raphael SOUBEYRAN (CEE-M, University of Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France), Wickramasooriya Ayodhya Vinodani Sovis (University of Kiel, Germany), Cristina Specchi (University of Bologna), Bhavya Srivastava (World Bank Group), Brian Stacy (World Bank), Paul Stainier (University of Pennsylvania), Dennis Steinle (Ulm University, Study Commission of Management and Economics), Dustin S. Stoltz (Lehigh University), Graham Straus (Texas A&M University), Andreas Strobl (MCI Management Center Innsbruck), Julie Subervie (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France), Pu Sun (Dongbei University of Finance and Economics), Vsevolod Suschevskiy (Linköping University), Alex Sutherland (University of Oxford, Department of Social Policy & Intervention), Adhipradana Prabu Swasito (The University of Western Australia), Oleksandr Talavera (University of Birmingham), Jotham Talemwa (University of Mannheim), Tze Yong Tan (Department of Economics and SoDa Labs, Monash University), Saloni Taneja (University of Southern California), Georgy Tarasenko (Cornell University), Boyd Tarlinton (Queensland Department of Primary Industries, Australia), Karin Teichmann (University of Innsbruck), Vincent Teixeira (University Mohammed VI Polytechnic, Africa Institute for Research in Economics and Social Sciences (AIRESS)), Dagmawe Tenaw (University of Padova), Sofia Tirraf (European centre for advanced research in economics and statistics (ECARES), at ULB), Martial Toniotti (LIDAM/CORE, UCLouvain), Matthew Tonkli (University of Sydney), Bastien Tourenc (Université Lumière Lyon 2, LARHRA), Dario Trujano-Ochoa (Texas A&M University), Theodore L. Turocy (University of East Anglia), Maria Ulasik (Department of Economics, University of Innsbruck), Justin Valasek (Norwegian School of Economics (NHH)), David Valenta (University of Ottawa), Giovanni van Empel (The University of Melbourne), Carlos Eduardo van Hombeeck (International Monetary Fund), Vidhulaa Vangal (Centre for Effective Governance of Indian States, India), Simon Varaine (Univ. Grenoble Alpes, IEPG, CNRS, INRAE, Grenoble INP, GAEL, Grenoble 38000, France), Gabrielle Vasey (Concordia University), Francisco Vazquez (International Monetary Fund), Hoa Vo (American University), David S Waller (University of Technology Sydney), Julian Walterskirchen (University of Gothenburg), Eric Wang (University of Washington), Jiayi (Harper) Wang (Columbia University), Albert Ward (University of Cambridge), Taylor Watson (Boston University), Christoph Weber (ESSCA School of Management), Xinyu Wei (Université libre de Bruxelles), Michael Wiebe (Independent researcher), Timo Wochner (ETH Zurich), Soyoung Woo (University of Oxford), Amanda Woodman Deza (Ifo Institute), Victor Y. Wu (Stanford University), Wanrui Wu (Peking University), Zhuangchen Wu (University of

Birmingham), Feihong Xu (Northwestern University), Qianyan Xu (Lund university), Yaroslav Yakymovych (Uppsala University), Yifan Yang (Stockholm School of Economics), Shuo Yu (UC Berkeley), Minjae Yun (Lowe Institute of Political Economy, Claremont McKenna College), Tahreen Zahra (Carleton University), Mirela Zaneva (University of Oxford), Donald Zarate (University of California, Riverside), Arezou Zaresani (University of Sydney), Aline Zayat (University of Ottawa), Chi ZHANG (Kyushu University), Junjun Zhang (University of East Anglia PhD candidate), Le Zhang (Macquarie University), mao zhang (university of st andrews), Tianqi Zhang (University of Southern California), Yueshu Zhao (International Monetary Fund), Yu Zheng (Queen Mary University of London), Lingyun Zhou (University of Chicago), Ziyue Zhu (Research Centre for Education and the Labour Market (ROA), Maastricht University), Federico Zilio (University of Melbourne), Abel Brodeur (University of Ottawa and Institute for Replication)

A.2 Authors' Contribution

Preparation of tables, figures, and manuscript (manuscript's author order within category). Nikolai Cook (Wilfrid Laurier University), Derek Mikola (University of Ottawa and Institute for Replication), Juan P. Aparicio (University of Ottawa and Institute for Replication), Lenka Fiala (University of Ottawa and Institute for Replication), Ghina Abdul Baki (University of Ottawa and Institute for Replication), Abel Brodeur (University of Ottawa and Institute for Replication)

Conception or design of the work (manuscript's author order within category). Nikolai Cook (Wilfrid Laurier University), Derek Mikola (University of Ottawa and Institute for Replication), Juan P. Aparicio (University of Ottawa and Institute for Replication), Lenka Fiala (University of Ottawa and Institute for Replication), Abel Brodeur (University of Ottawa and Institute for Replication)

Analysis or interpretation of data (manuscript's author order within category). Nikolai Cook (Wilfrid Laurier University), Derek Mikola (University of Ottawa and Institute for Replication), Juan P. Aparicio (University of Ottawa and Institute for Replication), Lenka Fiala (University of Ottawa and Institute for Replication), Praveena Bandara (Net Zero Industrial Policy Lab), Anand Butler (Trinity College), Attila Gáspár (ELTE Centre for Economic and Regional Studies), Rida Hameed (Lahore University of Management Sciences), Olle Hammar (Linnaeus University), Barry Hashimoto (Emory University), Felix Holzmeister (University of Innsbruck), Diego Marino Fages (Durham University), Nicolas Nastri (American University), Marcello Sartarelli (Complutense University of Madrid), Oleksandr Talavera (University of Birmingham), Tze Yong Tan (Monash University), Georgy Tarasenko (Cornell University), Hoa Vo (American University), Christoph Weber (ESSCA School of Management), Abel Brodeur (University of Ottawa and Institute for Replication)

Author of reproduction report (surname alphabetical). Alexander Cassiel Abajian (Yale School of the Environment), Kseniya Abanokova (World Bank), Mamdouh Abdelkader (Department of Economics, University of Ottawa), Sahar Abi-Hassan (Department of Political Science, University of Tulsa), Mohamed Alhadi A Abubraidaa (Simon Fraser University), Joop Adema (University of Innsbruck), George Agyeah (Walmart Inc), Farida Akhtar (Macquarie University), Shumi Akhtar (The University of Sydney), Dinara Alpysbayeva (Faculty of Economics and Administration, Masaryk University), Micah Altman (Center for Research in Equitable and Open Scholarship, Massachusetts Institute of Technology), Flavia Alves (Carleton University), Satyam Anand (Georgetown University), Mirinda Lee (Martin) Anderson (Brigham Young University-Idaho), Zubaria Andlib (Lancaster University, the UK, Federal Urdu University of Arts, Science and Technology, Islamabad, Pakistan), David Angenendt (SKEMA Business School, Université Côte d'Azur (GREDEG),), Konstantin Ash (StrategEast), Thomas Ash (Anderson School of Management, University of California, Los Angeles), Tobias Auer (London School of Economics), Antonio Avila-Uribe (Institution and university), Idrissa Aladji AYA (Université Clermont Auvergne, CNRS, IRD, CERDI, F-63000 CLERMONT FERRAND, France.), Abdelkarim Azami idrissi (University of lorraine, Bureau d'Économie Théorique et Appliquée (BETA)), Marc Bachmeier (Freie Universität Berlin), Lior Bakalo (The Faculty of Data and Decision Sciences, Technion – Israel Institute of Technology, Haifa, Israel), E. Grant Baldwin (University of California, Los Angeles), Amanda Baptista Denham (University-

UTS), Söhnke M. Bartram (University of Warwick and CEPR), Gabriel Bayle (GATE, CNRS, Université Lumière Lyon 2, Université Jean-Monnet Saint-Etienne, emlyon business school), Dana Bazarkulova (Nazarbayev University), Clayton Becker (ULCA), Louis-Philippe Beland (Carleton University, Department of Economics), Kendon Bell (Scarlatti), David Benatia (HEC Montréal), Guillaume Bérard (Luxembourg Institute of Socio-Economic Research (LISER)), Marta Bernardi (Research Associate and PhD Candidate at the Dresden University of Technology), Tim Bersak (Wofford College), Simone Bertoli (Université Clermont Auvergne, CNRS, IRD, CERDI), Konrad Bierl (Resource Economics Group, Humboldt University Berlin; Berlin School of Economics), Paul Binder (LMU Munich), Boris Blagov (RWI - Leibniz Institute for Economic Research), Lisa Bogler (Department of Economics and Centre for Modern Indian Studies, Georg-August-Universität Göttingen), Carl Bonander (Karlstad University), Arianna Bondi (McGill University), Stephanie Bonds (RAND Corporation), Adil Boufarsi (Université Grenoble Alpes, INRAE, UMR GAEL, 38000 Grenoble, France.), Adam Bouyamourn (Center for Data Science, NYU), Thomas Brailey (University of Oxford), Michelle Bronsard (Stanford University), Roberto Brunetti (LEMMA, Université Paris Panthéon-Assas), Lukas Brunninger (Vienna University of Economics and Business), Jens Christian Brunstad (University of Bergen), Calvin Bryan (Washington and Lee University), Nino Buliskeria (Nazarbayev University), Emil Bustos (Independent (previously: Research Institute of Industrial Economics)), Nienke Buters (Technical University Munich (TUM)), Alistair Cameron (Laterite Ltd), Gewei Cao (Max Planck Institute for the Study of Crime, Security and Law), Michelle Escobar Carias (Newcastle University Business School, University of Melbourne (Department of Economics)), Florian Caro (Yale University), Alice Carraro (University of Padova), Fortuna Casoria (Université Bourgogne Europe, Burgundy School of Business, CEREN EA 7477, 21000 Dijon, France.), Antoine Castet (University of Rouen Normandy, LERN), Brian Castro-Aguirre (University of Houston), Simonas Čepėnas (ISM University of Management and Economics), Luca Cermak (Institute of Economics, Ulm University), Marek Chadim (Yale University), Cédric Chambru (École normale supérieure de Lyon, CERGIC, EMR CNRS EqCo 5001), Patrick Chester (Stevens Institute of Technology), Erik Chi (Texas A&M University), Hongbi Choi (Texas A&M University (Department of Political Science)), Taesoo Choi (Department of Agricultural and Applied Economics, University of Wisconsin-Madison), Shih-Hsien Chuang (Northwest Missouri State University), Costin Ciobanu (Aarhus University), Elías Cisneros (The University of Texas at Dallas), Melchior Clerc (Université Clermont Auvergne, CNRS, IRD and CERDI), Guila Cohen (McGill University), Adrien Coiffard (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France), Jason Collins (University of Technology Sydney), Matthew Collins (University of Galway), Kenneth Colombe (University of Bonn, Department of Economics), Mariarosaria Comunale (International Monetary Fund), Marie Connolly (Université du Québec à Montréal), Mattia Coppo (Multilateral Investment Guarantee Agency (MIGA)), Guillaume Coqueret (EMLYON Business School), Cesar Cornejo (London School of Hygiene & Tropical Medicine (LSHTM), University College London (UCL)), Claudio Costanzo (ECARES), Jacques Courbe (UCLA Department of Political Science), Mathieu Couttenier (Ecole Normale Superieure de Lyon), Paolo Crosetto (INRAE), Andrés Cruz (University of Texas at Austin), Efrén Cruz Cortés (Northwestern University), John Cullinan (University of Galway), Tommaso D'Amelio (ECARES, Université Libre de Bruxelles), Sarah C. Dahmann (The University of Melbourne), Shangze Dai (University of Florida, Food and Resource Economics Department), Cem Mert Dalli (Aarhus University/University of Gothenburg), Justin Dang (University of San Diego), Thi Kim

Dung Dang (KDI School of Public Policy and Management), Daniel Dao (University of Oxford), Valdonė Darškuvienė (ISM University of Management and Economics), Amreeta Das (Graduate Student, University of California Merced), Cole Davis (Brown University), Piersilvio De Bortoli (University of Trento), Florian De Bundel (UCLouvain & Fonds de la Recherche Scientifique - FNRS (F.R.S.-FNRS)), Alison de Farias (University of Southern California), Alexandra de Gendre (The University of Melbourne), Marjorie De la Cruz (Department of Public Policy and Administration, Florida International University, Miami, FL), Charles de Pierpont de Burnot (Université catholique de Louvain (UCLouvain), Louvain Institute of Data Analysis and Modeling in economics and statistics (LIDAM)), Lachlan Deer (University of Melbourne, Department of Marketing and Management), Matthew DeHaven (Brown University), Léonie Delpeyrou (University of Lausanne, IDHEAP), Michael Denly (Texas A&M University), Rose Deperrois (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France; GAEL, Univ. Grenoble Alpes, CNRS, INRAE, Grenoble INP, Grenoble, France), Scott Desposato (University of California, San Diego), Elodie Djemai (Université Paris-Dauphine PSL), Jesugnon Ezechias Djima (University of Houston), Ruggero Doino (World Bank), Yuanqing Dong (Stockholm School of Economics), Pierce Donovan (University of Nevada, Reno), Flóra Drucker (Düsseldorf Institute for Competition Economics, HHU Düsseldorf and ELTE KRTK), Zhuoran Du (University of New South Wales), Dimitri DUBOIS (CEE-M, Univ. Montpellier, CNRS, INRAe, Institut Agro, Montpellier, France), Julian Duggan (University of Southern California), Erwan Dujeancourt (Stockholm University and Linnaeus University), Thibaut Duprey (Bank of Canada), Adam Akmal Dzulkipli (Centre for Health Economics, Monash University), Nour Elashmawy (World Bank), Ali Elminejad (Nazarbayev University), Erik Engberg (Örebro University), Julia Francesca Engel (Kiel University), Joseph Enguehard (Kiel Institute), Aaron Erlich (McGill University, Centre for the Study of Democratic Citizenship (CSDC)), Antal Ertl (ELTE KRTK & University of Iceland), Francesco Fallucchi (University of Bergamo), Valeria Fanghella (EDHEC Business School), Ismail El Fassi (University of St.Gallen), Youssef El Fassi (HEC Lausanne, University of Lausanne), Joshua Ferrer (American University), Maria Elena Filippin (Central Bank of Ireland and Uppsala University), Kira Finan (University College Dublin), Olivia Finan (University College Dublin), Blaine Finstein (University of California, Los Angeles), Jack Fitzgerald (Stanford University, Vrije Universiteit Amsterdam, and Institute for Replication), Francesca Foliano (University of Southampton), Frank M. Fossen (University of Nevada-Reno), Felix Atta Fosu (Smith School of Business, Queen's University), Giotto Frean (Victoria University of Wellington), Dimitria Freitas (Humboldt University, Thuenen Institute, TU Dresden), Joris Frese (Nuffield College, University of Oxford), Johann D Gaebler (Harvard University), Sebastian Gallegos (UAI Business School), Rachel Ganly (University of Oxford), Thomas Gareau-Paquette (Cornell University), Romain Gauriot (Deakin University), Diogo Geraldes (University College Dublin; Universidade de Coimbra), Forough Ghadamyari (University of Victoria), Michele Giannetti (University of Padua), Julia Girard (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France; GAEL, Univ. Grenoble Alpes, CNRS, INRAE, Grenoble INP, Grenoble, France), Alexandros-Christos Gkotinakos (Aristotle University of Thessaloniki), Paula E. Gobbi (ECARES, Université Libre de Bruxelles), Joshua Goetz (University of California, Los Angeles (Political Science)), Felipe Gonzalez (King's College London and PUC-Chile), Pedro Gonzalez-Fernandez (Heidelberg University), Adrien Gosselin-Pali (University of Namur, CRED, DEFIPP, Belgium), Mark Granberg (Ratio institute), Pau Grau-Vilalta (Princeton University), Míde Griffin (Trinity College Dublin), Hasmik Grigoryan (Uni-

versity of Galway, J.E. Cairnes School of Business and Economics), Jiancheng Gu (University for Continuing Education Krems), Xhezaira Guleksi (University of Erlangen-Nuremberg), Gaygysyz Guljanov (University of Muenster), Liwen Guo (Department of Social Medicine and Health Management, Xiangya School of Public Health, Central South University), Florian Habermann (University of Lausanne), Joanne Haddad (Georgetown University), Fabian Haggerty (LMU Munich), Timo Haller (Berlin School of Economics (BSoE)), Adam Hamdan (Brown University), Marie-lin Hamel (Université Laval), Christina Sarah Hauser (University College Dublin and UCD Geary Institute for Public Policy), Karl Hauser (Université Grenoble Alpes), Reginald Hebert (Georgia State University), Florian Heine (Vrije Universiteit Amsterdam), Benjamin Helms (Texas A&M University), Moritz Hennicke (Maastricht University, Université libre de Bruxelles), Matthew Hepplewhite (University of Oxford), Andrew Hess (UCLA), John Hicks (University of Wisconsin–Madison, Department of Political Science), Kien Hoang-Le (Queen Mary University of London), Matthew Holian (San Jose State University), Joonkyo Hong (National Taiwan University), Daniel Horn (Corvinus University Budapest; ELTE KRTK), Sven Hörner (University of Bayreuth), Yue Hua (International University of Japan), Jingzhou (Ron) Huang (Chicago Booth), Kathleen Hui (Department of Justice), Sara Hutchinson Tovar (Monash University), John Iselin (The Budget Lab at Yale), Boris Ivanov (Institute for Employment Research (IAB)), Niklas Jakobsson (Karlstad Business School, Karlstad University), Bo-Yeon Jang (University College Dublin), Michael Jankowski (Independent Researcher), Joakim Jansson (Linnaeus University), Joachim Jarreau (Université Paris-Dauphine PSL), Michael Jetter (University of Western Australia), Mofei Jia (Xi'an Jiaotong-Liverpool University), Donghua Karry Jiao (University of Cambridge), Dominik Jockers (Heidelberg Institute of Global Health, Heidelberg University), Magnus Johannesson (Department of Economics, Stockholm School of Economics), Naimi Johansson (Department of clinical science and education, Karolinska Institutet), Indigo Jones (Forward College & Elinor Ostrom Research Unit, France), Jan Kabatek (The University of Melbourne), Kanishka Kacker (Indian Statistical Institute), Zhanna Kapsalyamova (Nazarbayev University), Lamis Kattan (Georgetown University), Edmund Kelly (University of Oxford), Aicha Kharazi (University of Warwick), Martin Kihlstedt (Copenhagen Business School), Do Won Kim (University of Maryland), Inae Kim (Seoul National University), Kathleen Kimmey, Sorravich Kingsuwankul (Emlyon Business School, GATE, CNRS, Université Lyon 2, Université Jean Monnet Saint-Etienne, Lyon, 69007, France), Kiss Hubert János (ELTE KRTK & Corvinus University of Budapest), Slawomir Konopa (Warsaw University of Life Sciences; Indiana University Bloomington, Ostrom Workshop), Orestis Kopsacheilis (Technical University of Munich; University of Crete), Noah Kouchekinia (University of California, Irvine), Yao Thibaut Kpegli (University of Pau et des Pays de l'Adour), Adithya Krishna (Macquarie University), Pijus Krūminas (ISM University of Management and Economics), Jona Krutaj (Center for Decision Research and Experimental Economics, University of Nottingham, UK), Essi Kujansuu (University of Innsbruck), Ezgi Kurt (QMUL), Emily Lucinda Kurtz (Carleton College), Christoph F. Kurz (Ludwig-Maximilians-Universität Munich, Germany), Matthew Kutam (Brown University), Andrea La Nauze (Deakin University, University of Queensland and CESifo), Asher Labovich (Brown University), Martial Laguerre (EMLYON Business School & Université Claude Bernard Lyon 1), Parijat Lal (World Bank Group), Gregor Lauth (Humboldt University Berlin and RFBerlin), Antoine Leblois (CEE-M (Univ Montpellier, CNRS, INRAE, Institut Agro), Montpellier, France), John Lee (university), Lukas Lehner (University College London), Emilio Lehoucq (Northwestern University), Andrea Marietta Leina (Uni-

versity of Bologna), Francesca Lipari (Facultad de Ciencias Económicas y Empresariales Universidad Complutense de Madrid), Sarah Listabarth (TU Dresden), Haoxuan Liu (Cornerstone Research), Jingle Liu (World Bank), Yang Liu (Boston College), Yang Liu (The University of Melbourne), Chenhui Lu (University of Florida, Department of Economics), Shijia Luo (International Monetary Fund), Davide Luparello (Florida State University), Félix Lupien (Université du Québec à Montréal (UQAM)), Alexander Lutsenko (Northwestern University), Ke Lyu (University of Nevada, Reno), Anna Mader (Department of Agricultural Economics, Humboldt University Berlin, Germany), Nicolas Mäder (University of San Diego, Knauss School of Business), Prateek Mahajan (University of Texas at Austin), Simon Maier (Ulm University - Institute of Economics), Christina Maimone (Northwestern University), Anne-Gaëlle Maltese (University of Lille, CLERSE), Andrea Manera (IMF), Chenyu Mao (World Bank), Steeve Marchand (University of Melbourne), Jan Marcus (Freie Universität Berlin), Giovanni Paolo Mariani (ECARES - Universite Libre de Bruxelles), Franco Mariuzzo (University of East Anglia), Aidas Masiliunas (University of Sheffield), Damien Mayaux (University Paris Dauphine - PSL), Nathan McCoslin (Texas A&M University (Department of Political Science)), Sean McCulloch (Brown University), Jack McLaughlin (Monash University), Ryan McWay (University of Minnesota), Maria Medellin Esguerra (Brown University), Muhammad Zia Mehmood (Occidental College), Kevin Mei (University of Texas at Austin), Francesca Meli (University of Bristol), Wietse Mesman (Tilburg University), Catherine Michaud-Leclerc (Université Laval), Fabian Mierisch (Independent Researcher), Cuichi Miess (LMU Munich), Joshua B. Miller (University of Melbourne), Dibya Deepta Mishra (Rice University), Belay Moges (Dilla University), Morteza Mohandes Mojarrad (Tilburg University), Victoria Mooers (Columbia University), Guillaume Morlet (ETH Zurich), Andrei Munteanu (Université du Québec à Montréal), Kapil Narwal (Ivey Business School, Western University, Canada), Frieder Neunhoeffer (Universidade de Lisboa, ISEG Lisbon School of Economics and Management, ISEG Research), Hang Anh (Tia) Nguyen (University of Wollongong), P. Linh Nguyen (French Institut for Demographic Studies, University of Mannheim), Viet Nguyen-Tien (London School of Economics), Giorgi Nikolaishvili (Wake Forest University), Vincent Notte (PhD student at UCLouvain (belgium)), Albert Nsengumuremyi (Department of Agricultural Economics, Humboldt University Berlin, Germany), Patrick Nüß (Max Planck Institute for Human Development, WZB Berlin), Stephen Daniel O'Connell (Emory university), Kento Ohara (University of Oxford), Cedric Okou (International Monetary Fund), Miguel Olivo-Villabrille (Australian National University), Tim Ölkers (Department of Agricultural Economics, Humboldt University Berlin, Germany), Joseph T. Ornstein (University of Georgia), Christian Oswald (University of the Bundeswehr Munich), Neerbewendé Abdoul Rachid PAFADNAM (Clermont Auvergne University (CERDI)), Fatma Selcen Palut (Erasmus University Rotterdam), Scott Robert Patterson (Université de Montréal), Nathaniel Pattison (Southern Methodist University), Stefanie Pechar (Technical University of Munich), Fidel Petros (WZB Berlin), Linh Pham (Lake Forest College), Tho Pham (University of York), Florian Ploeckl (Adelaide University), Maria Polipciuc (Friedrich Schiller University), Jakub Procházka (Masaryk University, Faculty of Economics and Administration), Roger Prudon (Lancaster University), Klajdi Puka (Western University), Pengzhan Qian (Royal Holloway University of London), Prasanthi Ramakrishnan (Southern Methodist University), Alejandra Ramos (Trinity College Dublin), Ryan D. Ratcliff (University of San Diego), Romy Reggiani Theodat (University of Ottawa), David Reinstein (Director of The Unjournal), Yohan Renard (University of Orléans, Department of Economics, France), Morgane Rigaux (Univer-

sité Libre de Bruxelles (ECARES)), Sabrina Rittinger (Ulm University), Lorenzo Rocco (University of Padova), Maria Clara Rodrigues (Yale University), Guido A. Romero (Emory University), Julie Rosaz (Université Bourgogne Europe, Burgundy School of Business, CEREN EA 7477, 21000 Dijon, France.), Paul Rosmer (Humboldt University of Berlin and Berlin School of Economics), Jonathan Roth (Brown University), Rebecca Chepkoech Rotich (University of Edinburgh), Christina Rott (Vrije Universiteit Amsterdam and Tinbergen Institute), Nicolas P. Rougier (Centre Inria de l'université de Bordeaux), Victor Rudakov (Freie Universität Berlin), Nicolas Rudolf (University of Lausanne), Meike Rudolph (Kiel University & Institute for Employment Research (IAB)), Alessia Russo (University of Padua), Erica Ryan (University of Maryland, College Park and Amazon), Martin Søyland (University of Oslo), Galiya Sagyndykova (Nazarbayev University), Camille Saint-Macary (IRD, UMR LEDa, PSL, Université Paris-Dauphine, France), Orkun Saka (City St George's, University of London), Nicolas Salamanca (The University of Melbourne), Zeeshan Samad (American University of Beirut), Alexandra Sandström (Uppsala University), Sudipta Sarangi (Virginia Tech), Theodoros Saroglou (Interdisciplinary Transformation University Austria), Alexandre Sauquet (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France), Viraj Sawant (University of California Los Angeles), Alvise Scarabosio (UC Berkeley), Francesco Scarazzato (Vienna University of Economics and Business), Katrin Schefbänker (University Innsbruck), David Schulze (Ruhr-University Bochum, Germany), Julia Schwarz (Kiel University), Stefan Seifert (University of Göttingen, Department of Agricultural Economics and Rural Development), Mateo Servent (Centro de Investigación y Docencia Económicas (CIDE)), Alex Sharples (Scarlatti), Isaac Shon (Columbia University), William Sidemo-Holm (Lund University), Claire Siebenmorgen (Georg-August Universität Göttingen), Christoph Siemroth (University of Essex), Rui Silva (University of East Anglia), Larissa Simoes (Instituto de Pesquisa Econômica Aplicada, Brazil), Wai Yan Siu (Institute for Coastal Adaptation and Resilience and Department of Economics, Old Dominion Univeristy), Balasurya Sivakumar (Oslo Metropolitan University), Ingvild Lindgren Skarpeid (FAIR and Norwegian School of Economics), Dislène Senan Sossou (University Laval), Raphael SOUBEYRAN (CEE-M, University of Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France), Wickramasooriya Ayodhya Vinodani Sovis (University of Kiel, Germany), Cristina Specchi (University of Bologna), Bhavya Srivastava (World Bank Group), Brian Stacy (World Bank), Paul Stainier (University of Pennsylvania), Dennis Steinle (Ulm University, Study Commission of Management and Economics), Dustin S. Stoltz (Lehigh University), Graham Straus (Texas A&M University), Andreas Strobl (MCI Management Center Innsbruck), Julie Subervie (CEE-M, Univ. Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France), Pu Sun (Dongbei University of Finance and Economics), Vsevolod Suschevskiy (Linköping University), Alex Sutherland (University of Oxford, Department of Social Policy & Intervention), Adhipradana Prabu Swasito (The University of Western Australia), Jotham Talemwa (University of Mannheim), Saloni Taneja (University of Southern California), Boyd Tarlinton (Queensland Department of Primary Industries, Australia), Karin Teichmann (University of Innsbruck), Vincent Teixeira (University Mohammed VI Polytechnic, Africa Institute for Research in Economics and Social Sciences (AIRESS)), Dagmawe Tenaw (University of Padova), Sofia Tirraf (European centre for advanced research in economics and statistics (ECARES), at ULB), Martial Toniotti (LIDAM/CORE, UCLouvain), Matthew Tonkli (University of Sydney), Bastien Tourenc (Université Lumière Lyon 2, LARHRA), Dario Trujano-Ochoa (Texas A&M University), Theodore L. Turocy (University of East Anglia), Maria Ulasik (Department of Eco-

nomics, University of Innsbruck), Justin Valasek (Norwegian School of Economics (NHH)), David Valenta (University of Ottawa), Giovanni van Empel (The University of Melbourne), Carlos Eduardo van Hombeeck (International Monetary Fund), Vidhulaa Vangal (Centre for Effective Governance of Indian States, India), Simon Varaine (Univ. Grenoble Alpes, IEPG, CNRS, INRAE, Grenoble INP, GAEL, Grenoble 38000, France), Gabrielle Vasey (Concordia University), Francisco Vazquez (International Monetary Fund), David S Waller (University of Technology Sydney), Julian Walterskirchen (University of Gothenburg), Eric Wang (University of Washington), Jiayi (Harper) Wang (Columbia University), Albert Ward (University of Cambridge), Taylor Watson (Boston University), Xinyu Wei (Université libre de Bruxelles), Michael Wiebe (Independent researcher), Timo Wochner (ETH Zurich), Soyoung Woo (University of Oxford), Amanda Woodman Deza (Ifo Institute), Victor Y. Wu (Stanford University), Wanrui Wu (Peking University), Zhuangchen Wu (University of Birmingham), Feihong Xu (Northwestern University), Qianyan Xu (Lund university), Yaroslav Yakymovych (Uppsala University), Yifan Yang (Stockholm School of Economics), Shuo Yu (UC Berkeley), Minjae Yun (Lowe Institute of Political Economy, Claremont McKenna College), Tahreen Zahra (Carleton University), Mirela Zaneva (University of Oxford), Donald Zarate (University of California, Riverside), Arezou Zaresani (University of Sydney), Aline Zayat (University of Ottawa), Chi ZHANG (Kyushu University), Junjun Zhang (University of East Anglia PhD candidate), Le Zhang (Macquarie University), mao zhang (university of st andrews), Tianqi Zhang (University of Southern California), Yueshu Zhao (International Monetary Fund), Yu Zheng (Queen Mary University of London), Lingyun Zhou (University of Chicago), Ziyue Zhu (Research Centre for Education and the Labour Market (ROA), Maastricht University), Federico Zilio (University of Melbourne)

B ONLINE APPENDIX B

B.1 Pre-Analysis Plan

Pre-Analysis Plan

2nd Meta Paper: Determinants of Reproducibility and Replicability*

Abel Brodeur, Derek Mikola and Nikolai Cook

April 6th, 2024

Abstract

In this study, we expand on our first meta paper and reproduce and replicate the main claims from additional empirical papers published in leading economic and political science outlets. We will also include a number of reproductions and replications from *Nature Human Behaviour*. For each study, we worked in small teams and either conducted a robustness reproduction — which is the ability to duplicate the results of a prior study using the same data but different procedures as were used by the original investigator — or recoded the study using the raw or intermediate data. Some teams did a combination of a robustness reproduction, recoding, or using new data. In this pre-analysis plan, we detail our 2nd meta paper's research questions and methodology.

*Corresponding Author: Brodeur: University of Ottawa and Institute for Replication. E-mail: abrodeur@uottawa.ca.

1 Introduction

In this study, we expand on our first meta paper and reproduce and replicate the main claims from additional empirical papers published in leading economic and political science outlets. We will also include a small number of reproductions and replications from *Nature Human Behaviour*. As with the first meta paper, for each study, we worked in small teams and either conducted a robustness reproduction — which is the ability to duplicate the results of a prior study using the same data but different procedures as were used by the original investigator — or recoded the study using the raw or intermediate data. Some teams did a combination of a robustness reproduction, recoding, or using new data.

Our main goal is to further document the determinants of robustness reproducibility and replicability. In what follows, we describe our main research questions and our methodology. More specifically, we tackle the following questions:

1. "Does reproducibility/replicability rate depend on the methods use?"
2. "Does reproducibility/replicability rate depend on the methods use?"
3. "Does reproducibility/replicability rate vary by subfield?"
4. "Are researchers able to predict robustness reproducibility rates by method and subfield?"

We will also tackle several secondary research questions. These questions will be the same as in the first meta paper by the Insitute for Replication (i.e., computational reproducibility rate, prevalence of coding errors and robustness reproducibility rate), but with our augmented sample.

This pre-analysis plan was written and pre-registered on April 6th, 2024.

2 Hypotheses

Based on our research questions, we will test the following hypotheses:

1. Reproducibility/replicability rates vary by method.
2. Reproducibility/replicability rates vary by subfield.
3. Researchers accurately predict robustness reproducibility rates by method and subfield.

3 Definitions

Reproducibility is the examination of whether the results and conclusions of original studies can be duplicated using the original studies' data, while **replicability** is defined as using data other than what was used in the original studies.

Reproducibility is further delineated into three categories. **Computational reproducibility** gauges the extent to which results in original studies can be reproduced using both the data and code from those studies. **Recreate reproducibility** assesses the extent to which results can be reproduced using the information in the original studies without access to the processed data set and/or the analysis code. **Robustness reproducibility** explores the extent to which results in original studies remain robust to alternative plausible analytical decisions, utilizing the same data as in the original studies.

Replicability is also classified into two types. **Direct replicability** evaluates the extent to which results in original studies can be repeated on new data using the original studies' research design and analysis. This classification is further subdivided based on whether data from the same population, a similar population, or a different population is employed. **Conceptual replicability** employs new data to assess the extent to which results in original studies can be repeated; however, this type of replication involves an alternative research design and/or alternative analysis to test the same hypothesis as in the original study. Conceptual replicability is also further subdivided into the same three categories based on populations which are the same, similar or different.

There are key definitions which we will use throughout the pre-analysis plan which better explains the different roles of people in relationship to this paper.

> **Original Author(s)** are the individual(s) who have published a paper in one of the targeted journals.
>
> **Original Paper** refers to the paper published in one of the targeted journals by the original authors.
>
> **Replicator(s)** are the individual(s) who conducted a reproduction or replication of an original paper written by an original author.
>
> **Replication Report** is the written report by the replicators documenting their findings while conducting the reproduction/replication of the original paper by the original authors.

We will rely on multiple dependent variables:

First Dependent Variable: dummy variable indicating whether the re-analysis is statistically significant at 5% level and same sign. For this dependent variable, we only keep original estimates statistically significant at the 5% level.

Second Dependent Variable: dummy variable indicating whether the re-analysis is statistically significant at 10% level and same sign. For this dependent variable, we only keep original estimates statistically significant at the 10% level.

Third Dependent Variable: dummy variable indicating whether the re-analysis remains not statistically significant at 5% level. For this dependent variable, we only keep original estimates statistically insignificant at the 5% level.

Fourth Dependent Variable: dummy variable indicating whether the re-analysis remains not statistically significant at 10% level. For this dependent variable, we only keep original estimates statistically insignificant at the 10% level.

We will also rely on relative t-statistics. As there can be multiple re-analysis estimates per original estimate, we first take the average of the re-analysis estimates by original estimate and take the ratio. Then, in order for all re-analyses to have the same effect, we average those ratios at the re-analysis level.

4 Methodology

For this paper, our focus is on the following nine economic journals: (1) *American Economic Review*, (2) *American Economic Review: Insights*, (3) *American Economic Journal: Applied Economics*, (4) *American Economic Journal: Economic Policy* and (5) *American Economic Journal: Macroeconomics*, (6) *The Economic Journal*, (7) *Journal of Political Economy*, (8) *Quarterly Journal of Economics* and (9) *Review of Economic Studies*. For political science, our focus is on three journals: (1) *American Journal of Political Science*, (2) *American Political Science Review* and (3) *Journal of Politics*. We have also begun reproducing and replicating papers from *Nature Human Behaviour* which will likely be added to our Meta Database.

5 Methods and Subfields

We will present results only for methods and subfields for which we have enough statistical power.

The methods which we will be comparing are: (i) instrumental variable, (ii) randomized control trial, (iii) regression discontinuity design, (iv) difference-in-differences, (v) laboratory experiment, (vi) online experiment, (vii) matching and

synthetic control methods, (viii) time series, (ix) structural modelling and (x) descriptibve/non-causal estimation.

The subfields which we will be comparing are: (i) American politics, (ii) comparative politics, (iii) international relations, (iV) JEL C, (v) JEL D, (vi) JEL E, (vii) JEL G, (viii) JEL H, (ix) JEL I, (x) JEL J, (xi) JEL L, (xii) JEL N and (xiii) JEL O. JEL refers to the Journal of Economic Literature classification. We exclude other JEL codes as we are almost certain we will not have enough statistical power at the time when we anticipate writing the second meta paper.

6 Reproduction and Replication Process

We have two streams to generate reproductions and replications.

I4R's Board. First, I4R has a board of editors (https://i4replication.org/people.html) who recommend potential replicators. All board members are nominated by the lead author, A.B. He then reaches out to the board for suggestions of replicators who could be a good fit for the studies in the targeted journals.

Replication games. Our second stream to generate reproductions and replications is the replication games (RGs). RGs are one-day meet-ups open to faculty, postdocs, graduate students and other researchers.

Replicators first identify the main claims, check for computational reproducibility, and are then free to conduct any robustness or recoding exercises.

Replication Reports

Two of the lead authors (A.B. and D.M.) and research assistants read replication reports and copy test statistics into an Excel file. We also code and group robustness reproducibility and replicability exercises, and information on computational reproducibility and coding errors. The work being entered by RAs is checked by A.B. or D.M. for completeness and accuracy. If any part of any entry was unclear, they were checked again and discussed. In some circumstances, we reach out to the replicator(s) to clarify parts of their reports.

7 Meta Database

The Meta Database forms the basis for our analysis. As of March 2024, 110 replication reports were coded into the Meta Database. We hope to have at least 200 replication reports at the time of completing this project.

At the time of writing this pre-analysis plan, we did not code the methods nor the field for these 110 replication reports.

8 Computational Reproducibility and Coding Errors

We will first document computational reproducibility and coding errors by method and subfield.

For computational reproducibility, we will rely on the Social Science Reproduction Platform (SSRP)'s 10-point scale to document computational reproducibility. (Visit here for more details on SSRP and this scale).

We will also document the prevalence and severity of coding errors by method and subfield. The errors will be broadly categorized into errors of the dependent variable, main independent variable, control variables, estimation, inference, sample/observations and other.

9 Changes in Statistical Significance, Effect Size, and the Reproducibility and Replicability Rate

For each method and subfield, we will compare the re-analysis estimates in the meta-database against their originally published counterparts. First, we will compare statistical significance both visually and with a suite of state-of-the-art tests of publication bias. Second, we will compare the relative effect size of re-analysis estimates. Third, we will detail how originally published estimates 'move' from statistical significance to insignificance (and vice versa) during the re-analysis process. We then identify which types of re-analysis have the best (and worst) replication rates.

9.1 Statistical Significance

We will present both t-statistics and p-curves. We will restrict the visualization to $|t| \in [0, 5]$, present bins of width 0.1, and present an Epanechnikov kernel (with standard errors in blue) which softens valleys and peaks. We will provide reference lines at the conventional two-tailed significance levels.

For the p-curves, we will illustrate the p-values $\in [0.0025, 0.1500]$, with bins of width 0.0025. We will remove $p < 0.0025$ for illustrative purposes only.

9.2 Relative t-Statistics

As mentioned before, we will also rely on relative t-statistics. We will first take the average of the re-analysis estimates by original estimate and take the ratio. Then,

we will average those ratios at the re-analysis level. We will present those relative t-statistics by method and subfield.

9.3 Relative Effect Size

For relative effect size, we restrict the analysis to originally published estimates that were statistically significant.

10 Many-Analysts Approach: Authors' Experience and Prestige, and Data and Code Availability

The main text will present figures and tables describing computational reproducibility and robustness reproducibility/replicability across methods and subfields. We then move to a more careful empirical examination of the relationship between robustness reproducibility/replicability, methods and subfields.

Data for Meta-Analysts

The backbone of the data to be provided to meta analysts will be the Meta Database, of which questions from surveys will be added.

Method

The goal is to have each team answer each research question independently. Each team received the same instructions and data. We will allow full flexibility to all teams. Teams will be allowed to use any statistics package, statistical model, inference, weighting scheme, *etc.* Teams will be free to choose the independent variables and how to code them. Teams will also be free to construct their own derived variables from the dataset given to them.

11 Predictions

We will gather predictions from researchers for the following questions (in addition to sociodemographic questions):

1. Have you heard anything about the results of this study, even informally?
2. Do you have prior experience with reproduction and replication, or other open science initiatives?
3. What percent of results/tests do you expect to be fully computationally reproducible for the following methods and subfields?

4. Excluding minor issues like missing packages, broken pathways, and versioning issues, what percent of studies do you expect to have at least one coding error for the following methods and subfields?

5. What do you expect the robustness reproducibility rate across claims to be for the following methods and subfields?

6. What percent of re-analyses effect size estimates do you expect to be smaller (in magnitude) than the original published estimates for the following methods and subfields?

C ONLINE APPENDIX C

C.1 List of Reproduction Reports

C.1.1 Reproduction Report

Title Original Study: A Friend like Me: The Effect of International Organization Membership on State Preferences

doi: https://doi.org/10.1086/722348, Journal of Politics

Link to Full Report and Authors' Response: https://osf.io/t5gah/

Original Authors' Package: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/5XG84N

C.1.2 Reproduction Report

Title Original Study: Acceptance of Inequality Between Children: Large-Scale Experimental Evidence from China and Norway

doi: https://doi.org/10.1093/ej/ueae097, Economic Journal

Link to Full Report and Original Authors' Response: https://osf.io/hr358/

Original Authors' Package: https://doi.org/10.5281/zenodo.13981874

C.1.3 Reproduction Report

Title Original Study: Age Set versus Kin: Culture and Financial Ties in East Africa

doi: 10.1257/aer.20211856, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/3tqb9/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/192224/version/V1/view

C.1.4 Reproduction Report

Title Original Study: Agenda Control and Timing of Bill Initiation: A Temporal Perspective on Coalition Governance in Parliamentary Democracies

doi: https://doi.org/10.1017/S0003055421000897, American Political Science Review

Link to Full Report: https://osf.io/y3cms/

Link to Original Authors' Response: https://osf.io/y3cms/

Original Authors' Package: https://doi.org/10.7910/DVN/OFCAAE

C.1.5 Reproduction Report

Title Original Study: Agricultural Diversity, Structural Change, and Long-run Development

doi: https://doi.org/10.1257/mac.20190285, American Economic Journal: Macroeconomics

Link to Full Report and Authors' Response: https://osf.io/mb9hj/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/119556/version/V3/view

C.1.6 Reproduction Report

Title Original Study: An Outbreak of Selective Attribution: Partisanship and Blame in the COVID-19 Pandemic

doi: https://doi.org/10.1017/S0003055423000047, American Political Science Review

Link to Full Report: https://osf.io/5gjpt/

Link to Original Authors' Response: https://osf.io/5gjpt/

Original Authors' Package: https://doi.org/10.7910/DVN/3FF05Y

C.1.7 Reproduction Report

Title Original Study: Anchored Inflation Expectations

doi: https://doi.org/10.1257/mac.20200080, American Economic Journal: Macroeconomics

Link to Full Report: https://osf.io/v7jtq/

Link to Original Authors' Response: https://osf.io/v7jtq/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/147021/version/V1/view

C.1.8 Reproduction Report

Title Original Study: Anticipatory Anxiety and Wishful Thinking

doi: https://doi.org/10.1257/aer.20191068, American Economic Review

Link to Full Report: https://osf.io/pxz42/

Original Authors' Response: https://osf.io/pxz42/ "On behalf of the authors I [Peter Schwardmann] thank the replicators for their effort and for helping to improve science. We have nothing to add to the report. The replicators uncovered a typo in Table 1 and the description of Experiment 3. The maximum number of trials in this experiment was indeed 96 and not 64 as we had erroneously stated."

Original Authors' Package: https://doi.org/10.3886/E195781V1

C.1.9 Reproduction Report

Title Original Study: Antinormative Messaging, Group Cues, and the Nuclear Ban Treaty

doi: https://doi.org/10.1086/714924, Journal of Politics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/97.htm

Replication Package: https://osf.io/xbvzg/

Link to Original Authors' Response: https://econpapers.repec.org/paper/zbwi4rdps/98.htm

Original Authors' Package: https://doi.org/10.7910/DVN/GLT4FX

C.1.10 Reproduction Report

Title Original Study: Ascriptive Characteristics and Perceptions of Impropriety in the Rule of Law: Race, Gender, and Public Assessments of Whether Judges Can Be Impartial

doi: https://doi.org/10.1111/ajps.12599, American Journal of Political Science

Link to Full Report: https://osf.io/yf48r/

Replication Package: https://osf.io/yf48r/

Link to Original Authors' Response: No response.

Original Authors' Package: https://doi.org/10.7910/DVN/ZHOL6Y

C.1.11 Reproduction Report

Title Original Study: Asset-Based Microfinance for Microenterprises: Evidence from Pakistan

doi: https://doi.org/10.1257/aer.20210169, American Economic Review

Full Report: The reproducers only conducted computational reproducibility.

Original Authors' Package: https://doi.org/10.3886/E195984V1

C.1.12 Reproduction Report

Title Original Study: Assortative Matching at the Top of the Distribution: Evidence from the World's Most Exclusive Marriage Market

doi: https://doi.org/10.1257/app.20180463, American Economic Journal: Applied Economics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/47.htm

Link to Replicators' Package: https://osf.io/pqsem/

Original Author's Response: "I now reviewed the report carefully and with interest, and I am glad to see that the authors succeeded in replicating all results and found no coding errors. I hope the replication package was clear and easy to work with. I am also happy to see that they performed several additional robustness checks and heterogeneity analysis, and that these show that the original estimates "are robust and are not significantly affected using these alternative specifications" (p. 1). Given this, and the replicators' conclusion that "the study's main findings demonstrate robustness and reliability" (p. 7), I think that there is nothing substantial for me to write in a response in the form of a discussion paper. This is because both the replication exercise and the additional analysis found no major issues in the original work to respond to. I would also like to thank the authors for the fairness and professionalism of their report, and also for the time and effort they put in producing it, from which I ultimately benefit — as it adds to the credibility of my original paper — as well as the profession as a whole benefits — as making replication exercises more common is important for economics.

Please let me know if I can be of any further assistance regarding this Reproduction Report in the future. I am at your or the authors' disposal, in case I can be of help in clarifying anything in the replication package or in the analysis of the original paper. As I stated above, I believe that increasing replication rates is important for our field, as it is making original datasets publicly available — even when, as in the case of my paper, the data collection is an important part of my contribution, and in this situation, many do not grant public access to the original data."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/140921/version/V1/view

C.1.13 Reproduction Report

Title Original Study: Belief Elicitation and Behavioral Incentive Compatibility

doi: https://doi.org/10.1257/aer.20201248, American Economic Review

Link to Full Report: https://osf.io/jf479

Authors' Response: We did not get a response from the authors.

Original Authors' Package: https://doi.org/10.3886/E157161V1

C.1.14 Reproduction Report

Title Original Study: Black Workers in White Places: Daytime Racial Diversity and White Public Opinion

doi: https://doi.org/10.1086/716289, Journal of Politics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/61.htm

Link to Replicators' Package: https://osf.io/ue4pm/

Original Authors' Response: "We enjoyed reading the replication, and don't see a need to write a response.

Thank you for doing this important work."

Original Authors' Package: https://doi.org/10.7910/DVN/FMOR6K

C.1.15 Reproduction Report

Title Original Study: Brahmin Left Versus Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948–2020

doi: https://doi.org/10.1093/qje/qjab036, Quarterly Journal of Economics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/19.htm

Link to Replicators' Package: https://osf.io/2hpeq/

Original Authors' Response: "Thank you for your mail and for your interesting report! We are happy to see that you were able to easily replicate our results and that our main conclusions were found to be largely robust. In this context, we do not think that an answer from our side would be particularly useful: we are happy with the report as it is.

Thank you for the very valuable work that your institute is producing in testing the replicability and robustness of published studies!"

Original Authors' Package: https://doi.org/10.7910/DVN/XUSWG6

C.1.16 Reproduction Report

Title Original Study: Bridging the Intention-Behavior Gap? The Effect of Plan-Making Prompts on Job Search and Employment

doi: https://doi.org/10.1257/app.20170566, American Economic Journal: Applied Economics

Link to Full Report: Only conducted computational reproducibility.

Original Authors' Package: https://doi.org/10.3886/E113734V1

C.1.17 Reproduction Report

Title Original Study: Bubbles, Crashes, and Economic Growth: Theory and Evidence

doi: https://doi.org/10.1257/mac.20220015, American Economic Journal: Macroeconomics

Link to Full Report: https://osf.io/d76tn/

Link to Replicators' Package: https://osf.io/d76tn/

Original Authors' Response: Provided a short response and answered a question. Did not provide a final response as of November 2025.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/173441/version/V1/view

C.1.18 Reproduction Report

Title Original Study: Building Nations through Shared Experiences: Evidence from African Football

doi: https://doi.org/10.1257/aer.20180805, American Economic Review

Link to Full Report and Original Authors' Response: https://osf.io/gj7v3/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/115011/version/V1/view

C.1.19 Reproduction Report

Title Original Study: Bundling Genetic and Financial Technologies for More Resilient and Productive Small-Scale Farmers in Africa

doi: https://doi.org/10.1093/ej/ueae012, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/b4vs3/

Original Authors' Package: https://doi.org/10.5281/zenodo.10629265

C.1.20 Reproduction Report

Title Original Study: Campaign Contributions and Roll-Call Voting in the U.S. House of Representatives: The Case of the Sugar Industry

doi: https://doi.org/10.1017/S0003055422000466, American Political Science Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/57.htm

Link to Replicators' Package: https://osf.io/4hjb9/

Original Authors' Final Response: "We thank the Institute for Replication for their diligent work replicating and performing some extensions to our 2023 APSR paper. Replication is an important and often undervalued work in the scientific process. Of course we are quite pleased to see that our results do replicate and that the extensions performed largely support the results and ideas we advanced in our paper. Keep up the good work!"

Original Authors' Package: https://doi.org/10.7910/DVN/2IFZR9

C.1.21 Reproduction Report

Title Original Study: Can Financial Incentives to Firms Improve Apprenticeship Training? Experimental Evidence from Ghana

doi: https://doi.org/10.1257/aeri.20220696, American Economic Review: Insights

Link to Full Report and Authors' Response: https://osf.io/7wjmy/

Original Authors' Package: https://doi.org/10.3886/E192366V1

C.1.22 Reproduction Report

Title Original Study: Can Information Reduce Ethnic Discrimination? Evidence from Airbnb

doi: https://doi.org/10.1257/app.20190188, American Economic Journal: Applied Economics

Link to Full Report: https://osf.io/zn98a/

Link to Replicators' Package: https://github.com/TuanNguyen04/Replication_Airbnb

Original Authors' Response: The authors provided initial feedback which the replicators took it into account.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/120078/version/V1/view

C.1.23 Reproduction Report

Title Original Study: Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution

doi: https://doi.org/10.1257/aeri.20200373, American Economic Review: Insights

Link to Full Report: https://osf.io/b7dn2/

Link to Replicators' Package: https://osf.io/m8hfr/?view_only=9f6632ec96c0451daf0f8889b9ad2b25

Original Authors' Response: https://osf.io/b7dn2/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/125321/version/V1/view

C.1.24 Reproduction Report

Title Original Study: Can't We All Just Get Along? How Women MPs Can Ameliorate Affective Polarization in Western Publics

doi: https://doi.org/10.1017/S0003055422000491, American Political Science Review

Link to Full Report: https://osf.io/69px3/

Link to Replicators' Package: https://osf.io/69px3/

Original Authors' Response: Thank you for replicating our paper Can't We All Just Get Along? How Women MPs Can Ameliorate Affective Polarization in Western Publics (APSR 2023) as part of the Montreal Replication Games. We appreciate the attention to detail and rigor applied to the replication project. We are pleased that our initial results replicate well. We appreciate your robust approach to testing the stability of our findings using a country and year 'leave-one-out' cross-validation strategy. We also thank you for catching the coding error which dropped a handful of cases from the original analysis; we are glad that the results remain substantively the same when this error is corrected. We also are interested in the results from the extension you undertook, finding that our results are primarily driven by left-wing parties' supporters, in particular parties from the green, radical left and social Democratic parties. On the other hand, the point estimates are positive for all parties excepting the conservative and radical right parties, which can be expected to have the most conservative views on gender roles. We note that the authors' interpretation, that "the portion of women MPs affects the attitudes of left-wing voters and not the attitudes of the voters most likely to undermine democracy" is true, but that the results also suggest that far-right parties, who most aggressively challenge liberal democratic norms, may be able to "soften" their image among left-wing voters by running female candidates. This is consistent with the argument made by Catalano Week et al (2023), that radical right parties strategically run women to broaden their appeal. Again, we deeply appreciate your replication and insightful extension of our research.

Original Authors' Package: https://doi.org/10.7910/DVN/AHQRVR

C.1.25 Reproduction Report

Title Original Study: Carbon Taxes and CO2 Emissions: Sweden as a Case Study

doi: https://doi.org/10.1257/pol.20170144, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/j85gb/

Original Authors' Package: https://doi.org/10.1257/pol.20170144

C.1.26 Reproduction Report

Title Original Study: Changing Hearts and Minds? Why Media Messages Designed to Foster Empathy Often Fail

doi: https://doi.org/10.1086/719416, Journal of Politics

Link to Full Report and Authors' Response: https://osf.io/zes6g/

Original Authors' Package: https://doi.org/10.7910/DVN/FUCDTT

C.1.27 Reproduction Report

Title Original Study: Changing Tides: Public Attitudes on Climate Migration

doi: https://doi.org/10.1086/715163, Journal of Politics

Link to Full Report: https://www.socialsciencereproduction.org/reproductions/791/published/index

Original Authors' Response: "Thank you very much for reaching out! We are very pleased to hear that the results of our study were replicated, and do not need to provide an answer."

Original Authors' Package: https://doi.org/10.7910/DVN/FDML2N

C.1.28 Reproduction Report

Title Original Study: Checking and Sharing Alt-Facts

doi: https://doi.org/10.1257/pol.20210037, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/34.htm

Link to Replicators' Package: https://doi.org/10.5281/zenodo.7858829

Link to Original Authors' Response: "Many thanks! No, we won't be writing a response."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/140161/version/V1/view

C.1.29 Reproduction Report

Title Original Study: Child Marriage Bans and Female Schooling and Labor Market Outcomes: Evidence from Natural Experiments in 17 Low- and Middle-Income Countries

doi: https://doi.org/10.1257/pol.20200008, American Economic Journal: Economic Policy

Link to Full Report: https://osf.io/5yhxc/

Link to Replicators' Package: https://osf.io/5yhxc/

Original Authors' Response: We could not reach out the author.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/130784/version/V1/view

C.1.30 Reproduction Report

Title Original Study: Climate Change and Labor Reallocation: Evidence from Six Decades of the Indian Census (Team 1)

doi: https://doi.org/10.1257/pol.20210129, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/k9t7z/

Original Authors' Package: https://doi.org/10.3886/E150441V1

C.1.31 Reproduction Report

Title Original Study: Climate Change and Labor Reallocation: Evidence from Six Decades of the Indian Census (Team 2)

doi: https://doi.org/10.1257/pol.20210129, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/k9t7z/

Original Authors' Package: https://doi.org/10.3886/E150441V1

C.1.32 Reproduction Report

Title Original Study: Coase and Cap-and-Trade: Evidence on the Independence Property from the European Carbon Market (Team 1)

doi: https://doi.org/10.1257/pol.20210028, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/9g8rp

Original Authors' Package: https://doi.org/10.3886/E152861V1

C.1.33 Reproduction Report

Title Original Study: Coase and Cap-and-Trade: Evidence on the Independence Property from the European Carbon Market (Team 2)

doi: https://doi.org/10.1257/pol.20210028, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/9g8rp

Original Authors' Package: https://doi.org/10.3886/E152861V1

C.1.34 Reproduction Report

Title Original Study: Coercive Contract Enforcement: Law and the Labor Market in Nineteenth Century Industrial Britain

doi: https://doi.org/10.1257/aer.103.1.107, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Response: No formal response.

Original Authors' Package: https://doi.org/10.3886/E112585V1

C.1.35 Reproduction Report

Title Original Study: Cognitive Behavioral Therapy among Ghana's Rural Poor Is Effective Regardless of Baseline Mental Distress

doi: https://doi.org/10.1257/aeri.20210612, American Economic Review: Insights

Link to Full Report and Authors' Response: https://osf.io/ay83e/

Original Authors' Package: https://doi.org/10.3886/E164481V1

C.1.36 Reproduction Report

Title Original Study: Cognitive Performance in Remote Work: Evidence from Professional Chess

doi: https://doi.org/10.1093/ej/ueab094, Economic Journal

Full Report: Only computational reproducibility.

Original Authors' Package: https://doi.org/10.5281/zenodo.5674798

C.1.37 Reproduction Report

Title Original Study: College Enrollment and Mandatory FAFSA Applications: Evidence from Louisiana

doi: https://doi.org/10.1257/pol.20210360, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/hg98m/

Original Authors' Package: https://doi.org/10.3886/E161021V1

C.1.38 Reproduction Report

Title Original Study: Commercial Imperialism? Political Influence and Trade during the Cold War

doi: https://doi.org/10.1257/aer.103.2.863, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Authors' Response: No response.

Original Authors' Package: https://doi.org/10.3886/E112606V1

C.1.39 Reproduction Report

Title Original Study: Competition and Pass-Through: Evidence from Isolated Markets

doi: https://doi.org/10.1257/app.20200863, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/qhbmp/

Original Authors' Package: https://doi.org/10.3886/E135621V1

C.1.40 Reproduction Report

Title Original Study: Concentration Bias in Intertemporal Choice

doi: https://doi.org/10.1093/restud/rdab043, Review of Economic Studies

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/42.htm

Link to Replicators' Package: https://osf.io/d42xr/

Original Authors' Response: "We thank Deer, Ellingsrud, Heuer, and Kordt (2023) for conducting the Reproduction Report and appreciate that their "results confirm the sign, magnitude and statistical significance of [our] reported estimates across each of [our] five main findings" (p. 1). We don't have anything substantive to add to this. "

Original Authors' Package: https://zenodo.org/records/5091975

C.1.41 Reproduction Report

Title Original Study: Contested Killings: The Mobilizing Effects of Community Contact with Police Violence

doi: https://doi.org/10.1017/S0003055423000321, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/f9sdc/

Original Authors' Package: https://doi.org/10.7910/DVN/DMG0XD

C.1.42 Reproduction Report

Title Original Study: Convex Supply Curves

doi: https://doi.org/10.1257/aer.20210811, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/172781/version/V1/view

C.1.43 Reproduction Report

Title Original Study: Cooperative Property Rights and Development: Evidence from Land Reform in El Salvador

doi: https://doi.org/10.1086/717042, Journal of Political Economy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/20.htm

Link to Replicators' Package: https://doi.org/10.7910/DVN/AMD3NO

Link to Original Authors' Response: https://doi.org/10.1086/725234

Original Authors' Package: https://doi.org/10.1086/717042/suppl_file/20190161data.zip

C.1.44 Reproduction Report

Title Original Study: Coordination and Bandwagon Effects: How Past Rankings Shape the Behavior of Voters and Candidates

doi: https://doi.org/10.1257/app.20210840, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/r2jc5/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/172161/version/V1/view

C.1.45 Reproduction Report

Title Original Study: Corrupted by Algorithms? How AI-generated and Human-written Advice Shape (Dis)honesty

doi: https://doi.org/10.1093/ej/uead056, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/gqukr/

Original Authors' Package: https://doi.org/10.5281/zenodo.8170048

C.1.46 Reproduction Report

Title Original Study: Corruption and Co-Optation in Autocracy: Evidence from Russia

doi: https://doi.org/10.1017/S0003055424000340, American Political Science Review

Link to Full Report and Author's Response: https://osf.io/fq8vr/

Original Authors' Package: https://doi.org/10.7910/DVN/P2U84B

C.1.47 Reproduction Report

Title Original Study: Credit, Attention, and Externalities in the Adoption of Energy Efficient Technologies by Low-Income Households

doi: https://doi.org/10.1257/aer.20210766, American Economic Review

Link to Full Report: https://osf.io/ztwgq

Authors' Response: No response.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/166661/version/V1/view

C.1.48 Reproduction Report

Title Original Study: Decentralization Can Increase Cooperation among Public Officials

doi: https://doi.org/10.1111/ajps.12606, American Journal of Political Science

Link to Full Report: https://osf.io/q3dpt/

Link to Replicators' Package: https://osf.io/q3dpt/

Link to Original Authors' Response: https://osf.io/q3dpt/

Original Authors' Package: https://doi.org/10.7910/DVN/ZLHYSZ

C.1.49 Reproduction Report

Title Original Study: Declining Worker Turnover: The Role of Short-Duration Employment Spells

doi: https://doi.org/10.1257/mac.20190230, American Economic Journal: Macroeconomics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/93.htm

Link to Replicators' Package: https://github.com/AlexandrePavlov/PriesRogerson2022Replication

Original Authors' Response: Declined to respond.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/120568/version/V1/view

C.1.50 Reproduction Report

Title Original Study: Demand for Commitment in Credit and Saving Contracts: A Field Experiment

doi: https://doi.org/10.1093/ej/ueae053, Economic Journal

Link to Full Report and Authors' Response: Back and forth with authors not completed.

Original Authors' Package: https://doi.org/10.5281/zenodo.10651755

C.1.51 Reproduction Report

Title Original Study: Detecting Mother-Father Differences in Spending on Children: A New Approach Using Willingness-to-Pay Elicitation

doi: https://doi.org/10.1257/aeri.20220159, American Economic Review: Insights

Link to Full Report and Authors' Response: https://osf.io/3s7j8/

Original Authors' Package: https://doi.org/10.3886/E179162V1

C.1.52 Reproduction Report

Title Original Study: DETER-ing Deforestation in the Amazon: Environmental Monitoring and Law Enforcement

doi: https://doi.org/10.1257/app.20200196, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/rnezh/

Original Authors' Package: https://doi.org/10.3886/E132281V1

C.1.53 Reproduction Report

Title Original Study: Did Unconventional Interventions Unfreeze the Credit Market?

doi: https://doi.org/10.1257/mac.20160023, American Economic Journal: Macroeconomics

Link to Full Report and Authors' Response: https://osf.io/93ex6/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/231501/version/V1/view

C.1.54 Reproduction Report

Title Original Study: Digital Addiction

doi: https://doi.org/10.1257/aer.20210867, American Economic Review

Link to Full Report: https://osf.io/8kvdf/

Link to Replicators' Package: https://osf.io/8kvdf/

Link to Original Authors' Response: https://osf.io/8kvdf/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/163822/version/V2/view

C.1.55 Reproduction Report

Title Original Study: Digital Information Provision and Behavior Change: Lessons from Six Experiments in East Africa

doi: https://doi.org/10.1257/app.20220072, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/6ucyn/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/186241/version/V1/view

C.1.56 Reproduction Report

Title Original Study: Discriminatory Lending: Evidence from Bankers in the Lab

doi: https://doi.org/10.1257/app.20210180, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/yvxwd/

Original Authors' Package: https://doi.org/10.3886/E157121V1

C.1.57 Reproduction Report

Title Original Study: Discretionary Tax Changes and the Macroeconomy: New Narrative Evidence from the United Kingdom

doi: https://doi.org/10.1257/aer.103.4.1507, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Authors' Response: No response.

Original Authors' Package: https://doi.org/10.3886/E112651V1

C.1.58 Reproduction Report

Title Original Study: Distinguishing Common Ratio Preferences from Common Ratio Effects Using Paired Valuation Tasks

doi: https://doi.org/10.1257/aer.20221535, American Economic Review

Link to Full Report: Back and forth with authors not completed.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/194003/version/V1/view

C.1.59 Reproduction Report

Title Original Study: District Magnitude and Female Representation: Evidence from Argentina and Latin America

doi: https://doi.org/10.1111/ajps.12575, American Journal of Political Science

Link to Full Report: Back and forth with authors not completed.

Original Authors' Package: https://doi.org/10.7910/DVN/FXI5ZO

C.1.60 Reproduction Report

Title Original Study: Do Policy Makers Listen to Experts? Evidence from a National Survey of Local and State Policy Makers

doi: https://doi.org/10.1017/S0003055421000800, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/sw7f9/

Original Authors' Package: https://doi.org/10.7910/DVN/S2SNOT

C.1.61 Reproduction Report

Title Original Study: Do Thank-You Calls Increase Charitable Giving? Expert Forecasts and Field Experimental Evidence

doi: https://doi.org/10.1257/app.20210068, American Economic Journal: Applied Economics

Link to Full Report: https://osf.io/fe2tr/

Link to Replicators' Package: https://gitlab.com/c3754/replication-games/-/tree/main/replication%20games%20MTL%202023%20charity?ref_type=heads

Link to Original Authors' Response: Waiting for the authors' response.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/149481/version/V1/view

C.1.62 Reproduction Report

Title Original Study: Do Transitional Justice Museums Persuade Visitors? Evidence from a Field Experiment

doi: https://doi.org/10.1086/714765, Journal of Politics

Link to Full Report: https://osf.io/m3hwg/

Link to Replicators' Package: https://osf.io/m3hwg/

Original Authors' Response: "We thank all involved for their interest in our work. We are happy to hear that the results from our paper successfully replicated. We are intrigued by the additional analyses performed by the replicators. We hope their insights and results can inform future theorizing and empirical studies of the impact of Transitional Justice."

Original Authors' Package: https://doi.org/10.7910/DVN/TNFDDX

C.1.63 Reproduction Report

Title Original Study: Does Competence Make Citizens Tolerate Undemocratic Behavior?

doi: https://doi.org/10.1017/S0003055422000119, American Political Science Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/28.htm

Link to Replicators' Package: https://github.com/tjbrailey/nottingham_replication_2023

Link to Original Authors' Final Response: "Thanks a lot for this initiative and not least for replicating my results."

Original Authors' Package: https://doi.org/10.7910/DVN/NGFLRO

C.1.64 Reproduction Report

Title Original Study: Does Directed Innovation Mitigate Climate Damage? Evidence from U.S. Agriculture

doi: https://doi.org/10.1093/qje/qjac039, Quarterly Journal of Economics

Link to Full Report and Authors' Response: https://osf.io/f2jbz/

Original Authors' Package: https://doi.org/10.7910/DVN/5ELEPA

C.1.65 Reproduction Report

Title Original Study: Does Discipline Decrease Police Misconduct? Evidence from Chicago Civilian Allegations

doi: https://doi.org/10.1257/app.20200568, American Economic Journal: Applied Economics

Link to Full Report: https://osf.io/cx47j/

Original Authors' Response: No response.

Original Authors' Package: https://doi.org/10.3886/E143841V1

C.1.66 Reproduction Report

Title Original Study: Does Patient Demand Contribute to the Overuse of Prescription Drugs?

doi: https://doi.org/10.1257/app.20190722, American Economic Journal: Applied Economics

Link to Full Report: https://osf.io/x7g9z/

Link to Replicators' Package: https://osf.io/x7g9z/

Original Authors' Response: Provided feedback to an initial report. Final response: " Thank you very much for sharing the updated report. We appreciate that the authors of the replication reworked the paper and have no further response or comments."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/126722/version/V1/view

C.1.67 Reproduction Report

Title Original Study: Does Public Opinion Affect the Preferences of Foreign Policy Leaders? Experimental Evidence from the UK Parliament

doi: https://doi.org/10.1086/719007, Journal of Politics

Link to Full Report: https://osf.io/bqz6w/

Link to Replicators' Package: https://osf.io/vwt2n/?view_only=84e52a7c684942a4880410b3c89ff4c6

Original Authors' Response: " Thank you for your note and engaging with our work. We don't have a formal reply, though this is an honest question: isn't it standard practice to use weights when using YouGov's data, since making valid claims about representativeness depends on using their weights? YouGov's MP panels operate similarly to their public opinion poll, in that their claims to representativeness rely on using weights, provided by YouGov. I [Chu] think your write-up mentioned that there's a debate about using weights, and cited MTurk data, but I think that MTurk is quite different, and yes, I agree I do not use weights for MTurk data except unless requested by a reviewer for robustness checks, etc.. But I don't think MTurk and the MP representative poll we used is a good comparison in the context of evaluating the validity of weighting. In any case, happy to adapt if there is a clear consensus on this. Thanks again."

Original Authors' Package: https://doi.org/10.7910/DVN/BNINNL

C.1.68 Reproduction Report

Title Original Study: Do Motivated Classmates Matter for Educational Success?

doi: https://doi.org/10.1093/ej/ueae060, Economic Journal

Link to Full Report and Original Authors' Response: https://osf.io/2ac69/

Original Authors' Package: https://doi.org/10.5281/zenodo.11911078

C.1.69 Reproduction Report

Title Original Study: Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia

doi: https://doi.org/10.1093/qje/qjx040, Quarterly Journal of Economics

Link to Full Report: Successful computational reproducibility; no coding errors.

Authors' Response: Not contacted.

Original Authors' Package: https://doi.org/10.1093/qje/qjx040

C.1.70 Reproduction Report

Title Original Study: Economic preferences across generations and family clusters: A large-scale experiment in a developing country

doi: https://doi.org/10.1086/720395, Journal of Political Economy

Link to Full Report and Authors' Response: https://osf.io/wjzsm/

Original Authors' Package: https://www.journals.uchicago.edu/doi/suppl/10.1086/720395/suppl_file/20200664data.zip

C.1.71 Reproduction Report

Title Original Study: Effective for Whom? Ethnic Identity and Nonviolent Resistance

doi: https://doi.org/10.1017/S0003055421000940, American Political Science Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/86.htm

Link to Replicators' Package: https://zenodo.org/records/10193470

Original Authors' Response: Cannot provide a response.

Original Authors' Package: https://doi.org/10.7910/DVN/SHHVCA

C.1.72 Reproduction Report

Title Original Study: Electoral Cycles in Macroprudential Regulation

doi: https://doi.org/10.1257/pol.20200626, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/73hk9/

Original Authors' Package: https://doi.org/10.3886/E163121V1

C.1.73 Reproduction Report

Title Original Study: Enabling or Limiting Cognitive Flexibility? Evidence of Demand for Moral Commitment

doi: https://doi.org/10.1257/aer.20201333, American Economic Review

Link to Full Report: https://osf.io/nwds7

Link to Replicators' Package: https://osf.io/yfdet/

Link to Original Authors' Response: https://doi.org/10.1257/aer.20201333.appx

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/180741/version/V1/view

C.1.74 Reproduction Report

Title Original Study: Energy Saving May Kill: Evidence from the Fukushima Nuclear Accident

doi: https://doi.org/10.1257/app.20200505, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/tek63

Original Authors' Package: https://doi.org/10.3886/E170502V1

C.1.75 Reproduction Report

Title Original Study: Entertaining Beliefs in Economic Mobility

doi: https://doi.org/10.1111/ajps.12702, American Journal of Political Science

Link to Full Report: https://osf.io/xf5w2/

Link to Replicators' Package: https://osf.io/xf5w2/

Original Author's Response: "Thanks for this! I have no particular response per se. I'm grateful for your collective efforts to make social science much more transparent."

Original Authors' Package: https://doi.org/10.7910/DVN/FVRZYU

C.1.76 Reproduction Report

Title Original Study: Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty?

doi: https://doi.org/10.1257/aeri.20210014, American Economic Review: Insights

Link to Full Report and Authors' Response: https://osf.io/bfr3y/

Original Authors' Package: https://doi.org/10.3886/E183613V1

C.1.77 Reproduction Report

Title Original Study: Evaluating Deliberative Competence: A Simple Method with an Application to Financial Choice

doi: https://doi.org/10.1257/aer.20210290, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/scgbt/

Original Authors' Package: https://doi.org/10.3886/E171681V1

C.1.78 Reproduction Report

Title Original Study: Exploiting Externalities to Estimate the Long-Term Effects of Early Childhood Deworming

doi: https://doi.org/10.1257/app.20160183, American Economic Journal: Applied Economics

Full Report: Computational reproducibility conducted during the World Bank-IMF games.

Authors' Response: Not contacted.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/113689/version/V1/view

C.1.79 Reproduction Report

Title Original Study: Exposure and Preferences: Evidence from Indian Slums

doi: https://doi.org/10.1111/ajps.12570, American Journal of Political Science

Link to Full Report: No report.

Original Author's Response: "Thanks for your email. I am not interested in participating."

Original Authors' Package: https://doi.org/10.7910/DVN/AV8PLT

C.1.80 Reproduction Report

Title Original Study: Extraction, Assimilation, and Accommodation: The Historical Foundations of Indigenous–State Relations in Latin America

doi: https://doi.org/10.1017/S0003055423000333, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/5gxuf/

Original Authors' Package: https://doi.org/10.7910/DVN/GS838F

C.1.81 Reproduction Report

Title Original Study: Factor Market Failures and the Adoption of Irrigation in Rwanda

doi: https://doi.org/10.1257/aer.20210059, American Economic Review

Full Report and Authors' Response: Successful reproduction.

Original Authors' Package: https://doi.org/10.3886/E159061V1

C.1.82 Reproduction Report

Title Original Study: Family Heterogeneity, Human Capital Investment, and College Attainment

doi: https://doi.org/10.1257/mac.20200014, American Economic Journal: Macroeconomics

Link to Full Report and Original Authors' Response: Back and forth with authors not completed.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/135321/version/V1/view

C.1.83 Reproduction Report

Title Original Study: Fathers' Leave Reduces Sexist Attitudes

doi: https://doi.org/10.1017/S0003055423000369, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/j4csn/

Original Authors' Package: https://doi.org/10.7910/DVN/4DUB7X

C.1.84 Reproduction Report

Title Original Study: Female Representation and Legitimacy: Evidence from a Harmonized Experiment in Jordan, Morocco, and Tunisia

doi: https://doi.org/10.1017/S0003055423000357, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/hkc5b/

Original Authors' Package: https://doi.org/10.7910/DVN/5GP5SI

C.1.85 Reproduction Report

Title Original Study: Finance and Green Growth

doi: https://doi.org/10.1093/ej/ueac081, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/95.htm

Link to Replicators' Package: https://osf.io/h8ct2/

Link to Original Authors' Response: https://econpapers.repec.org/paper/zbwi4rdps/96.htm

Original Authors' Package: https://zenodo.org/records/7220094

C.1.86 Reproduction Report

Title Original Study: Flight to Safety: COVID-Induced Changes in the Intensity of Status Quo Preference and Voting Behavior

doi: https://doi.org/10.1017/S0003055421000691, American Political Science Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/36.htm

Link to Replicators' Package: https://github.com/Dmscates/Bisbee-and-Honig-2022-Flight-to-Safety-Replication

Original Authors' Response: "You guys are amazing. Thank you for doing this! We are impressed by your rigor and grateful for the introduction to DCD'H DiD estimator that we'll have to add to the repertoire. We were working on the conditional accept when the flurry of generalized DiD work (Goodman-Bacon, Callaway, etc.) was blowing up [...] We also appreciate the manner in which you communicated with us during the course of your re-analysis, and the thoughtfulness of your report. [...] Although I'm sure it is a logistical nightmare and likely would add even more delays to the publication pipeline, it would be very pro-science if this type of replication were part of a journal's own pre-publication process. (This is what I naively thought replication meant back when I got my first publication, and have been disappointed in the process ever since.)"

Original Authors' Package: https://doi.org/10.7910/DVN/S5YMS7

C.1.87 Reproduction Report

Title Original Study: Formative Experiences and the Price of Gasoline

doi: https://doi.org/10.1257/app.20200407, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/d53bc

Original Authors' Package: https://doi.org/10.3886/E127261V1

C.1.88 Reproduction Report

Title Original Study: From Hashtag to Hate Crime: Twitter and Antiminority Sentiment

doi: https://doi.org/10.1257/app.20210211, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/zyndc/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/158541/version/V1/view

C.1.89 Reproduction Report

Title Original Study: Fuelling Organised Crime: the Mexican War on Drugs and Oil Theft

doi: https://doi.org/10.1093/ej/ueae047, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/nr8vt/

Original Authors' Package: https://doi.org/10.5281/zenodo.11236210

C.1.90 Reproduction Report

Title Original Study: Gender Differences in Cooperative Environments? Evidence from The U.S. Congress

doi: https://doi.org/10.1093/ej/ueab069, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/75.htm

Link to Replicators' Package: https://www.dropbox.com/scl/fo/dmvgx5wlgql3tz98dx47u/h?rlkey=af83xiqrkw70rbjicdunoach9&dl=0

Link to Original Authors' Response: https://osf.io/w48tf/

Original Authors' Package: https://zenodo.org/records/5111360

C.1.91 Reproduction Report

Title Original Study: Getting Dynamic Implementation to Work

doi: https://doi.org/10.1086/721153, Journal of Political Economy

Full Report: Successful reproduction.

Original Authors' Package: https://www.journals.uchicago.edu/doi/suppl/10.1086/721153/suppl_file/20200049data.zip

C.1.92 Reproduction Report

Title Original Study: Going to a Better School: Effects and Behavioral Responses

doi: https://doi.org/10.1257/aer.103.4.1289, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/w7vpu/files/osfstorage

Original Authors' Package: https://doi.org/10.3886/E112645V1

C.1.93 Reproduction Report

Title Original Study: Gone with the wind: The consequences of US drone strikes in Pakistan

doi: https://doi.org/10.1093/ej/ueac049, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/fm4ky/

Original Authors' Package: https://zenodo.org/record/5830531

C.1.94 Reproduction Report

Title Original Study: Good Reverberations? Teacher Influence in Music Composition since 1450

doi: https://doi.org/10.1086/718370, Journal of Political Economy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/27.htm

Link to Replicators' Package: https://www.dropbox.com/scl/fo/6hecmgjsq3mjo5ekkv8lm/h?rlkey=ftuoe4mf5f9jon0hiabb4brtn&dl=0

Link to Original Authors' Response: https://osf.io/79e2z/

Original Authors' Package: https://doi.org/10.1086/718370/suppl_file/20210405data.zip

C.1.95 Reproduction Report

Title Original Study: Has the Information Channel of Monetary Policy Disappeared? Revisiting the Empirical Evidence

doi: https://doi.org/10.1257/mac.20200068, American Economic Journal: Macroeconomics

Link to Full Report: https://osf.io/v4w9t/

Authors' Response: No response.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/169301/version/V1/view

C.1.96 Reproduction Report

Title Original Study: Hate Crimes and Gender Imbalances: Fears over Mate Competition and Violence against Refugees

doi: https://doi.org/10.1111/ajps.12595, American Journal of Political Science

Link to Full Report (and Initial Version of the Report): https://osf.io/5n3ds/

Link to Replicators' Package: https://osf.io/5n3ds/

Link to Original Authors' Response: https://osf.io/5n3ds/

Original Authors' Package: https://doi.org/10.7910/DVN/QXJDJ5

C.1.97 Reproduction Report

Title Original Study: Helping Children Catch Up: Early Life Shocks and the PROGRESA Experiment (Team 1)

doi: https://doi.org/10.1093/ej/uead067, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/sp2zh/

Original Authors' Package: https://zenodo.org/records/8206694

C.1.98 Reproduction Report

Title Original Study: Helping Children Catch Up: Early Life Shocks and the PROGRESA Experiment (Team 2)

doi: https://doi.org/10.1093/ej/uead067, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/sp2zh/

Original Authors' Package: https://zenodo.org/records/8206694

C.1.99 Reproduction Report

Title Original Study: Hidden Income and the Perceived Returns to Migration

doi: https://doi.org/10.1257/app.20210571, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/p9twk/

Original Authors' Package: https://doi.org/10.3886/E176081V1

C.1.100 Reproduction Report

Title Original Study: Historical Lynchings and the Contemporary Voting Behavior of Blacks

doi: https://doi.org/10.1257/app.20190549, American Economic Journal: Applied Economics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/32.htm

Link to Replicators' Package: https://osf.io/hv7wp/

Original Author's Response: No response to our emails.

Original Author's Package: https://www.openicpsr.org/openicpsr/project/136741/version/V1/view

C.1.101 Reproduction Report

Title Original Study: Hobo Economicus

doi: https://doi.org/10.1093/ej/ueab103, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/55.htm

Link to Replicators' Package: https://osf.io/s4bca/

Original Authors' Response: Declined to respond.

Original Authors' Package: https://zenodo.org/records/5719541

C.1.102 Reproduction Report

Title Original Study: Housing in Medicaid: Should It Really Change?

doi: https://doi.org/10.1257/pol.20200178, American Economic Journal: Economic Policy

Full Report: Computational reproducibility successful.

Original Authors' Package: https://doi.org/10.3886/E147421V1

C.1.103 Reproduction Report

Title Original Study: How Distributional Conflict over In-Kind Benefits Generates Support for Far-Right Parties

doi: https://doi.org/10.1086/720643, The Journal of Politics

Original Authors' Package: https://doi.org/10.7910/DVN/SYNP73

C.1.104 Reproduction Report

Title Original Study: How Do Beliefs about the Gender Wage Gap Affect the Demand for Public Policy?

doi: https://doi.org/10.1257/pol.20200559, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/12.htm

Link to Replicators' Package: https://osf.io/j2ubt/

Original Authors' Response: "I very much appreciate the effort of you and your team to replicate not just my paper, but many others too. It is quite impressive to see the scope of your project and I am curious about your future plans with this initiative.

I just read the Reproduction Report for my paper and I think it is great. In particular, Figures 2 and 3 are really cool. (They are actually new, and offer a really insightful way of looking at the data. I should have come up with them myself!)

Regarding your question, I don't think the Reproduction Report requires a formal response from my side. I fully agree with the authors' interpretation of the results, and just want to say thank you for their great work. Please go ahead and publish the report whenever it suits you."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/134041/version/V1/view

C.1.105 Reproduction Report

Title Original Study: How do Campaigns Shape Vote Choice? Multicountry Evidence from 62 Elections and 56 TV Debates

doi: https://doi.org/10.1093/qje/qjad002, Quarterly Journal of Economics

Link to Full Report and Authors' Response: https://osf.io/nj3fw/

Original Authors' Package: https://doi.org/10.7910/DVN/XMDFQO

C.1.106 Reproduction Report

Title Original Study: How Effective Are Monetary Incentives to Vote? Evidence from a Nationwide Policy?

doi: https://doi.org/10.1257/app.20200482, American Economic Journal: Applied Economics

Report's Abstract: Successful computational reproducibility. No re-analyses conducted.

Authors' Response: Not contacted.

Original Authors' Package: https://doi.org/10.1086/720458/suppl_file/20190733data.zip

C.1.107 Reproduction Report

Title Original Study: How Merchant Towns Shaped Parliaments: From the Norman Conquest of England to the Great Reform Act (Team 1)

doi: https://doi.org/10.1257/aer.20200885, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/nfvbg/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/166222/version/V1/view

C.1.108 Reproduction Report

Title Original Study: How Merchant Towns Shaped Parliaments: From the Norman Conquest of England to the Great Reform Act (Team 2)

doi: https://doi.org/10.1257/aer.20200885, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/nfvbg/

Original Authors' Package: https://doi.org/10.3886/E166222V1

C.1.109 Reproduction Report

Title Original Study: How Much Should We Trust the Dictator's GDP Growth Estimates?

doi: https://doi.org/10.1086/720458, Journal of Political Economy

Link to Full Report: https://osf.io/4sk52/

Link to Replicators' Package: https://osf.io/4sk52/

Original Author's Final Response: "As before, please extend my gratitude to the replicators for their thoughtful work and for taking into consideration my previous comments. I am reassured by the fact that they were able to replicate all the original results in the paper. I also find it reassuring that the results prove robust to additional robustness tests concerning alternative clustering structures for the standard errors or alternative data sources on political regimes (albeit with some loss of precision). The heterogeneous effects by subperiod are also quite intriguing and potentially reflect changes in the geopolitical incentives to overstate GDP growth in non-democracies. My paper is certainly not the final word on this topic and these results could be the first step towards new and exciting research."

Original Authors' Package: https://doi.org/10.1086/720458/suppl_file/20190733data.zip

C.1.110 Reproduction Report

Title Original Study: How the Party Commands the Gun: The Foreign-Domestic Threat Dilemma in China

doi: https://doi.org/10.1111/ajps.12739

Link to Full Report and Authors' Response: https://osf.io/2ce4q/

Original Authors' Package: https://doi.org/10.7910/DVN/R3XPEJ

C.1.111 Reproduction Report

Title Original Study: How Wage Announcements Affect Job Search - A Field Experiment

doi: https://doi.org/10.1257/mac.20200116, American Economic Journal: Macroeconomics

Link to Full Report and Authors' Response: https://osf.io/u6xz9/

Original Authors' Package: https://doi.org/10.3886/E152721V1

C.1.112 Reproduction Report

Title Original Study: Ideological Asymmetries and the Determinants of Politically Motivated Reasoning

doi: https://doi.org/10.1111/ajps.12624, American Journal of Political Science

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/79.htm

Link to Replicators' Package: https://osf.io/mh5sk/

Link to Original Authors' Response: "Thank you again for examining our paper so closely [...] we changed the codebook and appreciate this replication effort."

Original Authors' Package: https://doi.org/10.7910/DVN/CGHTPZ

C.1.113 Reproduction Report

Title Original Study: Immigration and Redistribution

doi: https://doi.org/10.1093/restud/rdac011, Review of Economic Studies

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/40.htm

Link to Replicators' Package: https://osf.io/ajm9g/

Original Authors' Response: "Dear Institute for Replication team,

Thank you very much for taking the time to replicate our paper. We appreciate the important work you do. We are happy to see that our results replicated well and that our robustness checks were confirmed.

With best wishes, Armando Miano and Stefanie Stantcheva"

Original Authors' Package: https://zenodo.org/records/5997521

C.1.114 Reproduction Report

Title Original Study: Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis

doi: https://doi.org/10.1086/707764, Journal of Political Economy

Link to Full Report and Original Authors' Response: https://osf.io/9b4ku/

:

Original Authors' Package: https://www.journals.uchicago.edu/doi/suppl/10.1086/707764/suppl_file/2016347data.zip

C.1.115 Reproduction Report

Title Original Study: Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion

doi: https://doi.org/10.1257/aer.20170765, American Economic Review

Link to Full Report and Original Authors' Response: https://osf.io/wdbc3/

:

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/113187/version/V1/view

C.1.116 Reproduction Report

Title Original Study: Improving Business Practices and the Boundary of the Entrepreneur: A Randomized Experiment Comparing Training, Consulting, Insourcing, and Outsourcing

doi: https://doi.org/10.1086/717044, Journal of Political Economy

Full Report: Computational reproducibility conducted during the World Bank-IMF games.

Authors' Response: Not contacted.

Original Authors' Package: https://www.journals.uchicago.edu/doi/full/10.1086/717044#supplementary-materials

C.1.117 Reproduction Report

Title Original Study: Improving Workplace Climate in Large Corporations: A Clustered Randomized Intervention

doi: https://doi.org/10.1093/qje/qjac034, Quarterly Journal of Economics

Link to Full Report and Authors' Response: https://osf.io/txqjw/

Original Authors' Package: https://doi.org/10.7910/DVN/A7GEVV

C.1.118 Reproduction Report

Title Original Study: Indecent Disclosures: Anticorruption Reforms and Political Selection

doi: https://doi.org/10.1111/ajps.12646, American Journal of Political Science

Link to Full Report: https://osf.io/gx4d6/

Link to Replicators' Package: https://osf.io/gx4d6/

Original Author's Response: "Many thanks to the replicator for taking the time to replicate and extend the paper. The placebo tests are very helpful in illuminating whether the identifying assumptions hold. I will make sure to run versions of these in future analyses."

Original Authors' Package: https://doi.org/10.7910/DVN/KDUMRM

C.1.119 Reproduction Report

Title Original Study: Inflammatory Political Campaigns and Racial Bias in Policing

doi: https://doi.org/10.1093/qje/qjac037, Quarterly Journal of Economics

Link to Full Report: https://osf.io/xadb6/

Link to Replicators' Package: https://osf.io/c7j58/

Original Authors' Response: https://osf.io/xadb6/

Original Authors' Package: https://doi.org/10.7910/DVN/A3B9HE

C.1.120 Reproduction Report

Title Original Study: Influence Motives in Social Signaling: Evidence from COVID-19 Vaccinations in Germany

doi: https://doi.org/10.1257/aeri.20220163, American Economic Review: Insights

Link to Full Report and Authors' Response: https://osf.io/bn65z/

Original Authors' Package: https://doi.org/10.3886/E177962V1

C.1.121 Reproduction Report

Title Original Study: Information Aggregation Under Ambiguity: Theory and Experimental Evidence

doi: https://doi.org/10.1093/restud/rdae009, Review of Economic Studies

Link to Full Report and Authors' Response: https://osf.io/pv7eh/

Original Authors' Package: https://doi.org/10.5281/zenodo.10033669

C.1.122 Reproduction Report

Title Original Study: Information and Spillovers from Targeting Policy in Peru's Anchoveta Fishery

doi: https://doi.org/10.1257/pol.20210812, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/5uz6d/

Original Authors' Package: https://doi.org/10.5281/zenodo.8006639

C.1.123 Reproduction Report

Title Original Study: Innovation and Institutional Ownership

doi: https://doi.org/10.1257/aer.103.1.277, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Response: No formal response.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/112591/version/V1/view

C.1.124 Reproduction Report

Title Original Study: Informed Enforcement: Lessons from Pollution Monitoring in China

doi: https://doi.org/10.1257/app.20210386, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/wb5x6/

Original Authors' Package: https://doi.org/10.3886/E174901V1

C.1.125 Reproduction Report

Title Original Study: Interaction, Stereotypes, and Performance: Evidence from South Africa

doi: https://doi.org/10.1257/aer.20181805, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files

Link to Replicators' Package: https://osf.io/w7vpu/files

Link to Original Authors' Response: Did not receive formal response.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/174501/version/V1/view

C.1.126 Reproduction Report

Title Original Study: Interregional Contact and the Formation of a Shared Identity

doi: https://doi.org/10.1257/pol.20210237, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/26jh8/

Original Authors' Package: https://doi.org/10.3886/E163261V1

C.1.127 Reproduction Report

Title Original Study: Interventions and Cognitive Spillovers

doi: https://doi.org/10.1093/restud/rdab087, Review of Economic Studies

Link to Full Report: https://www.econstor.eu/bitstream/10419/272845/1/I4R-DP043.pdf

Link to Replicators' Package: https://osf.io/kugbs/

Original Authors' Response: "We do not have any comments regarding the replication. We just want to briefly express our gratitude for the thorough and excellent work of the authors of the replication study."

Original Authors' Package: https://zenodo.org/records/5652808

C.1.128 Reproduction Report

Title Original Study: Intrahousehold Resource Allocation and Individual Poverty: Assessing Collective Model Predictions using Direct Evidence on Sharing

doi: https://doi.org/10.1093/ej/ueab085, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/fw7gp/

Original Authors' Package: https://doi.org/10.5281/zenodo.5525753

C.1.129 Reproduction Report

Title Original Study: Intrinsic Information Preferences and Skewness

doi: https://doi.org/10.1257/aer.20171474, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/5d4q9/

Original Authors' Package: https://doi.org/10.3886/E190641V1

C.1.130 Reproduction Report

Title Original Study: Jobs and Political Participation: Evidence from a Field Experiment in Ethiopia

doi: https://doi.org/10.1086/726929, Journal of Politics

Link to Full Report and Authors' Response: https://osf.io/cpmnt/

Original Authors' Package: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/QIH3D7

C.1.131 Reproduction Report

Title Original Study: Jumping the Gun: How Dictators Got Ahead of Their Subjects

doi: https://doi.org/10.1093/ej/ueac073, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/50.htm

Link to Replicators' Package: https://osf.io/4cx86/

Original Authors' Response: "we do not have any comments to the Reproduction Report, so I'm just sending you this email to applaud the initiative. You and the Institute for Replication is doing a great service to the profession."

Original Authors' Package: https://zenodo.org/records/7077694

C.1.132 Reproduction Report

Title Original Study: Juvenile Crime and Anticipated Punishment

doi: https://doi.org/10.1257/pol.20210530, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/pyjmc/

Original Authors' Package: https://doi.org/10.3886/E176501V1

C.1.133 Reproduction Report

Title Original Study: Kin Networks and Institutional development

doi: https://doi.org/10.1093/ej/ueac027, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/2d5hv/

Original Authors' Package: https://zenodo.org/records/6373189

C.1.134 Reproduction Report

Title Original Study: Learning-through-Survey in Inflation Expectations

doi: https://doi.org/10.1257/mac.20200387, American Economic Journal: Macroeconomics

Link to Full Report and Authors' Response: https://osf.io/s3k8h/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/166481/version/V1/view

C.1.135 Reproduction Report

Title Original Study: Liberty, Security, and Accountability: The Rise and Fall of Illiberal Democracies

doi: https://doi.org/10.1093/restud/rdab063, Review of Economic Studies

Link to Full Report and Authors' Response:: https://osf.io/8euy2/

Original Authors' Package: https://doi.org/10.5281/zenodo.7302150

C.1.136 Reproduction Report

Title Original Study: Liquidity Constraints in the U.S. Housing Market

doi: https://doi.org/10.1093/restud/rdab063, Review of Economic Studies

Report's Abstract: Successful computational reproducibility. No re-analyses conducted.

Original Authors' Response: Authors provided feedback and suggestions.

Original Authors' Package: http://doi.org/10.5281/zenodo.5112964

C.1.137 Reproduction Report

Title Original Study: Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo

doi: https://doi.org/10.1257/aer.20201159, American Economic Review

Link to Full Report: https://ideas.repec.org/p/zbw/i4rdps/191.html

Link to Replicators' Package: https://github.com/SossouAdjisse/LocalTaxReplicationProject.git

Link to Original Authors' Response: https://ideas.repec.org/p/zbw/i4rdps/192.html

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/147561/version/V1/view

C.1.138 Reproduction Report

Title Original Study: Major Reforms in Electricity Pricing: Evidence from a Quasi-Experiment

doi: https://doi.org/10.1093/ej/ueab076, Economic Journal

Link to Full Report: https://osf.io/bysa7/

Link to Replicators' Package: https://osf.io/bysa7/

Original Authors' Response: Original authors provided feedback. Multiple rounds of back and forth with replicators.

Original Authors' Package: https://zenodo.org/records/5423782

C.1.139 Reproduction Report

Title Original Study: Market Access and Quality Upgrading: Evidence from Four Field Experiments

doi: https://doi.org/10.1257/aer.20210122, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/72.htm

Link to Replicators' Package: https://journaldata.zbw.eu/dataset/bold-et-al-american-economic-review-2022

Original Authors' Response: "Thank you very much for sharing the report (and taking the time to replicate the study). We have no comments."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/158401/version/V1/view

C.1.140 Reproduction Report

Title Original Study: Market Power and Price Exposure: Learning from Changes in Renewable Energy Regulation

doi: https://doi.org/10.1257/pol.20210221, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/792jx/

Original Authors' Package: https://doi.org/10.3886/E171462V1

C.1.141 Reproduction Report

Title Original Study: Market-Based Monetary Policy Uncertainty (Team 1)

doi: https://doi.org/10.1093/ej/ueab086, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/77.htm

Link to Replicators' Package: https://osf.io/qx8aw/

Original Authors' Response: "Thank you, glad to see that this work found our results to be rock solid!

We won't write a response. Do let us know if you have any other questions about our work."

Original Authors' Package: https://zenodo.org/records/5566246

C.1.142 Reproduction Report

Title Original Study: Market-Based Monetary Policy Uncertainty (Team 2)

doi: https://doi.org/10.1093/ej/ueab086, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/76.htm

Link to Replicators' Package: https://github.com/YaolangZhong/Nottingham_Replication_Game/tree/main/replication_code

Original Authors' Response: "Thank you, glad to see that this work found our results to be rock solid!

We won't write a response. Do let us know if you have any other questions about our work."

Original Authors' Package: https://zenodo.org/records/5566246

C.1.143 Reproduction Report

Title Original Study: Maternal Investments in Children: The Role of Expected Effort and Returns

doi: https://doi.org/10.1093/ej/ueae059, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/6rpk2/

Original Authors' Package: https://doi.org/10.5281/zenodo.11649661

C.1.144 Reproduction Report

Title Original Study: Measuring Monetary Policy in the Euro Area Using SVARs with Residual Restrictions

doi: https://doi.org/10.1257/mac.20210035, American Economic Journal: Macroeconomics

Link to Full Report and Authors' Response: https://osf.io/39vkb/

Original Authors' Package: https://doi.org/10.3886/E169401V1

C.1.145 Reproduction Report

Title Original Study: Measuring the Welfare Effects of Shame and Pride

doi: https://doi.org/10.1257/aer.20190433, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/64.htm

Link to Replicators' Package: https://github.com/tilmanfries/welfare-shame-pride-replication-report

Original Authors' Final Response: "Thanks again for all your hard work on this."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/145141/version/V1/view

C.1.146 Reproduction Report

Title Original Study: Mental Health Costs of Lockdowns: Evidence from Age-Specific Curfews in Turkey

doi: https://doi.org/10.1257/app.20200811, American Economic Journal: Applied Economics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/16.htm

Link to Replicators' Package: https://osf.io/25u7b/

Link to Original Authors' Response: https://econpapers.repec.org/paper/zbwi4rdps/17.htm

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/131981/version/V1/view

C.1.147 Reproduction Report

Title Original Study: Mortality, Temperature, and Public Health Provision: Evidence from Mexico

doi: https://doi.org/10.1257/pol.20180594, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/90.htm

Link to Replicators' Package: https://osf.io/q52e4/

Original Authors' Response: Cohen: "We thank The Institute for Replication. Next time, I will make sure I do not forget Feb. 29th in the code!"

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/125201/version/V1/view

C.1.148 Reproduction Report

Title Original Study: Motivated Beliefs and Anticipation of Uncertainty Resolution

doi: https://doi.org/10.1257/aeri.20200829, American Economic Review: Insights

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/65.htm

Link to Replicators' Package: https://osf.io/evt3a/

Original Authors' Response: "Thanks for sharing the report. I think it's a great initiative and feel free to publish this report on your webpage. I will not be able to provide an "answer"."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/139262/version/V1/view

C.1.149 Reproduction Report

Title Original Study: Motivated Errors

doi: https://doi.org/10.1257/aer.20191849, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/x29kf/

Original Authors' Package: https://doi.org/10.3886/E195862V1

C.1.150 Reproduction Report

Title Original Study: Multinationals' Sales and Profit Shifting in Tax Havens

doi: https://doi.org/10.1257/pol.20200203, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/37.htm

Link to Replicators' Package: https://osf.io/3sbmr/

Original Authors' Response: "Thanks for your email and for replicating our exercise. Your work is useful. We recognize that the results remain consistent even when considering different interpretations of the haven concept and a smaller sample of observations.

We are also pleased to hear that the replication file we shared with the AEJ: Policy has proven helpful."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/148301/version/V1/view

C.1.151 Reproduction Report

Title Original Study: Multiracial identity and political preferences

doi: https://doi.org/10.1086/714760, Journal of Politics

Link to Full Report: https://osf.io/ecymu/

Link to Replicators' Package: https://github.com/taylorjwright/r_and_p_versioning

Original Authors' Response: Back and forth between authors and replicators, but did not obtain a final response as of November 2025.

Original Authors' Package: https://doi.org/10.7910/DVN/BLVJJH

C.1.152 Reproduction Report

Title Original Study: News Shocks Under Financial Frictions

doi: https://doi.org/10.1257/mac.20170066, American Economic Journal: Macroeconomics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/51.htm

Link to Replicators' Package: https://github.com/gionikola/replication-game-ucsd

Original Authors' Final Response: "Thank you for the update and considering our work for replication."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/130141/version/V1/view

C.1.153 Reproduction Report

Title Original Study: No Man Left Behind: Effects of Emigration Prospects on Educational and Labour Outcomes of Non-migrants

doi: https://doi.org/10.11ll/ecqj.l2306, Economic Journal

Link to Full Report: https://osf.io/9d6au/

Original Authors' Response: No response.

Original Authors' Package: Supplementary Materials.

C.1.154 Reproduction Report

Title Original Study: Non-Linearities, State-Dependent Prices and the Transmission Mechanism of Monetary Policy

doi: https://doi.org/10.1093/ej/ueab049, Economic Journal

Link to Full Report: https://osf.io/kbwap/

Link to Replicators' Package: https://osf.io/kbwap/

Original Authors' Response: No response.

Original Authors' Package: https://academic.oup.com/ej/article/132/641/37/6302374#supplementary-data

C.1.155 Reproduction Report

Title Original Study: Not All Elections Are Created Equal: Election Quality and Civil Conflict

doi: https://doi.org/10.1086/714778, Journal of Politics

Link to Full Report: https://osf.io/unhkr/

Link to Replicators' Package: https://drive.google.com/drive/folders/1Vlwfr3_Q7c56XFPsQuKUNSnEU_lMFUQz?usp=sharing

Link to Original Authors' Response: https://osf.io/unhkr/

Original Authors' Package: https://doi.org/10.7910/DVN/8B31FG

C.1.156 Reproduction Report

Title Original Study: Optimal Deposit Insurance

doi: https://doi.org/10.1086/723633, Journal of Political Economy

Link to Full Report and Authors' Response: https://osf.io/nye3s/

Original Authors' Package: https://doi.org/10.7910/DVN/Q3I619

C.1.157 Reproduction Report

Title Original Study: Pandemic Schooling Mode and Student Test Scores: Evidence from US School Districts

doi: https://doi.org/10.1257/aeri.20210748, American Economic Review: Insights

Link to Full Report and Authors' Response: https://osf.io/gkjqe/

Original Authors' Package: https://doi.org/10.3886/E168843V1

C.1.158 Reproduction Report

Title Original Study: Parties as Disciplinarians: Charisma and Commitment Problems in Programmatic Campaigning

doi: https://doi.org/10.1111/ajps.12638, American Journal of Political Science

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/54.htm

Link to Replicators' Package: https://osf.io/93gfx/?view_only=7063353244d646ffaf7bfd53013e3143

Original Authors' Response: "Thanks for your note and for all the work of Kelly, Odermatt, and Metson in replicating our paper. [...] Our read of the Reproduction Reports that the findings in our paper hold in the Kelly et al replication. [...] Our sense is that the discrepancies between the replication and original paper are sufficiently small, and the task of comparing the replication R code to the original Stata code is likely to be sufficiently demanding of time, that the opportunity cost of a thorough response is high. So, I think we'll forgo the opportunity to draft a response, and just let the replication stand without reply.

We'll leave it to any sufficiently interested parties with expertise in both Stata and R to iron out the discrepancies between the replication and original paper."

Original Authors' Package: https://doi.org/10.7910/DVN/AWSQTW

C.1.159 Reproduction Report

Title Original Study: Partisanship and Fiscal Policy in Economic Unions: Evidence from US States

doi: https://doi.org/10.1257/aer.20210147, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/177001/version/V1/view

C.1.160 Reproduction Report

Title Original Study: Patience, Risk-Taking, and Human Capital Investment Across Countries

doi: https://doi.org/10.1093/ej/ueab105, Economical Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/48.htm

Link to Replicators' Package: https://osf.io/kgt8z/

Link to Original Authors' Response: "We would like to thank the replicators and compliment them for their thoughtful replication and extension of our paper. We are particularly impressed by the extension to the TIMSS data, which is actually great support for the underlying idea. We do not see a reason to formulate a formal response for your website.

Thank you all for your valuable work!"

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/153101/version/V2/view

C.1.161 Reproduction Report

Title Original Study: Peer Effects in Academic Research: Senders and Receives

doi: https://doi.org/10.1093/ej/ueac031, Economical Journal

Link to Full Report: https://osf.io/czkgw/

Link to Replicators' Package: https://osf.io/czkgw/

Link to Original Authors' Response: The authors responded to the replicators' questions. Bosquet then responded to the final report: "I would simply thank the team of replicators and I am happy to see that the tested results are robust to the tested alternatives. As written in my previous email, I think those kinds of efforts are very useful for the community and the credibility of published results so thanks as well for that."

Original Authors' Package: https://zenodo.org/records/6457037

C.1.162 Reproduction Report

Title Original Study: Persuasion through Slanted Language: Evidence from the Media Coverage of Immigration

doi: https://doi.org/10.1257/aer.20211537, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/182482/version/V1/view

C.1.163 Reproduction Report

Title Original Study: Physical Disability and Labor Market Discrimination: Evidence from a Video Résumé Field Experiment (Team 1)

doi: https://doi.org/10.1257/app.20210633, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/qxnb5/

Original Authors' Package: https://doi.org/10.3886/E175061V1

C.1.164 Reproduction Report

Title Original Study: Physical Disability and Labor Market Discrimination: Evidence from a Video Résumé Field Experiment (Team 2)

doi: https://doi.org/10.1257/app.20210633, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/qxnb5/

Original Authors' Package: https://doi.org/10.3886/E175061V1

C.1.165 Reproduction Report

Title Original Study: Playing Politics with Environmental Protection: The Political Economy of Designating Protected Areas

doi: https://doi.org/10.1086/718978, Journal of Politics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/73.htm

Link to Replicators' Package: https://osf.io/t76jd/

Original Authors' Response: "We are grateful to Laura Villalobos, Jill Caviglia-Harris, Tharaka Jayalath, and the team at the Institute for Replication for generously replicating our work. We encourage readers to follow their alternative software routines for faster estimations."

Original Authors' Package: https://doi.org/10.7910/DVN/N6LIMH

C.1.166 Reproduction Report

Title Original Study: Policy Deliberation and Voter Persuasion: Experimental Evidence from an Election in the Philippines

doi: https://doi.org/10.1111/ajps.12566, American Journal of Political Science

Link to Full Report: https://osf.io/y8ubt/

Link to Replicators' Package: https://osf.io/y8ubt/

Link to Original Authors' Response: https://osf.io/y8ubt/

Original Authors' Package: https://doi.org/10.7910/DVN/S3HACJ

C.1.167 Reproduction Report

Title Original Study: Political Misinformation and Factual Corrections on the Facebook News Feed: Experimental Evidence

doi: https://doi.org/10.1086/719271, The Journal of Politics

Link to Full Report and Authors' Response: https://osf.io/fzy8x/

Original Authors' Package: https://doi.org/10.7910/DVN/5E3D0F

C.1.168 Reproduction Report

Title Original Study: Political Responsiveness to Conflict Victims: Evidence from a Countrywide Audit Experiment in Colombia

doi: https://doi.org/10.1017/S0003055422001277, American Political Science Review

Link to Full Report and Authors' Response: To be posted soon on I4R's Webpage.

Original Authors' Package: https://doi.org/10.7910/DVN/EORLRJ

C.1.169 Reproduction Report

Title Original Study: Political Turnover, Bureaucratic Turnover, and the Quality of Public Services

doi: https://doi.org/10.1257/aer.20171867, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/39.htm

Link to Replicators' Package: https://osf.io/q43vz/

Link to Original Authors' Response: https://osf.io/kv4pj/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/150323/version/V1/view

C.1.170 Reproduction Report

Title Original Study: Politicians' Private Sector Jobs and Parliamentary Behavior

doi: https://doi.org/10.1111/ajps.12721, American Journal of Political Science

Link to Full Report: https://osf.io/pnk24/

Link to Original Authors' Response: https://osf.io/pnk24/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/150323/version/V1/view

C.1.171 Reproduction Report

Title Original Study: Populist Leaders and the Economy

doi: https://doi.org/10.1257/aer.20202045, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/6uef8/

Original Authors' Package: https://doi.org/10.3886/E184362V1

C.1.172 Reproduction Report

Title Original Study: Pre-Colonial Warfare and Long-Run Development in India

doi: https://doi.org/10.1093/ej/ueab089, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/35.htm

Link to Replicators' Package: https://osf.io/af6m2/

Link to Original Authors' Response: https://osf.io/af6m2/

Original Authors' Package: https://zenodo.org/records/5583263

C.1.173 Reproduction Report

Title Original Study: Predictable Effects of Visual Salience in Experimental Decisions and Games

doi: https://doi.org/10.1093/qje/qjac025, Quarterly Journal of Economics

Link to Full Report and Authors' Response: https://osf.io/rb7zv/

Original Authors' Package: https://doi.org/10.7910/DVN/9LCYKG.

C.1.174 Reproduction Report

Title Original Study: Preference Discovery in University Admissions: The Case for Dynamic Multioffer Mechanisms

doi: https://doi.org/10.1086/718983, Journal of Political Economy

Link to Full Report and Authors' Response: https://osf.io/zwt7n/

Original Authors' Package: https://www.journals.uchicago.edu/doi/suppl/10.1086/718983/suppl_file/20190479data.zip

C.1.175 Reproduction Report

Title Original Study: Price Discrimination by Negotiation: a Field Experiment in Retail Electricity

doi: https://doi.org/10.1093/qje/qjac021, Quarterly Journal of Economics

Link to Full Report and Authors' Response: https://osf.io/47hg6/

Original Authors' Package: https://doi.org/10.7910/DVN/KRHAWJ

C.1.176 Reproduction Report

Title Original Study: Propagation and Insurance in Village Networks

doi: https://doi.org/10.1257/aer.20220892, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/fq5vz/

Original Authors' Package: https://doi.org/10.3886/E193220V1

C.1.177 Reproduction Report

Title Original Study: Public Pension Reforms and Retirement Decisions: Narrative Evidence and Aggregate Implications

doi: https://doi.org/10.1257/pol.20200881, American Economic Journal: Economic Policy

Link to Full Report and Original Authors' Response: https://osf.io/rqc9p/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/173182/version/V1/view

C.1.178 Reproduction Report

Title Original Study: Public Infrastructure and Economic Development: Evidence from Postal Systems

doi: https://doi.org/10.1111/ajps.12594, American Journal of Political Science

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/92.htm

Link to Replicators' Package: https://osf.io/j3ydr/?view_only=ad14a07cb3a741ca9bbfab391ad7c6fe

Original Authors' Response: "Thanks so much for reproducing the findings in our paper and exploring extensions of our results. We also appreciate your sharing the report with us. [...] I [Rogowski] confirm that we are comfortable letting your report stand and that we will not write a response to it. "

Original Authors' Package: https://doi.org/10.7910/DVN/33K3EF

C.1.179 Reproduction Report

Title Original Study: Publishing While Female: are Women Held to Higher Standards? Evidence from Peer Review

doi: https://doi.org/10.1093/ej/ueac032, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/zb73y/

Original Authors' Package: https://doi.org/10.5281/zenodo.6491871

C.1.180 Reproduction Report

Title Original Study: Re-Assessing Elite-Public Gaps in Political Behavior

doi: https://doi.org/10.1111/ajps.12583, American Journal of Political Science

Link to Full Report: https://www.econstor.eu/bitstream/10419/266385/1/I4R-DP010.pdf

Link to Replicators' Package: https://osf.io/93urk/

Original Authors' Response: "Thank you for your email and for the invitation. [...] please send my appreciation to the authors for their interest in the manuscript; I find their analysis very interesting."

Original Authors' Package: https://doi.org/10.7910/DVN/LHOTOK

C.1.181 Reproduction Report

Title Original Study: Rebel on the Canal: Disrupted Trade Access and Social Conflict in China, 1650–1911

doi: https://doi.org/10.1257/aer.20201283, American Economic Review

Link to Full Report: https://osf.io/dhn6e/

Link to Replicators' Package: https://osf.io/dhn6e/

Link to Original Authors' Response: https://osf.io/dhn6e/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/157781/version/V1/view

C.1.182 Reproduction Report

Title Original Study: Recessions, Mortality, and Migration Bias: Evidence from the Lancashire Cotton Famine

doi: https://doi.org/10.1257/app.20190131, American Economic Journal: Applied Economics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/25.htm

Link to Replicators' Package: https://www.openicpsr.org/openicpsr/project/192272/version/V1/view

Original Authors' Response: "Thanks for the interest in our work. We've had a chance to review the report and it looks like everything replicated. Since there are no outstanding queries, we are happy to sign off on this."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/128521/version/V1/view

C.1.183 Reproduction Report

Title Original Study: Reducing Crime and Violence: Experimental Evidence from Cognitive Behavioral Therapy in Liberia

doi: https://doi.org/10.1257/aer.20150503, American Economic Review

Full Report: Computational reproducibility conducted during the World Bank-IMF games.

Original Authors' Response: Not contacted.

Original Authors' Package: https://doi.org/10.3886/E113056V1

C.1.184 Reproduction Report

Title Original Study: Religious Festivals and Economic Development: Evidence from the Timing of Mexican Saint Day Festivals

doi: https://doi.org/10.1257/aer.20211094, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/161801/version/V2/view

C.1.185 Reproduction Report

Title Original Study: Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India

doi: https://doi.org/10.1257/aer.20201112, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/24.htm

Link to Replicators' Package: https://osf.io/r5jfe/

Final Original Authors' Response: "Thanks. the revision looks good. I actually don't think we need to have a formal response any more. [...] Thus, I don't think there is anything substantive for us to include in a discussion paper/response. That reflects the fact that the Reproduction Reports fair and there is nothing major to respond to, so it's good news, from both the perspective of the integrity of our original paper and the professionalism of the replication."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/149882/version/V1/view

C.1.186 Reproduction Report

Title Original Study: Reshaping Global Trade: The Immediate and Long-Run Effects of Bank Failures

doi: https://doi.org/10.1093/qje/qjac016, Quarterly Journal of Economics

Link to Full Report and Authors' Response: https://osf.io/yf2mt/

Original Authors' Package: https://doi.org/10.7910/DVN/EIEQIS

C.1.187 Reproduction Report

Title Original Study: Resisting Social Pressure in the Household Using Mobile Money: Experimental Evidence on Microenterprise Investment in Uganda

doi: https://doi.org/10.1257/aer.20220717, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/j5rge/

Original Authors' Package: https://doi.org/10.3886/E194886V2

C.1.188 Reproduction Report

Title Original Study: Retail Pharmacies and Drug Diversion During the Opioid Epidemic

doi: https://doi.org/10.1257/aer.20210357, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Response: No formal response.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/170881/version/V1/view

C.1.189 Reproduction Report

Title Original Study: Run-off Elections in the Laboratory

doi: https://doi.org/10.1093/ej/ueab051, Economic Journal

Link to Full Report: https://osf.io/a8cev/

Link to Replicators' Package: https://github.com/carinahausladen/runoff-elections

Original Authors' Response: The authors provided feedback which was taken into account.

Original Authors' Package: https://doi.org/10.1093/ej/ueab051

C.1.190 Reproduction Report

Title Original Study: Safe Assets

doi: https://doi.org/10.1093/ej/ueac017, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/p3yjd/

Original Authors' Package: https://zenodo.org/record/5979683

C.1.191 Reproduction Report

Title Original Study: School Spending and Student Outcomes: Evidence from Revenue Limit Elections in Wisconsin

doi: https://doi.org/10.1257/pol.20200226, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/88.htm

Link to Replicators' Package: https://osf.io/m2w4x/

Original Author's Response: "Thank you for sharing the Reproduction Report. Please pass on my thanks to the replicators for their important work. First and foremost, I'm glad to see that the results in the paper are reproducible without any issues arising. The report explores two additional sources of heterogeneity. I have no additional comments on these. I do briefly want to clarify the last sentence in the report's abstract, which reads "Lastly, we find that the estimated effects on student outcomes rely heavily on recent observations." While I am not entirely sure what the replicators are referring to, my guess is that they refer to Table 2 in the report. In this table, they discuss that they are unable to study heterogeneity in the impacts of passing a referendum on test scores and postsecondary enrollment from 1996-2005, because data on these outcomes are unavailable prior to 2005. The availability of each dataset was discussed in the published version of the paper (see, for example, Table 1). Perhaps a more accurate statement would be to explain that the replicators couldn't explore the impact of passing a referendum on these specific outcomes in the early period due to data constraints—and that this was acknowledged in the published version."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/125821/version/V1/view

C.1.192 Reproduction Report

Title Original Study: Se Habla Español: Spanish-Language Appeals and Candidate Evaluations in the United States

doi: https://doi.org/10.1017/S0003055423000084, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/xfp4k/

Original Authors' Package: https://doi.org/10.7910/DVN/GT15SC

C.1.193 Reproduction Report

Title Original Study: Self-Persuasion: Evidence from Field Experiments at International Debating Competitions

doi: https://doi.org/10.1257/aer.20200372, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/vh5wb/

Original Authors' Package: https://doi.org/10.3886/E148242V1

C.1.194 Reproduction Report

Title Original Study: Social Class and (Un)Ethical Behaviour: Causal and Correlational Evidence

doi: https://doi.org/10.1093/ej/ueac022, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/29.htm

Link to Replicators' Package: https://github.com/ha0ye/replication-gsottbauer-2022

Original Authors' Response: Declined to respond.

Original Authors' Package: https://zenodo.org/records/6226207

C.1.195 Reproduction Report

Title Original Study: Solar Eclipses and the Origins of Critical Thinking and Complexity

doi: https://doi.org/10.1093/ej/uead117, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/cu4t5/

Original Authors' Package: https://zenodo.org/records/10408724

C.1.196 Reproduction Report

Title Original Study: Sorting or Steering: The Effects of Housing Discrimination on Neighborhood Choice

doi: https://doi.org/10.1086/720140, Journal of Political Economy

Link to Full Report: https://osf.io/vwgxd/

Link to Replicators' Package: https://github.com/mattdwebb/HUDreplication

Original Authors' Response: Authors mentioned that they are currently writing a response.

Original Authors' Package: https://doi.org/10.1086/720140/suppl_file/20191181data.zip

C.1.197 Reproduction Report

Title Original Study: Spatial Polarisation

Link to Full Report and Authors' Response: https://osf.io/w9nbm/

Original Authors' Package: https://zenodo.org/record/6726014

C.1.198 Reproduction Report

Title Original Study: Spillover Effects of Intellectual Property Protection in the Interwar Aircraft Industry

doi: https://doi.org/10.1093/ej/ueab091, Economic Journal

Link to Full Report: https://osf.io/t4avf/

Link to Replicators' Package: https://osf.io/t4avf/

Link to Original Authors' Response: https://osf.io/t4avf/

Original Authors' Package: https://zenodo.org/records/5627298

C.1.199 Reproduction Report

Title Original Study: State Action to Prevent Violence against Women: The Effect of Women's Police Stations on Men's Attitudes toward Gender-Based Violence

doi: https://doi.org/10.1086/714931, Journal of Politics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/67.htm

Link to Replicators' Package: https://osf.io/yjwr8/

Link to Original Authors' Response: Responded to our emails but no formal response as of February 2024.

Original Authors' Package: https://doi.org/10.7910/DVN/D2WL5I

C.1.200 Reproduction Report

Title Original Study: State Violence, Party Formation, and Electoral Accountability: The Political Legacy of the Marikana Massacre

doi: https://doi.org/10.1017/S0003055423000448, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/wdk2s/

Original Authors' Package: https://doi.org/10.7910/DVN/SAA71U

C.1.201 Reproduction Report

Title Original Study: Sticky Expectations and Consumption Dynamics

doi: https://doi.org/10.1257/mac.20180286, American Economic Journal: Macroeconomics

Full Report: Computational reproducibility conducted during the World Bank-IMF games.

Authors' Response: Not contacted.

Original Authors' Package: https://doi.org/10.3886/e231543v1

C.1.202 Reproduction Report

Title Original Study: Strategic Complexity and the Value of Thinking

doi: https://doi.org/10.1093/ej/ueac070, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/24v36/

Original Authors' Package: https://doi.org/10.5281/zenodo.7019108

C.1.203 Reproduction Report

Title Original Study: Student Performance, Peer Effects, and Friend Networks: Evidence from a Randomized Peer Intervention

doi: https://doi.org/10.1257/pol.20200563, American Economic Journal: Economic Policy

Link to Full Report: https://osf.io/9hx3b/

Original Authors' Response: The authors provided feedback which was taken into account.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/149262/version/V2/view

C.1.204 Reproduction Report

Title Original Study: Subways and Urban Air Pollution

doi: https://doi.org/10.1257/app.20180168, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/e8x63/

Original Authors' Package: https://doi.org/10.3886/E126401V1

C.1.205 Reproduction Report

Title Original Study: Talking Shops: The Effects of Caucus Discussion on Policy Coalitions

doi: https://doi.org/10.1111/ajps.12636, American Journal of Political Science

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/69.htm

Link to Replicators' Package: https://osf.io/tmfyj/

Link to Original Authors' Response: "One purpose of replication, among others, is to evaluate whether published results are sensitive to modeling decisions. Do alternative, reasonable approaches generate the same, or different, results? Did the author's approach provide an outlier estimate that is indicative of p-hacking or, to be kinder about it, sensitivity of results to modeling decisions? That seems incredibly useful. That purpose is not advanced, in my view, by testing 'incorrect' methods or models. We do not learn about the robustness of results from testing alternative approaches that introduce bias, or by estimating different estimands that are a combination of treatment effects and selection bias. While it doesn't seem to matter too much in this case — selection bias appears relatively small, and in the same direction as treatment effects — I think this issue matters for the exercise in general for several reasons. First, do the analyses justify the inferences being made? In my view, changing the estimand or estimating biased models cannot justify saying anything about the robustness of the original results. Second, what would have happened if the new results did not match the original? Are we willing to claim published results are not robust when applying estimators with known flaws generates different results? And third, shouldn't we just generally aim to use 'correct' estimators for a given situation? While IPW is not perfect, ignoring differential treatment probabilities is a conscious decision to ignore selection bias. Why would we want to run that model if our goal is inference about treatment effects? I appreciate the work everyone is doing on this enterprise. Hopefully these comments, whether correct or not, help advance the goal of publishing robust, valid empirical research."

Original Authors' Package: https://doi.org/10.7910/DVN/S3M5AX

C.1.206 Reproduction Report

Title Original Study: Targeting High Ability Entrepreneurs Using Community Information: Mechanism Design in the Field

doi: https://doi.org/10.1257/aer.20200751, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/49.htm

Link to Replicators' Package: https://doi.org/10.7910/DVN/DI7RR9

Link to Original Authors' Response: "We are very grateful to Isabella Masetto, Diego Ubfal, and to the team at I4R for their excellent work. We verified the coding error and we agree that it did not meaningfully alter the conclusion of our paper that community information is informative over and above the predictive power of observable characteristics. We will post a link to this correction

on our websites and will consult the editors of the AER as to whether this error rises to the level of requiring a formal correction."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/151841/version/V1/view

C.1.207 Reproduction Report

Title Original Study: Taste-Based Gender Favouritism in High-Stake Decisions: Evidence from The Price is Right

doi: https://doi.org/10.1093/ej/uead087, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/5sv6m/

Original Authors' Package: https://doi.org/10.5281/zenodo.8372384

C.1.208 Reproduction Report

Title Original Study: Taxing Billionaires: Estate Taxes and the Geographical Location of the Ultra-Wealthy

doi: https://doi.org/10.1257/pol.20200685, American Economic Journal: Economic Policy

Original Authors' Package: https://doi.org/10.3886/E157481V1

C.1.209 Reproduction Report

Title Original Study: Teaching Norms: Direct Evidence of Parental Transmission

doi: https://doi.org/10.1093/ej/ueac074, Economic Journal

Link to Full Report: https://osf.io/qnbfa/

Link to Replicators' Package: https://zenodo.org/records/8114738

Original Authors' Response: The replicators took into account the authors' feedback. They wrote at the end of the back and forth: "We thank you and the replication team for the replication and the successive interactions. We created an OSF project including the data replication package enabling the reproduction of the analysis presented in our article. The package comprises a source file (in Stata format and in TXT) and a Stata do-file that allows the reconstruction of the master file used in the replication package submitted to the Economic Journal."

Original Authors' Package: https://zenodo.org/records/7045559

C.1.210 Reproduction Report

Title Original Study: Technological Change and the Consequences of Job Loss

doi: https://doi.org/10.1257/aer.20210182, American Economic Review

Link to Full Report: https://osf.io/qws2p/

Link to Replicators' Package: https://osf.io/qws2p/

Original Authors' Response: Did not get a response as of November 2025.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/181166/version/V1/view

C.1.211 Reproduction Report

Title Original Study: Technology Gaps, Trade, and Income

doi: https://doi.org/10.1257/aer.20201940, American Economic Review

Link to Full Report and Author's Response: https://osf.io/czyx4/

Original Authors' Package: https://doi.org/10.3886/E182522V1

C.1.212 Reproduction Report

Title Original Study: Temperature, Worker Productivity, and Adaptation: Evidence from Survey Data Production

doi: https://doi.org/10.1257/app.20200547, American Economic Journal: Applied Economics

Link to Full Report and Authors' Response: https://osf.io/pg3m7/

Original Authors' Package: https://doi.org/10.3886/E141721V1

C.1.213 Reproduction Report

Title Original Study: Temporary Disenfranchisement: Negative Side Effects of Lowering the Voting Age

doi: https://doi.org/10.1017/S000305542200034X, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/hw7tv/

Original Authors' Package: https://doi.org/10.7910/DVN/T5LYWS

C.1.214 Reproduction Report

Title Original Study: The Ant and the Grasshopper: Seasonality and the Invention of Agriculture

doi: https://doi.org/10.1093/qje/qjae012, Quarterly Journal of Economics

Link to Full Report and Original Authors' Response: https://osf.io/krm68/

Original Authors' Package: See supplementary materials.

C.1.215 Reproduction Report

Title Original Study: The Common-Probability Auction Puzzle

doi: https://doi.org/10.1257/aer.20191927, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/74.htm

Link to Replicators' Package: https://osf.io/7bq4s/

Original Authors' Response: "Thank you for putting the effort in replicating our study! Your results are also quite interesting to us as we haven't thought of all the robustness checks you've made. At this point, we do not have any major comments to make and refrain from submitting a response."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/184041/version/V1/view

C.1.216 Reproduction Report

Title Original Study: The Curious Case of Theresa May and the Public That Did Not Rally: Gendered Reactions to Terrorist Attacks Can Cause Slumps Not Bumps

doi: https://doi.org/10.1017/S0003055421000861, American Political Science Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/41.htm

Link to Replicators' Package: https://doi.org/10.5683/SP3/6SYCML

Link to Original Authors' Response: https://econpapers.repec.org/paper/zbwi4rdps/44.htm

Original Authors' Package: https://doi.org/10.7910/DVN/VHNPUO

C.1.217 Reproduction Report

Title Original Study: The Curse of Good Intentions: Why Anticorruption Messaging Can Encourage Bribery

doi: https://doi.org/10.1017/S0003055421001398, American Political Science Review

Link to Full Report and Authors' Response: https://osf.io/3t4j7/

Original Authors' Package: https://doi.org/10.7910/DVN/JV8GBI

C.1.218 Reproduction Report

Title Original Study: The Dynamics and Spillovers of Management Interventions: Evidence from the Training within Industry Program

doi: https://doi.org/10.1086/719277, Journal of Political Economy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/66.htm

Link to Replicators' Package: https://github.com/cwestheide/i4r_dp66_code

Original Authors' Final Response: "Thanks a lot for sharing the updated report with us. We don't have anything to add at this point."

Original Authors' Package: https://doi.org/10.1086/719277/suppl_file/20200781data.zip

C.1.219 Reproduction Report

Title Original Study: The Ecological Impact of Transportation Infrastructure

doi: https://doi.org/10.1093/ej/ueaa013, Economic Journal

Full Report: Successful computational reproducibility conducted during the World Bank-IMF games.

Authors' Response: Not contacted.

Original Authors' Package: https://academic.oup.com/ej/article/130/629/1173/5798996#supplementary-data

C.1.220 Reproduction Report

Title Original Study: The Economic Effects of Long-Term Climate Change: Evidence from the Little Ice Age

doi: https://doi.org/10.1086/720393, Journal of Political Economy

Link to Full Report: https://osf.io/tmn2j/

Link to Replicators' Package: https://osf.io/tmn2j/

Link to Original Authors' Response: https://osf.io/tmn2j/

Original Authors' Package: https://doi.org/10.1086/720393/suppl_file/2015548data.zip

C.1.221 Reproduction Report

Title Original Study: The Effect of Early Childhood Education on Adult Criminality: Evidence from the 1960s through 1990s

doi: https://doi.org/10.1257/pol.20200660, American Economic Journal: Economic Policy

Original Authors' Package: https://doi.org/10.3886/E140241V1

C.1.222 Reproduction Report

Title Original Study: The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets

doi: https://doi.org/10.1257/aer.103.2.690, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Response: No formal response.

Original Authors' Package: https://doi.org/10.3886/E116114V1

C.1.223 Reproduction Report

Title Original Study: The Effects of Banking Competition on Growth and Financial Stability: Evidence from the National Banking Era

doi: https://doi.org/10.1086/717453, Journal of Political Economy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/81.htm

Link to Replicators' Package: https://doi.org/10.7910/DVN/BB864R

Original Authors' Final Response: "We thank the replication team (Andrea Calef, Sya In Chzhen, Marco Mandas, and Fabio Motoki) for the detailed Reproduction Report. We are glad to hear that the replicating team affirms the robustness of the paper's findings. We are also glad that the replicators were able to successfully replicate all tables and figures. We thank the replicating team for identifying various smaller issues regarding the code structure which fortunately did not affect our original findings. We believe that the report as such does not require us to respond in any further detail. We highly appreciate the effort of both the replicating team and the I4R."

Original Authors' Package: https://doi.org/10.1086/717453/suppl

C.1.224 Reproduction Report

Title Original Study: The Effects of Import Competition on Unionization

doi: https://doi.org/10.1257/pol.20200709, Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/c7eq2/

Original Authors' Package: https://doi.org/10.3886/E176801V1

C.1.225 Reproduction Report

Title Original Study: The Effects of Racial Diversity in Citizen Decision-Making Bodies

doi: https://doi.org/10.1086/726946, The Journal of Politics

Original Authors' Package: https://doi.org/10.7910/DVN/FQJ8WF

Link to Full Report and Authors' Response: https://osf.io/xqgru/

C.1.226 Reproduction Report

Title Original Study: The Evolution of Access to Public Accommodations in the United States

doi: https://doi.org/10.1093/qje/qjac035, Quarterly Journal of Economics

Link to Full Report and Authors' Response: https://osf.io/f9sdc

Original Authors' Package: https://doi.org/10.7910/DVN/NXFB5R

C.1.227 Reproduction Report

Title Original Study: The Gender Gap in Self-Promotion (Team 1)

doi: https://doi.org/10.1093/qje/qjac003, Quarterly Journal of Economics

Link to Full Report: https://osf.io/xph7u/

Original Authors' Response: Authors did not respond.

Original Authors' Package: https://doi.org/10.7910/DVN/YSWKHY

C.1.228 Reproduction Report

Title Original Study: The Gender Gap in Self-Promotion (Team 2)

doi: https://doi.org/10.1093/qje/qjac003, Quarterly Journal of Economics

Link to Full Report: https://osf.io/xph7u/

Original Authors' Response: Authors did not respond.

Original Authors' Package: https://doi.org/10.7910/DVN/YSWKHY

C.1.229 Reproduction Report

Title Original Study: The Geography of Repression and Opposition to Autocracy

doi: https://doi.org/10.1111/ajps.12614, American Journal of Political Science

Link to Full Report: https://www.socialsciencereproduction.org/reproductions/789/published/index?step=4

Link to Replicators' Package: https://github.com/pjesscarter/repression-replication

Original Authors' Package: https://doi.org/10.7910/DVN/EYAWES

C.1.230 Reproduction Report

Title Original Study: The Heterogeneous Tax Pass-Through Under Different Vertical Relationships

doi: https://doi.org/10.1093/ej/ueac007, Economic Journal

Link to Full Report: https://osf.io/a5py6/

Original Authors' Response: We appreciate the effort you made in replicating our study and sharing with us your conclusions. After carefully reading the report, we consider there is no need to write a response, as everything is fine from our side.

Original Authors' Package: https://doi.org/10.5281/zenodo.5824590

C.1.231 Reproduction Report

Title Original Study: The Impact of Presidential Appointment of Judges: Montesquieu or the Federalists?

doi: https://doi.org/10.1257/app.20210176, American Economic Journal: Applied Economics

Link to Full Report: https://osf.io/gesyq/

Original Authors' Response: Thank you. It is fine. I have no issues with their replication and their report seems to do a great job. I thank them and you for their time spent on my paper.

Original Authors' Package: https://doi.org/10.3886/E148001V1

C.1.232 Reproduction Report

Title Original Study: The Interplay Among Savings Accounts and Network-Based Financial Arrangements: Evidence from a Field Experiment

doi: https://doi.org/10.1093/ej/ueac053, Economic Journal

Link to Full Report: https://osf.io/qzgu7

Link to Original Authors' Response: https://osf.io/qzgu7

Original Authors' Package: https://doi.org/10.5281/zenodo.6985683

C.1.233 Reproduction Report

Title Original Study: The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund

doi: https://doi.org/10.1257/pol.20190299, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/80.htm

Link to Replicators' Package: https://osf.io/6atfw/

Original Authors' Final Response: "Thanks for putting in all this effort to evaluate the robustness of our results! I [Marinescu] think this is really a worthwhile endeavor."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/140121/version/V1/view

C.1.234 Reproduction Report

Title Original Study: The Long-Run Effects of Sports Club Vouchers for Primary School Children

doi: https://doi.org/10.1257/pol.20200431, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/46.htm

Link to Replicators' Package: https://osf.io/4bnjt/

Original Authors' Response: "We would like to thank the authors for their interest in our paper. We greatly appreciate their careful reading of the paper and the insightful robustness exercises they conducted. We are pleased that our results were successfully reproduced using different software packages, and that the additional robustness analyses performed by the authors further strengthen and support our conclusions."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/138922/version/V1/view

C.1.235 Reproduction Report

Title Original Study: Long-Term Effects of Equal Sharing: Evidence from Inheritance Rule for Land

doi: https://doi.org/10.1093/ej/ueae040, Economic Journal

Link to Full Report and Authors' Response: https://osf.io/eqbpu/

Original Authors' Package: https://doi.org/10.5281/zenodo.11186567

C.1.236 Reproduction Report

Title Original Study: The Long-Term Effects of Measles Vaccination on Earnings and Employment (Team 1)

doi: https://doi.org/10.1257/pol.20190509, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/33.htm

Link to Replicators' Package: https://osf.io/jv7kx/

Link to Original Authors' Response: https://osf.io/qxjnk/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/138401/version/V1/view

C.1.237 Reproduction Report

Title Original Study: The Long-Term Effects of Measles Vaccination on Earnings and Employment (Team 2)

doi: https://doi.org/10.1257/pol.20190509, American Economic Journal: Economic Policy

Link to Full Report: https://michaelwiebe.com/assets/atwood/atwood_comment.pdf

Original Authors' Response: No response.

Original Authors' Package: https://doi.org/10.3886/E138401V1

C.1.238 Reproduction Report

Title Original Study: The Macroeconomic Impact of Europe's Carbon Taxes

doi: https://doi.org/10.1257/mac.20210052, American Economic Journal: Macroeconomics

Link to Full Report: https://osf.io/q9knf/

Original Authors' Response: Thanks for doing a replication analysis of our paper. We have no concerns about the report and will not provide a formal response.

Original Authors' Package: https://doi.org/10.3886/E164521V1

C.1.239 Reproduction Report

Title Original Study: The Macroeconomics of Sticky Prices with Generalized Hazard Functions

doi: https://doi.org/10.1093/qje/qjab042, Quarterly Journal of Economics

Link to Full Report: https://github.com/atyho/Ottawa-Replication-Games-2023/blob/main/Ho_Huynh_Rea_Replication_Report.pdf

Link to Replicators' Package: https://github.com/atyho/Ottawa-Replication-Games-2023/

Link to Original Authors' Response:

Original Authors' Package: https://doi.org/10.7910/DVN/IBM0IE

C.1.240 Reproduction Report

Title Original Study: The Mechanics of the Industrial Revolution

doi: https://doi.org/10.1086/720890, Journal of Political Economy

Link to Full Report: https://osf.io/puknf/

Original Authors' Response: Beyond expressing our satisfaction that our findings were robust we have very little to say except to thank the authors for their meticulous analysis, and the Institute for its valuable work.

Original Authors' Package: https://www.journals.uchicago.edu/doi/suppl/10.1086/720890/suppl_file/20200559data.zip

C.1.241 Reproduction Report

Title Original Study: The Morning After: Cabinet Instability and the Purging of Ministers after Failed Coup Attempts in Autocracies

doi: https://doi.org/10.1086/716952, Journal of Politics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/45.htm

Link to Replicators' Package: https://doi.org/10.7910/DVN/21HZCC

Link to Original Authors' Response: https://osf.io/sm526/

Original Authors' Package: https://doi.org/10.7910/DVN/GCDJ25

C.1.242 Reproduction Report

Title Original Study: The Opportunity Cost of Debt Aversion

doi: https://doi.org/10.1257/aer.20221509, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/rhvyx/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/194543/version/V1/view

C.1.243 Reproduction Report

Title Original Study: The Origin of the State: Land Productivity or Appropriability?

doi: https://doi.org/10.1086/718372, Journal of Political Economy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/82.htm

Link to Replicators' Package: https://osf.io/ekzdg/

Original Authors' Response: Comments taken into account in the report.

Original Authors' Package: https://doi.org/10.1086/718372/suppl_file/2018030data.zip

C.1.244 Reproduction Report

Title Original Study: The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development

doi: https://doi.org/10.1257/aer.103.1.1, American Economic Review

Link to Full Report: https://osf.io/w7vpu/files/osfstorage

Link to Replicators' Package: https://osf.io/w7vpu/files/osfstorage

Original Authors' Response: No formal response.

Original Authors' Package: https://doi.org/10.3886/E112588V1

C.1.245 Reproduction Report

Title Original Study: The PhD Pipeline Initiative Works: Evidence from a Randomized Intervention to Help Underrepresented Students Prepare for PhDs in Political Science

doi: https://doi.org/10.1086/726954, The Journal of Politics

Link to Full Report: https://osf.io/75kqp/

Original Authors' Response: Thank you for sharing the report. I do not plan to write a formal follow up. One point to note is that the report says that "the stargazer package requires that the models included in the same table have the same number of observations." I have never found this to be the case, so perhaps the settings used by the reproducer were different. I only mention this, since I think this is an inaccurate description of what the Stargazer package can do.

Original Authors' Package: https://doi.org/10.7910/DVN/UPLZAK

C.1.246 Reproduction Report

Title Original Study: The People Think What I Think: False Consensus and Unelected Elite Misperception of Public Opinion

doi: https://doi.org/10.1111/ajps.12833, American Journal of Political Science

Link to Full Report: https://osf.io/fz2xs/

Original Authors' Response: The authors were satisfied with the report and did not write a formal response. Furnas: "Happy to see it out there!"

Original Authors' Package: https://doi.org/10.7910/DVN/3VFVS7

C.1.247 Reproduction Report

Title Original Study: The Political Consequences of Green Policies: Evidence from Italy

doi: https://doi.org/10.1017/S0003055423000308, American Political Science Review

Link to Full Report: https://osf.io/za5t8/

Link to Original Authors' Response: https://osf.io/za5t8/

Original Authors' Package: https://doi.org/10.7910/DVN/NSC8TJ

C.1.248 Reproduction Report

Title Original Study: The Power of Hydroelectric Dams: Historical Evidence from the United States over the Twentieth Century

doi: https://doi.org/10.1093/ej/ueac059, Economic Journal

Report's Abstract: Successful computational reproducibility. No coding errors uncovered.

Authors' Response: Not contacted.

Original Authors' Package: https://doi.org/10.1093/ej/ueac059

C.1.249 Reproduction Report

Title Original Study: The Psychosocial Value of Employment: Evidence from a Refugee Camp

doi: https://doi.org/10.1257/aer.20211616, American Economic Review

Link to Full Report: https://osf.io/cvkdz

Link to Original Authors' Response: https://osf.io/cvkdz

Original Authors' Package: https://doi.org/10.3886/E174781V1

C.1.250 Reproduction Report

Title Original Study: The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation

doi: https://doi.org/10.1257/aer.20191852, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/59.htm

Link to Replicators' Package: https://osf.io/fge7z/

Original Author's Response: "Thanks for replicating the paper. I don't have any comments to add to the report."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/146041/version/V1/view

C.1.251 Reproduction Report

Title Original Study: The Returns to Microenterprise Support among the Ultrapoor: A Field Experiment in Postwar Uganda

doi: https://doi.org/10.1257/app.20150023, American Economic Journal: Applied Economics

Full Report: Successful computational reproducibility conducted during the World Bank-IMF games.

Authors' Response: Not contacted.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/113629/version/V1/view

C.1.252 Reproduction Report

Title Original Study: The Side Effects of Immunity: Malaria and African Slavery in the United States

doi: https://doi.org/10.1257/app.20190372, American Economic Journal: Applied Economics

Link to Full Report: https://osf.io/728ud/

Link to Replicators' Package: https://osf.io/728ud/

Original Authors' Response: Original author provided feedback. No final response on the updated version.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/120483/version/V1/view

C.1.253 Reproduction Report

Title Original Study: The Wheels of Change: Technology Adoption, Millwrights and the Persistence in Britain'S Industrialisation

doi: https://doi.org/10.1093/ej/ueab102, Economic Journal

Link to Full Report: https://osf.io/gdne3/

Link to Replicators' Package: https://osf.io/tws8n/

Original Authors' Response: No response.

Original Authors' Package: https://zenodo.org/records/5734954

C.1.254 Reproduction Report

Title Original Study: They're Still There, He's All Gone: American Fatalities in Foreign Wars and Right-Wing Radicalization at Home

doi: https://doi.org/10.1017/S0003055423000904, American Political Science Review

Link to Full Report and Original Authors' Response: Back and forth with authors not completed.

Original Authors' Package: https://doi.org/10.7910/DVN/4GLPII

C.1.255 Reproduction Report

Title Original Study: Top of the Batch: Interviews and the Match

doi: https://doi.org/10.1257/aeri.20200800, American Economic Review: Insights

Link to Full Report: https://osf.io/63qe2/

Original Authors' Response: Thank you for the replication report; very interesting. We won't write a response.

Original Authors' Package: https://doi.org/10.3886/E141541V1

C.1.256 Reproduction Report

Title Original Study: Transfers, Diversification and Household Risk Strategies: Can Productive Safety Nets Help Households Manage Climatic Variability?

doi: https://doi.org/10.1093/ej/ueac018, Economic Journal

Link to Full Report: https://osf.io/5nrvu/

Original Authors' Response: Authors did not respond.

Original Authors' Package: https://zenodo.org/record/5933958

C.1.257 Reproduction Report

Title Original Study: Transmitting Rights: Effective Cooperation, Inter-Gender Contact, and Student Achievement.

doi: https://doi.org/10.1093/10.1257/pol.20230620, American Economic Journal: Economic Policy

Link to Full Report and Original Authors' Response: https://osf.io/c5s8b/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/208265/version/V1/view

C.1.258 Reproduction Report

Title Original Study: UN Peacekeeping and Democratization in Conflict-Affected Countries

doi: https://doi.org/10.1017/S0003055422001319, American Political Science Review

Link to Full Report: https://osf.io/jakq7/

Original Authors' Response: Thank you for this report, and for taking the time to replicate our work. We are of course very pleased to hear that the results proved to be robust.

Original Authors' Package: https://doi.org/10.7910/DVN/UOYDHN

C.1.259 Reproduction Report

Title Original Study: Understanding Ethnolinguistic Differences: The Roles of Geography and Trade

doi: https://doi.org/10.1093/ej/ueab065, Economic Journal

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/62.htm

Link to Replicators' Package: https://osf.io/k3p7g/

Link to Original Authors' Response: https://econpapers.repec.org/paper/zbwi4rdps/63.htm

Original Authors' Package: https://doi.org/10.1093/ej/ueab065

C.1.260 Reproduction Report

Title Original Study: Unpacking P-hacking and Publication Bias

doi: https://doi.org/10.1257/aer.20210795, American Economic Review

Link to Full Report: https://osf.io/rvynd/

Link to Original Authors' Response: https://osf.io/rvynd/

Original Authors' Package: https://doi.org/10.3886/E192271V1

C.1.261 Reproduction Report

Title Original Study: Vulnerability and Clientelism (Team 1)

doi: https://doi.org/10.1257/aer.20190565, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/83.htm

Link to Replicators' Package: https://osf.io/q2tw6/

Link to Original Authors' Response: https://econpapers.repec.org/paper/zbwi4rdps/84.htm

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/173341/version/V1/view

C.1.262 Reproduction Report

Title Original Study: Vulnerability and clientelism (Team 2)

doi: https://doi.org/10.1257/aer.20190565, American Economic Review

Link to Full Report: https://osf.io/w7vpu/

Link to Replicators' Package: https://osf.io/w7vpu/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/173341/version/V1/view

C.1.263 Reproduction Report

Title Original Study: Vulnerable Growth

doi: https://doi.org/10.1257/aer.20161923, American Economic Review

Full Report: Computational reproducibility conducted during the World Bank-IMF games.

Authors' Response: Not contacted.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/113169/version/V1/view

C.1.264 Reproduction Report

Title Original Study: Wage Cyclicality and Labor Market Sorting

doi: https://doi.org/10.1257/aeri.20210161, American Economic Review: Insights

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/78.htm

Link to Replicators' Package: https://osf.io/a8hcg/

Original Authors' Response: "I have read the report and I do not wish to write a reply.

Congratulations on this initiative – it is great!"

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/150581/version/V1/view

C.1.265 Reproduction Report

Title Original Study: War, Socialism, and the Rise of Fascism: an Empirical Exploration

doi: https://doi.org/10.1093/qje/qjac001, Quarterly Journal of Economics

Link to Full Report: https://osf.io/a672c/

Link to Replicators' Package: https://osf.io/a672c/

Link to Original Authors' Response: No response.

Original Authors' Package: https://doi.org/10.7910/DVN/CLJTSC

C.1.266 Reproduction Report

Title Original Study: Web of Power: How Elite Networks Shaped War and Politics in China

doi: https://doi.org/10.1093/qje/qjac041, Quarterly Journal of Economics

Link to Full Report: https://osf.io/7y9cf/

Link to Replicators' Package: https://osf.io/7y9cf/

Original Authors' Response: https://osf.io/7y9cf/

Original Authors' Package: https://doi.org/10.7910/DVN/0H4HG2

C.1.267 Reproduction Report

Title Original Study: We Are All Behavioural, More, or Less: A Taxonomy of Consumer Decision-Making

doi: https://doi.org/10.1093/restud/rdac055, Review of Economic Studies

Link to Full Report: https://osf.io/g4dp5/

Original Authors' Response: No formal response.

Original Authors' Package: https://doi.org/10.5281/zenodo.6783408

C.1.268 Reproduction Report

Title Original Study: What Makes Anticorruption Punishment Popular? Individual-Level Evidence from China

doi: https://doi.org/10.1086/715252, Journal of Politics

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/7.htm

Link to Replicators' Package: https://osf.io/czs6j/

Original Authors' Response: "We appreciate your efforts, both in replicating our paper and in doing so systematically for other studies in leading political science and economic journals. Your contribution is valuable to the entire academic community and to us especially.

We also appreciate your sharing Reproduction Reports with the original authors prior to dissemination and are glad to see from the Reproduction Report that our results and methods appear to be both valid and robust. Although a longer follow-up may not be necessary, we do wish to convey our gratitude to the replicator(s) and to the editorial team."

Original Authors' Package: https://doi.org/10.7910/DVN/XTRWKG

C.1.269 Reproduction Report

Title Original Study: When a Doctor Falls from the Sky: The Impact of Easing Doctor Supply Constraints on Mortality

doi: https://doi.org/10.1257/aer.20210701, American Economic Review

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/53.htm

Link to Replicators' Package: https://github.com/e-mcmanus/Okeke23_Replication

Original Authors' Response: "Thank you for sharing the Reproduction Report (and please pass on my thanks to the replicators). There does not appear to be much for me to respond to. It is gratifying to see that the results have held up well to additional scrutiny."

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/181581/version/V1/view

C.1.270 Reproduction Report

Title Original Study: When Choices Are Mistakes

doi: https://doi.org/10.1257/aer.20201550, American Economic Review

Link to Full Report and Authors' Response: https://osf.io/jubrh/

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/164661/version/V1/view

C.1.271 Reproduction Report

Title Original Study: When Do Männerparteien Elect Women? Radical Right Populist Parties and Strategic Descriptive Representation

doi: https://doi.org/10.1017/S0003055422000107, American Political Science Review

Link to Full Report: https://osf.io/p4m98/

Link to Original Authors' Response: https://osf.io/p4m98/

Original Authors' Package: https://doi.org/10.7910/DVN/SG55BJ

C.1.272 Reproduction Report

Title Original Study: When Do Politicians Appeal Broadly? The Economic Consequences of Electoral Rules in Brazil

doi: https://doi.org/10.1257/app.20210529, American Economic Journal: Applied Economics

Link to Full Report: https://osf.io/pnf8k/

Original Authors' Response: Thank you for reaching out. It looks like a great initiative, and I have no response to the exercise.

Original Authors' Package: https://doi.org/10.3886/E168901V1

C.1.273 Reproduction Report

Title Original Study: Where Motivated Reasoning Withers and Looms Large: Fear and Partisan Reactions to the COVID-19 Pandemic

doi: https://doi.org/10.1111/ajps.12808, American Journal of Political Science

Link to Full Report and Original Authors' Response: https://osf.io/rqc9p/

Original Authors' Package: https://doi.org/10.7910/DVN/TKLSNE

C.1.274 Reproduction Report

Title Original Study: Who Controls the Past: Far-Sighted Bargaining in International Regimes

doi: https://doi.org/10.1257/app.20210529, American Journal of Political Science

Link to Full Report and Authors' Response: Back and forth between authors and replications still ongoing: https://osf.io/26j5t/

Original Authors' Package: https://doi.org/10.7910/DVN/5M10KG

C.1.275 Reproduction Report

Title Original Study: Who Chooses Commitment? Evidence and Welfare Implications

doi: https://doi.org/10.1093/restud/rdab056, Review of Economic Studies

Link to Full Report: https://osf.io/752q9/

Link to Replicators' Package: https://osf.io/752q9/

Original Authors' Response: The authors provided feedback which was taken into account.

Original Authors' Package: https://zenodo.org/records/5173081

C.1.276 Reproduction Report

Title Original Study: Who Sells During a Crash? Evidence from Tax Return Data on Daily Sales of Stock

doi: https://doi.org/10.1093/ej/ueab059, Economic Journal

Link to Full Report: https://osf.io/b6s9k/

Link to Replicators' Package: https://www.dropbox.com/scl/fo/c3ysdlenysq391mugzprm/h?rlkey=riooyohci7i5vwx475r13jaqq&dl=0

Original Authors' Response: The authors provided feedback which was taken into account.

Original Authors' Package: https://doi.org/10.1093/ej/ueab059

C.1.277 Reproduction Report

Title Original Study: Why Don't Firms Hire Young Workers During Recessions?

doi: https://doi.org/10.1093/ej/ueab096, Economic Journal

Link to Full Report: https://osf.io/3pqbt/

Link to Replicators' Package: https://github.com/jcrechet/replication_forsythe_2022_EJ

Link to Original Authors' Response: The author provided feedback but did not provide a response.

Original Authors' Package: https://zenodo.org/records/5710784

C.1.278 Reproduction Report

Title Original Study: Why Don't Remittances Appear to Affect Growth?

doi: https://doi.org/10.1111/ecoj.12463, Economic Journal

Report: Successful computational reproducibility. The replicators did not conduct robustness checks.

Authors' Response: Not contacted.

Original Authors' Package: https://doi.org/10.1111/ecoj.12463

C.1.279 Reproduction Report

Title Original Study: Winter is Coming: Early-life Experiences and Politicians' Decisions

doi: https://doi.org/10.1093/ej/uead061, Economic Journal

Link to Full Report: https://osf.io/hvwbs/

Original Authors' Response: Thank you very much for your work! We think replication and dissemination are important for the reliability of scientific works, and we are glad that our works could be replicated by your team. We have read the report and think the results of the sensitivity analysis are not a threat to our main results, given the lack of biographic information of many county secretaries. Therefore, we prefer not to write a formal response in the current stage. Thank you for your understanding.

Original Authors' Package: https://doi.org/10.5281/zenodo.8188982

C.1.280 Reproduction Report

Title Original Study: Within-Firm Productivity Dispersion: Estimates and Implications

doi: https://doi.org/10.1086/720465, Journal of Political Economy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/181.htm

Link to Original Authors' Response: https://econpapers.repec.org/paper/zbwi4rdps/182.htm

Original Authors' Package: https://www.journals.uchicago.edu/doi/suppl/10.1086/720465/suppl_file/20200297data.zip

C.1.281 Reproduction Report

Title Original Study: Worth Your Weight: Experimental Evidence on the Benefits of Obesity in Low-Income Countries

doi: https://doi.org/10.1257/aer.20211879, American Economic Review

Link to Full Report: https://osf.io/5zfu4/

Original Author's Response: Thank you so much Abel for sharing this work with me! This is really great. I will post it on my website if that's okay with you?

Original Authors' Package: https://doi.org/10.3886/E181481V1

C.1.282 Reproduction Report

Title Original Study: Yellow Vests, Pessimistic Beliefs, and Carbon Tax Aversion

doi: https://doi.org/10.1257/pol.20200092, American Economic Journal: Economic Policy

Link to Full Report: https://econpapers.repec.org/paper/zbwi4rdps/58.htm

Original Authors' Response: Authors provided feedback which was taken into account. No response.

Original Authors' Package: https://www.openicpsr.org/openicpsr/project/128143/version/V1/view

C.1.283 Reproduction Report

Title Original Study: Your Place in the World: Relative Income and Global Inequality (Team 1)

doi: https://doi.org/10.1257/pol.20200343, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/896bd/

Original Authors' Package: https://doi.org/10.3886/E146541V1

C.1.284 Reproduction Report

Title Original Study: Your Place in the World: Relative Income and Global Inequality (Team 2)

doi: https://doi.org/10.1257/pol.20200343, American Economic Journal: Economic Policy

Link to Full Report and Authors' Response: https://osf.io/896bd/

Original Authors' Package: https://doi.org/10.3886/E146541V1