Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 24-1096

Integrating Open Science Principles into Quasi-Experimental Social Science Research

Blake H. Heller
University of Houston

Carly D. Robinson
Stanford University

Quasi-experimental methods are a cornerstone of applied social science, providing critical answers to causal questions that inform policy and practice. Although open science principles have influenced experimental research norms across the social sciences, these practices are rarely implemented in quasi-experimental research. In this paper, we explore how open science practices can enhance transparency, replicability, and credibility in quasi-experimental research. We discuss practical strategies to implement or adapt these practices for quasi-experimental researchers. We also emphasize the shared responsibility of external stakeholders, including data providers, journals, and funders to create the circumstances and incentives for open science practices to proliferate. We believe that all quasi-experimental work can benefit from an open science mentality, and this mindset shift will ultimately enhance the credibility, accessibility, replicability, and unbiasedness of quasi-experimental social science research.

VERSION: November 2024

Suggested citation: Heller, Blake H., and Carly D. Robinson. (2024). Integrating Open Science Principles into Quasi-Experimental Social Science Research. (EdWorkingPaper: 24 -1096). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/zr8a-rn80

Title: Integrating Open Science Principles into Quasi-Experimental Social Science Research*

Author Information:

Blake H. Heller
University of Houston
bhheller@central.uh.edu
ORCID ID: 0000-0003-2093-8170

Carly D. Robinson
Stanford University
carly_robinson@stanford.edu
ORCID ID: 0000-0002-0663-1589

Last Updated: November 19, 2024

Abstract

Quasi-experimental methods are a cornerstone of applied social science, providing critical answers to causal questions that inform policy and practice. Although open science principles have influenced experimental research norms across the social sciences, these practices are rarely implemented in quasi-experimental research. In this paper, we explore how open science practices can enhance transparency, replicability, and credibility in quasi-experimental research. We discuss practical strategies to implement or adapt these practices for quasi-experimental researchers. We also emphasize the shared responsibility of external stakeholders, including data providers, journals, and funders to create the circumstances and incentives for open science practices to proliferate. We believe that all quasi-experimental work can benefit from an open science mentality, and this mindset shift will ultimately enhance the credibility, accessibility, replicability, and unbiasedness of quasi-experimental social science research.

Keywords: Open science, quasi-experimental research, quasi-experiments, research methods, preregistration

JEL codes: O31, C81

---

* Financial support from the National Academy of Education and the National Academy of Education/Spencer Postdoctoral Fellowship is gratefully acknowledged. Opinions expressed do not represent the views of the National Academy of Education. Correspondence can be addressed to the authors at: bhheller@central.uh.edu (Heller) and carly_robinson@stanford.edu (Robinson). All errors are our own.

"True understanding of how best to structure and incentivize science will emerge slowly and will never be finished. That is how science works." (Munafò et al., 2017)

1. Motivation

Answering causal questions about human behavior is a central focus of applied social science research. Borrowing from and building upon empirical methods popularized in the agricultural and medical sciences, quantitative social scientists employ a wide range of tools to identify causal relationships in data (see Imbens, 2024 for a recent summary). While experimental research or randomized controlled trials (RCTs) are widely regarded as the gold-standard for causal inference, many questions cannot be answered in a laboratory, field, or survey experiment for practical or ethical reasons, and strictly observational or descriptive research is generally insufficient to answer causal questions (e.g., Lalonde, 1986).

In applied settings, quasi-experimental research designs are an increasingly important source of evidence to guide policy decisions or organizational behavior (Angrist & Pischke, 2010; Council of Economics Advisors, 2014; Currie, Kleven & Zwiers, 2020; Goldsmith-Pinkham, 2024). In quasi-experimental research, investigators may leverage idiosyncratic aspects of an intervention's timing, location, intensity, or eligibility criteria to identify as-good-as-random variation in exposure to a "treatment" and assess its impact on individuals or communities.

Open science principles have influenced *experimental* research norms across the social sciences (Munafò et al., 2017; Christensen et al., 2020; Logg & Dorison, 2021), however many common experimental open science practices are rarely implemented in *quasi-experimental* research (Christensen et al., 2020; Miguel, 2021; Hardwicke & Wagenmakers, 2023). In a 2019 essay posted on the World Bank's *Development Impact* blog, Berk Özler lamented the challenge of using existing preregistration platforms to accommodate quasi-experimental research designs. While he was able to find a repository that could be adapted to accommodate his propensity-score matching research design (the now-defunct EGAP registry, which has since been integrated into Open Science Framework [OSF] Registries), none of the dominant social science preregistration platforms were set up with quasi-experimental research in mind, and some continue to actively exclude non-experimental research.

Importantly, Özler's concerns were not that his study would struggle to be published in the absence of preregistration, but that he and his research team were vulnerable to the same unintentional biases that catalyzed the open science revolution to begin with: "The trouble is, I do not trust myself (or my collaborator) to not be affected in our choices by the impact findings, their statistical significance, etc. Not that I have a stake in the success of the program: it is more that I am worried about subconscious choices that can take the analysis in one direction than the other – exactly because I can see the consequences of these choices pretty easily if I have all the data[…]" (Özler, 2019).

Recent work interrogating the influence of these subconscious choices, which are typically invisible or very hard to identify in published work, suggests that these concerns are warranted. Decisions about variable definitions, sample exclusions, and model parameters can introduce substantial variation into quasi-experimental estimates of program effects—even when well-intentioned researchers use the same data to answer the same question (Huntington-Klein et al., 2021; 2024; Breznau et al., 2022; Holzmeister et al., 2024; Wuttke et al., 2024). Notably, these findings occur in the absence of publication incentives related to the direction, magnitude, or statistical significance of results. Munafò and colleagues (2017) similarly emphasized the need for “measures to counter the natural tendency of enthusiastic scientists who are motivated by discovery to see patterns in noise” (p.2).

The stakes of ensuring the credibility and reliability of evidence from quasi-experimental research are high, as the applied focus of quasi-experimental research means that it often yields the best available evidence to drive policy decisions. U.S. Presidents have cited quasi-experimental research to justify policy positions in every *Economic Report of the President* since at least 2010 (e.g., Council of Economic Advisors, 2014; 2018; 2022) as well as in speeches like the State of the Union (e.g., Obama, 2014; Biden, 2024). The credibility and accuracy of quasi-experimental studies can have real consequences that influence billions of dollars in public spending and philanthropic investment (Gibbons, 2023; Bill & Melinda Gates Foundation, n.d.; The White House, n.d.). Ultimately, if quasi-experimental research falls short, for whatever reason, the consequences can have profound real-world implications: misinformed policies, ineffective programs, and the perpetuation of inequities.

In this paper, we explore the potential for open science principles and practices to enhance quasi-experimental social science research. We describe open science research practices and identify opportunities to implement or adapt these practices to different quasi-experimental research scenarios. We echo Reich’s (2021) assertion that “[...]open science practices exist on a spectrum rather than as a binary, and researchers need not adopt whole cloth the approaches of experimentalists to find specific ways of increasing transparency in research design that work in particular subfields and methodological approaches” (p. 102). Not every quasi-experimental study can or should integrate every open science practice. However, we believe that all quasi-experimental work can benefit from adopting an open science mentality and can rely on many of the same tools to enhance the credibility, accessibility, replicability, and unbiasedness of quantitative social science research.

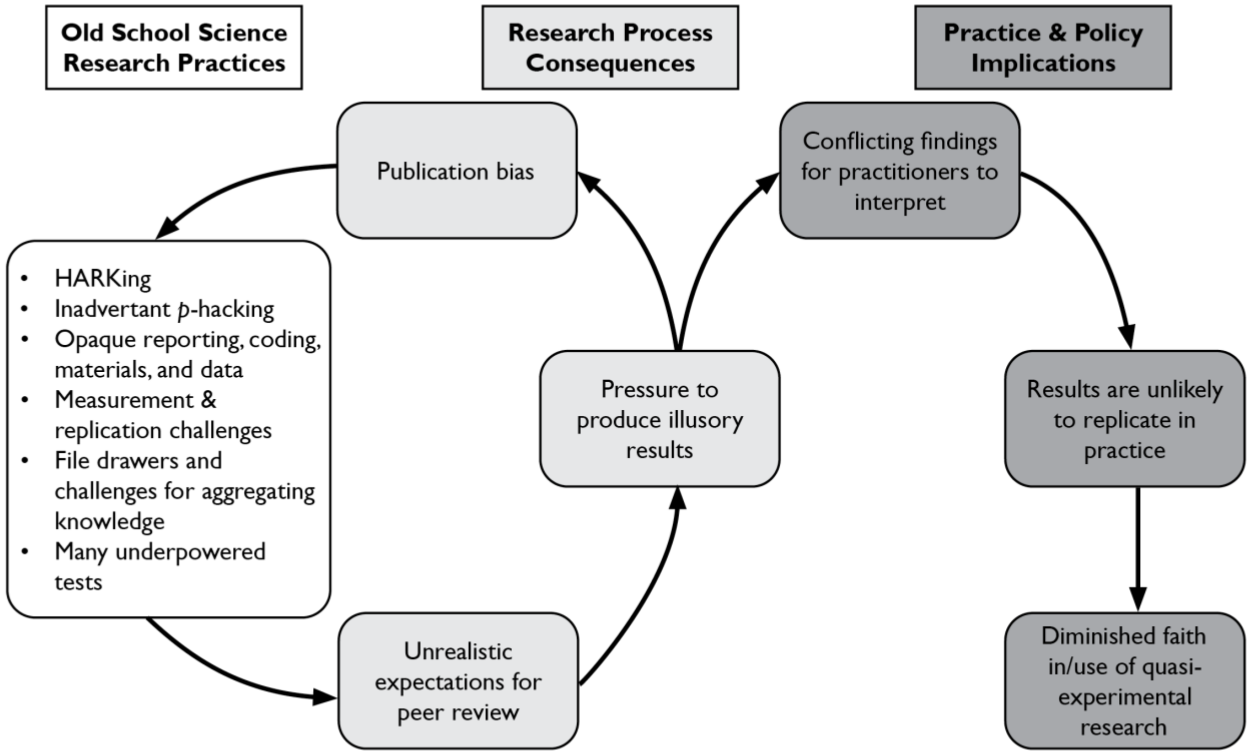
2. Background on Open Science and Causal Inference

The most popular and well-understood method to separate causation from correlation is the randomized controlled trial or experiment (RCT). Widely regarded as the “gold standard” for causal inference with respect to internal validity, experiments have been used in psychology since the early 1900s (e.g., Woodworth & Thorndike, 1901) and are growing in popularity in other social science disciplines (e.g., Currie, Kleven & Zwiers, 2020). In the early 2010s, the so-called “replication crisis” catalyzed substantial changes in experimental psychology research norms, many of which have spilled over to other fields.

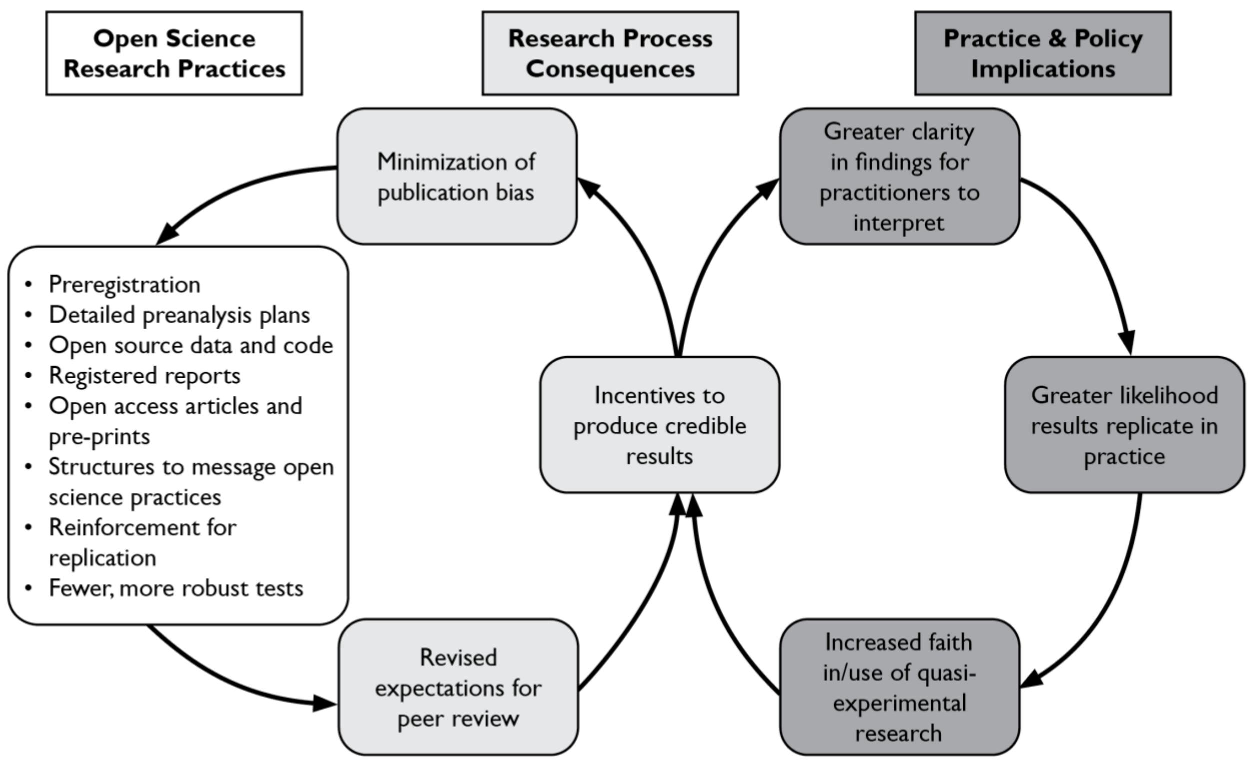
Figure 1, adapted from Gehlbach and Robinson (2021) synthesizes the contrast between “old school” experimental research practices that led to the replication crisis in psychology with “open science” research practices that align with the values of transparency, inclusivity, and honesty in pursuit of the truth. For example, while “old school” norms led to unchecked *p*-hacking via ex-post modeling decisions that are easily justified after data is analyzed (i.e., via

Figure 1: Contrasting the features and consequences of "Old school" research practices (panel a) with "Open science" research practices (panel b).

(a)



(b)



Note: Adapted from Robinson and Gehlbach (2021).

"Hypothesizing After Results are Known" or HARKing), today the default expectation is that experimental social science research is preregistered with a detailed preanalysis plan. In the past, research teams might run dozens of underpowered trials to achieve a statistically significant result, leaving untold numbers of null results to rot in file drawers. Preregistration encourages scholars to take power analyses seriously and adequately power their studies (van den Akker et al., 2024). When interventions do not, in fact, impact outcomes of interest, the resulting precise null results are, in theory, more likely to be published (Nosek et al., 2018)—and some are even accepted before data is collected as more and more highly-regarded journals accept registered reports.

Historically, replication efforts have been stifled by limited access to data, and replicators struggled to decipher researcher decisions about variable definitions, model parameters, and sample exclusions from the text of an article. Today, journals generally encourage (or require) scholars to facilitate replication by providing public access to the code used to clean and analyze their data, alongside de-identified datasets when possible. Researchers can also democratize access to the research frontier by adopting the open science practices of publishing preprints or publishing in open-access journals, rather than allowing slow publication processes to stall the release of cutting-edge research that remains inaccessible to the public behind expensive paywalls.

While recent research suggests that preregistration with preanalysis plans reduce *p*-hacking and publication bias in experimental research (Brodeur et al., 2024), it is unclear whether and how open science practices influence trust in research or perceptions of researchers' credibility (Field et al., 2020). As knowledge of the benefits and ease of implementing open science practices spreads, we expect perceptions and research norms to coevolve until these practices are close to universally adopted in causal research.

3. How can scholars integrate open science principles into quasi-experimental research?

Today, the process of conceiving, designing, implementing, and analyzing data to conduct a quasi-experimental research project involves a tremendous degree of researcher choice up to the point of publication (Simmons, Nelson & Simonsohn, 2011; Huntington-Klein et al., 2021; 2024). Hypothesis generation, sample selection, model choice, and myriad other parameters may be decided, adjusted, or finalized at any point of the research process, including during peer review, without documentation.

While some projects may deviate only slightly (or not at all) from a researcher's initial research design, there is very little transparency about how modeling decisions are made, and if replication materials are shared (e.g., data cleaning and/or analysis code), only final versions see the light of day. Even when dozens of models or sample selection criteria are tested to produce the final results, all other candidate models are essentially left on the cutting room floor. A growing body of evidence suggests that this process introduces variance to quasi-experimental impact estimates that is unaccounted for by standard errors (Huntington-Klein et al., 2021; 2024; Breznau et al., 2022; Holzmeister et al., 2024; Wuttke et al., 2024).

In this section, we discuss seven open science practices in greater detail to create a practical guide describing how each practice can be implemented by quasi-experimental researchers to increase transparency, replicability, and integrity of quasi-experimental social science research. We describe each practice, explain the problem(s) it solves or ameliorates, and the circumstances under which it can be successfully implemented in quasi-experimental studies.

3.1 PREREGISTRATION

The most basic preregistration establishes when a research project was formally initiated as well as the fundamental questions that project is designed to answer. A high-quality preregistration (especially when paired with a detailed preanalysis plan) accomplishes five goals:

1) Establishes the timeline of when a research project was initiated;
2) Establishes the primary hypotheses that the project is designed to test;
3) Increases the transparency of methods to facilitate replication and limit scope for HARKing;
4) Allows for a principled approach to multiple hypothesis correction; and
5) Enhances the credibility of prespecified results versus those that emerge from exploratory analyses.

We argue that these five goals are worthwhile—and often attainable—for any study that aims to identify the causal impact of a policy or practice, and researchers can benefit from preregistering quasi-experimental studies (and in some cases, descriptive research). Social scientists who conduct experiments now are expected, if not required to preregister their studies in order to publish in top peer-reviewed journals or receive funding from a growing list of private foundations (e.g., Arnold Ventures, the Russell Sage Foundation) and government agencies (e.g., NIH, IES). However, the vast majority of quasi-experimental studies are not preregistered and, at present, there are few incentives to do so. As of 2024, the American Economic Association (AEA) RCT registry (socialscienceregistry.org) does not even have a mechanism for preregistering non-experimental studies.

At the most basic level, preregistrations occur *before* researchers can exert any influence of the analysis process, ensuring that the study design and hypotheses are documented in advance, based on the considered judgements of the research team, without being influenced ex-post considerations related to statistical significance or other relationships in the data. In practice, researchers can preregister a study with just their hypotheses and outcome variables and very little additional detail. However, even among well-intentioned researchers, without a detailed preanalysis plan, preregistration alone is likely insufficient to prevent *p*-hacking and HARKing (Brodeur et al., 2024). In fact, some would argue (and we agree) that a preregistration without a preanalysis plan that documents the details of the actual analysis should not be considered a valid preregistration at all (e.g., Klonsky, 2024).

3.2 PREANALYSIS PLANS

Humans are natural storytellers (McAdams & McLean, 2013; Gehlbach & Robinson, 2018), making it easy to justify one data cleaning or modeling choice over others after seeing which option yields more favorable results. Thus, the general rule is to leave only one modeling path per hypothesis, which keeps researchers from making idiosyncratic post-hoc decisions, while also allowing for future work to replicate or extend the analysis.

For quasi-experimental studies, this means that insofar as is possible, researchers should commit to their primary hypotheses, primary outcomes, identification strategy, empirical model, model parameters, definition of treatment, principles of sample construction, and variable definitions before observing outcome data. For research questions that emerge after researchers begin analyzing outcomes, preregistration is not possible. However, for studies that are conceived as a result of a natural experiment, where the outcomes have not yet been observed (or, ideally, not yet been realized), preregistration is possible.

Although there are key differences between preregistering experimental and non-experimental studies, many of the general tenets are consistent. For instance, researchers conducting preregistered quasi-experimental or experimental research should be able to describe the study, list their hypotheses, explain principles of sample construction, define key variables, and distinguish between primary and exploratory outcomes or hypotheses before accessing outcome data that could influence those decisions based on ex-post considerations. However, where experimental studies require information on the experimental design, randomization procedures, sample recruitment, and treatment-control contrast, researchers would instead outline their quasi-experimental study's identification strategy, sample exclusions, empirical model, model parameters, and treatment-comparison contrast. This act of stating the key features of a study prior to analyses sends the signal of credibility and decreases the likelihood that the findings are spurious, guided by ad-hoc, ex-post modeling decisions.

Perhaps the greatest difference for preregistering studies with quasi-experimental designs is the existence of the data. While experimental studies always collect data prospectively, quasi-experimental studies can use a variety of data sources: existing data, data currently being collected, or data that will be collected in the future. For the latter two, the preregistration process follows a similar timeline as an experimental study might–the researchers can post the preregistration before the outcome data has been realized or observed. However, many quasi-experimental studies rely on outcome data that already exists, making the *timing* of researchers' access to the data a central consideration in the credibility of the preregistration.

Some cases allow researchers to approach quasi-experimental preregistration as if they are designing an experiment. For example, when a study is preregistered before a policy is implemented, it would be physically impossible to access outcome data that could influence any aspect of the preregistration. In other cases, there may be less clarity on when researchers had access to outcome data. Even if data is open source and has been accessible for years, quasi-experimental researchers can engage in practices that can increase the credibility of their work. For example, researchers like Özler, who have the awareness that they might fall victim to the very human trait of HARKing, can still preregister a study if they themselves have not observed the outcome data, even if it is in-principle accessible. They can report the timeline in which they preregistered their study and when they accessed the data, and reviewers can choose to believe them (or not). Other open science practices, like sharing well-documented code and replication files, also play a role in increasing transparency in these cases. Another important consideration (and potential solution) is how the data is shared by data providers, which we discuss more in section 4.1 below.

Of course, applied research is messy, and sometimes decisions made before working with the data turn out to be impractical—or even impossible—to implement. Perhaps the data structure was different than expected, the initial identification strategy did not work, or a variable was measured less precisely than anticipated. It is often reasonable, and even correct, to deviate from one's preanalysis plan. When these types of deviations are necessary, open science principles simply imply that researchers should document and justify those decisions (see discussion in section 3.6 below), and readers or reviewers can decide whether they agree or disagree with the analytic choices (Lakens, 2024).

Relying on post-hoc modeling decisions invites the skepticism that researchers made choices that increased the likelihood they found a statistically significant or otherwise favorable result, but sometimes ex-post deviations from one's considered ex-ante judgments are unavoidable. Currently, resolving this skepticism is a central focus of peer review for many quasi-experimental studies, often requiring researchers to produce dozens of robustness checks and alternative specifications to convince referees that their results are not spurious or driven by HARKing. Publishing a preanalysis plan and documenting deviations from these prespecified decisions allows readers and peer-reviewers to calibrate their confidence in results by considering the trade-offs of the different approaches with a greater understanding of the researcher's ex-ante thought process and beliefs.

3.3 OPEN SOURCE DATA AND CODE

The open science community is increasingly encouraging researchers to publicly post replication packages that include the code and/or data used to conduct a study. This access allows other researchers to more easily review studies, validate the findings, and replicate the results. Sharing code and data promotes transparency and fosters collaboration across the research community.

In practice, the open source sharing of data and code does not differ drastically between experimental and quasi-experimental studies, and by a wide margin, this is the open science practice that quasi-experimental researchers most commonly report implementing in their work (Christensen et al., 2020). The general tenet is that researchers share the data and code used to reach the stated conclusions in a particular study (Neal, 2022). Thus, if the data is being used for other studies, the researchers do not need to share *all* the data–just those that are necessary to conduct the specific analyses represented in that study. In cases where publicly sharing data poses a legitimate threat to related, ongoing research, scholars may embargo or delay the posting of data until related projects are completed.

There are several repositories where researchers can post data and code (e.g, Harvard Dataverse; ICPSR; OSF; ResearchBox; Scientific Data). The best place to share data and materials is one that is "independent, accessible, and persistent" (Neal, 2022). On OSF, researchers can give data and code a specific license that governs how they can be used and even get credit for the data they post, as the materials are citable.

Because quasi-experimental research is, by definition, applied work that studies real-world phenomena and policies, researchers rarely create or own the data they use as experimentalists might. Therefore, data ownership and privacy raises questions about the feasibility of publicly posting data for many researchers. In some cases, the data may be publicly available, in which case researchers can simply describe when and how they accessed the data and post the code they used for cleaning and analysis. In other cases, the data is not public and researchers must preserve the confidentiality of the people or organizations represented in the data. Thus, researchers cannot simply post the dataset, but it still may be possible to share data in a form that preserves confidentiality through de-identification (removing any information that would allow a sophisticated third-party to identify specific people or organizations with high confidence), aggregation (combining microdata into groups at higher levels of observation), or suppression (removing variables that could be used to identify specific people or organizations) (Neal, 2022). Or, if a particular dataset required researchers to apply for access, the authors can detail the process for others to gain access to the data. In section 4.1 below, we discuss the role of data providers in facilitating open data practices.

3.4 REGISTERED REPORTS

One of the potential drawbacks of preregistration and preanalysis plans is that researchers are less likely to find statistically significant results through HARKing and *p*-hacking (Kaplan & Irvin, 2015), which may make it harder to publish a study and contribute to the file drawer problem (Rosenthal, 1979). One open science practice introduced to counter the bias toward publishing studies with large, statistically significant findings is the use of registered reports.

Registered reports take the step of preregistration a step further and integrate it into the peer review process (Chambers et al., 2015; Reich, 2021). The process of preparing a registered report is quite similar whether the study is experimental, quasi-experimental, or even descriptive or qualitative. The key difference from the status quo is that the review process is split into two phases. First, the authors submit the introduction, background, context, and methods sections of an article for review *prior* to conducting their analysis. At this point, reviewers evaluate the research questions, methods, and contribution of the manuscript and decide to accept or reject *before* knowing the direction or magnitude of the results.

In the case of quasi-experimental studies, reviewers can evaluate and give feedback on the proposed research questions, primary hypotheses, primary outcomes, identification strategy, empirical model, model parameters, definition of treatment, principles of sample construction, and variable definitions, among other features of the study design and analytic plan. For quasi-experimental registered reports, sample sizes and units of intervention are usually out of the researchers control, thus decisions related to model parameters, primary versus exploratory hypotheses, and variable definitions are likely to receive the most scrutiny during the first phase of peer review. Reviewers can also provide feedback on the proposed motivation, theory, and framing.

After soliciting this feedback, the editor chooses whether to invite the authors to implement their analysis and submit the study without revision, to invite the authors to respond to the reviewers' critiques (similar to a revise and resubmit), or to reject the manuscript. If there are disagreements or additional justifications, the authors and reviewers might go through a couple of rounds in the first phase of the review–as they might in the standard peer review process.

After the first phase peer review process, if the reviewers decide a manuscript should be published based upon the quality and value of the front-end of the manuscript, the authors receive an "in-principle" acceptance. In the second phase of the process, the reviewers will evaluate the results and discussion sections of the manuscript. If the authors reasonably follow the process they outlined in the first phase, the paper will be published. If there are deviations or major issues that arise, the authors will have to justify any changes for the reviewers. There may be additional rounds of peer review, and the papers can still be rejected. However, the expectation is that registered reports carried out as planned will be published regardless of whether the authors find statistically significant, surprising, or "large" results.

At its core, the practice of registered reports involves reviewers providing feedback and making judgments about the quality of the work while authors can still reasonably implement changes, with neither party knowing how those changes will affect the outcome. In theory, this should reduce the likelihood of publication bias, help solve the file drawer problem, and increase confidence that researcher *and* reviewer decisions were motivated by their good-faith ex-ante beliefs, rather than ex-post considerations.

3.5 OPEN ACCESS ARTICLES AND PREPRINTS

The main goal of social science research is to produce reliable evidence that can guide practice, shape policy, and advance theory. However, access to this evidence is often fundamentally inequitable because so much published research is stored behind paywalls that are inaccessible or unaffordable to vast swaths of the public (Piwowar et al., 2018). The open access movement aims to democratize evidence by removing paywalls and making research available to anyone with an internet connection (see Fleming et al., 2021). Publishing pre-prints and publishing in open access journals are open science practices that are equally compatible with experimental, quasi-experimental, observational, descriptive, theoretical, or qualitative research.

Currently, individual researchers often have limited control over the accessibility of their publications. There are a few things researchers can do, however, to promote equitable access to their research. One avenue is to publish in open access journals or those that allow authors to pay for their article to be open access. There are generally fees associated with publishing open access, but these costs can be offset by universities and funders who are committed to promoting and supporting open science. Section 4 explores the roles of journals and funders in further democratizing access. When journals do not offer open access options or the fees are prohibitively high, authors might consider posting preprints, which are public versions of manuscripts posted on preprint servers or a personal website before peer review.

While some fields have long-standing traditions of posting preprints (e.g., in economics via the NBER Working Paper Series, which started in 1973), this practice is gaining popularity across the social sciences. In psychology, the PsyArXiv preprint repository was established in 2016; in sociology, the SocArXiv preprint repository was established in 2016; in education, the Annenberg Institute's EdWorkingPapers series began in 2019; and in political science, the APSA preprints platform was launched in 2019. Preprint repositories like arXiv (established in 1991) and SSRN (1994) have helped researchers from a wide range of fields disseminate timely research before it is officially published for decades.

Preprints enable immediate access to new research, especially critical in fields where timely evidence can inform pressing decisions. For quasi-experimental studies, which often inform real-world policy and practice, preprints facilitate feedback that can refine methodologies and interpretations. As Fleming and colleagues (2021) note, the thorny issues people raise about preprints (i.e., lack of peer-review; potential risks to blind review) are tractable, and preprints might actually broaden critique, allowing authors to strengthen their work prior to formal publication. Authors can also upload post-prints–accepted, peer-reviewed manuscripts formatted by the author–to ensure accuracy alongside accessibility. In this way, pre- and post-prints help make cutting-edge research readily available and less delayed by lengthy peer review processes.

3.6 MESSAGING OPEN SCIENCE PRACTICES (OR LACK THEREOF)

We recognize that not all quasi-experimental studies will engage in all (or any) open science practices, nor necessarily should they. However, in the spirit of transparency, we propose that readers should be able to easily identify the extent to which a given quasi-experimental study incorporated open science practices (or not). In Section 4, we discuss the growing role journals, funders, and other stakeholders play in incentivizing how researchers message open science practices. However, individual researchers have some control over how they report their results and indicate their use of open science practices.

For instance, authors can easily make it clear whether—and if so, where and how—the data and code can be accessed. Similarly, if a study was preregistered, authors can note that and provide the preregistration in a link or appendix, as is common in experimental studies. Conversely, if a study was not preregistered, authors can also highlight that for readers and recommend that the results should be considered exploratory until they can be replicated.

Of course, exploratory analyses can yield interesting and important findings, and we do not suggest suppressing those findings. Instead, authors can separate their results section into two parts: prespecified hypotheses and exploratory analysis (Gehlbach & Robinson, 2018). This signals that readers should have comparatively more faith in the results from those hypotheses that were prespecified and treat the exploratory findings as hypothesis-generating for future studies. Scholars can also avoid reporting exploratory findings in an abstract or executive summary without noting that they are the result of exploratory analyses rather than tests of a primary hypothesis.

Many ask, what happens when you have to deviate from your preregistration or preanalysis plan? Applied researchers regularly confront unexpected, complex, and nuanced challenges that can require them to modify their approach to evaluating a policy or intervention. Whether the discrepancies are due to an unforeseen issue with the data or a mistake, deviations will necessarily happen and they are not all problematic (Lakens, 2024). Sometimes the deviation is small, and can be easily addressed in the body of the manuscript. Sometimes—and we speak from experience here—the deviations are larger and less straightforward. To facilitate both the review process and to make it easier on the readers, authors of preregistered experimental and quasi-experimental studies can add an appendix that details any deviations from their preregistration or preanalysis plan.

3.7 REPLICATION

In RCTs, open science principles encourage researchers to test whether an experiment's findings generalize by testing the same hypothesis under different conditions or in different samples. In quasi-experimental studies, the details of the policy scenario will determine whether generalizability studies are feasible versus cases where replication should simply be viewed as demonstrating that a study's findings can be reproduced by third parties (e.g., as a class exercise or data validation process). Consider an example: during their study evaluating a tutoring program in Washington, D.C., Lee, Loeb, and Robinson (2024a) stumbled across an interesting (and unexpected) pattern in the data. Their within-student analysis revealed that students were more likely to attend school on days they had a tutoring session scheduled. Given that the finding emerged from an exploratory analysis, the hypothesis and analysis plan was not preregistered. However, because the same tutoring initiative was implemented the following year, the research team was able to preregister the hypothesis and analysis to determine if the result would replicate (Lee et al., 2024b).[1]

Similarly, Heller (2024a) used regression discontinuity methods to estimate the causal relationship between reaching a college readiness benchmark on the GED® test and post-secondary attainment. The quasi-experimental study relied upon extant administrative data and was not preregistered. The findings were imprecise, but suggested that there was no relationship between earning a GED® College Ready designation and college outcomes. Following this analysis, Heller (2024b) collected new administrative data from GED Testing Service, LLC., and preregistered[2] a follow-up quasi-experimental study that would revisit the same question using similar methods in a larger sample (94,000 observations instead of 15,000) to assess whether the null findings would replicate.

Of course, not all replications are as straightforward, nor can all quasi-experimental research be replicated on demand to assess the generalizability of findings. For example, in the absence of a similar mass-migration event, David Card's (1990) difference-in-differences analysis of the Mariel Boatlift on the Miami labor market is unlikely to be replicated in another context, with or without Card providing access to his open-source data and code. However, there are a wide range of quasi-experiments that might be replicated using data from other contexts or time periods to assess the generalizability or sensitivity of results and could benefit from researchers adopting open science practices.

For example, many studies assess the impact of the minimum Pell Grant on educational outcomes and borrowing behavior using a regression discontinuity design (e.g., Marx & Turner, 2018; Park & Scott-Clayton, 2018; Chan & Heller, 2023). On the one hand, the fact that the results of these studies are broadly consistent suggests that the impacts of the minimum Pell

[1] This would have been easier if there was a specific preregistration template for quasi-experimental designs, but in its absence they used the general template.
[2] Ibid.

grant on educational attainment and borrowing outcomes may generalize across contexts. However, each study has idiosyncrasies related to the population studied, sample exclusions, outcomes of interest, variable definitions, model parameters, subgroups of interest, covariates included, time periods covered, etc. that may be difficult to discern in the absence of replication packages that include open source data and code.

Although replication is a cornerstone of the scientific process, replication studies represent a miniscule share of published research in most fields, including psychology (1%; Makel, Plucker & Hagerty, 2012), education (0.13%; Makel & Plucker, 2014), and economics (0.1%; Mueller-Langer et al., 2019). This is, in part, due to the lack of incentives for individual researchers to replicate studies because journal editors and reviewers often prioritize novel findings over demonstrations of generalizability. Despite a growing movement to elevate the status of replication studies in experimental research (e.g., McShane et al., 2019), increasing rates of quasi-experimental replication will require a shift in incentives. External stakeholders can help drive this transition, as we discuss below in Section 4.

3.8 SUMMARY

Ultimately, there is no one-size-fits-all approach for quasi-experimental researchers want to align their work with open science principles. As we describe, the details of a given study may make it more or less compatible with specific practices. In general, the open science movement is not focused on promoting any single practice or bundle of practices, but in fostering norms that ensure transparency, rigor, and credibility are valued above novelty, impact, and incredibility (Mellor, 2022).

Open science practices are not meant to police research, but as disciplinary standards evolve, they have the scope to enhance our collective search for truth. The scope for open science practices to improve applied social science is predicated on honesty and good faith of the overwhelming majority of scholars. Bad actors can and will abuse open science tools—however, practices like replication and open data increase the probability unethical scholars will be discovered and discredited (Kennedy, 2024).

In this section, we described several open science practices and outlined how quasi-experimental researchers can integrate them into their own research. That said, individual researchers can only do so much to change norms and expectations within the quantitative social sciences. In the next section, we discuss how different stakeholders can support and incentivize quasi-experimentalists to adopt open science practices more systematically.

4. Practical advice for other stakeholders

Executing applied research often requires coordinated effort across multiple individuals and organizations. Data providers, journals, funders, and research repositories play important roles in facilitating and encouraging the adoption of open science practices, as well as reducing barriers to integrating open science practices into quasi-experimental research. In this section, we identify opportunities for different types of stakeholders to champion open science through their respective roles in quasi-experimental research.

4.1 DATA PROVIDERS

Unlike experimental studies, which generally must collect data prospectively, quasi-experimental studies can use a variety of data sources: existing data, data currently being collected, or data that will be collected in the future. The data that fuels policy research is often sensitive, private information that requires restrictions to protect human subjects. When data is not publicly accessible for ethical, legal, or business reasons, data providers have considerable influence on nearly all aspects of what, where, when, how, and by whom data is accessed, reported, and shared.

While complex bureaucratic processes and organizational risk aversion can create barriers to knowledge production, the data sharing process also presents opportunities for data providers to facilitate the integration of open science practices into research that relies on the retrospective or prospective analysis of administrative data. Data access is often preceded by a formal or informal proposal process wherein a researcher submits a description of the research questions they plan to answer with a provider's data, how the analysis will be conducted, and the potential benefits of the proposed work. Researchers and their institutions or organizations typically work with data providers to negotiate the terms of a legal agreement that governs how and when the data can be used and to what ends.

During this stage, data providers can exert influence to encourage or discourage parties interested in accessing restricted access data to evaluate a policy, product, or intervention to adopt open science practices. For example, data providers might encourage or even require researchers to preregister their hypothesis and/or submit a preanalysis plan detailing their empirical design and modeling decisions. Similarly, data providers can stipulate conditions under which deidentified microdata (or aggregated data) can be publicly posted to facilitate replication or, as is often the case, explicitly disallow the public sharing of deidentified data.

While there are legal or ethical reasons why some datasets cannot be shared publicly, even in deidentified or aggregated forms, we encourage data providers to think carefully about what *can* be shared with minimal risks to the individuals represented in the data. Similarly, researchers can support data providers in this effort by identifying the minimum viable datasets necessary to replicate their primary results. While posting of replication *code* has become a relatively common practice in quantitative social science, this does little to facilitate replication for the many studies that rely on proprietary data. Data providers and researchers can also work together to archive the specific files (and file-structures) required to use replication code to replicate published work and create transparent, streamlined processes for replicators to apply for access to those specific restricted-use files.

As the gatekeepers of the administrative data that fuels quasi-experimental research, data providers also have considerable scope to influence the exclusivity of data access in the short- and long-term. Relationships between data providers and researchers are built upon trust, which can create a self-reinforcing cycle where the same small groups of researchers obtain and maintain access to restricted-use data from specific providers while others are denied access. These exclusive relationships have benefits to these inner-circle researchers, who face less competition and whose work is more likely to be novel, and data providers, who limit their exposure to risks related to data security or politics by sharing data with trusted partners. However, granting individual researchers or a single team of researchers exclusive permission to conduct a particular quasi-experimental evaluation creates circumstances that are particularly conducive to practices like HARKing and unintentional *p*-hacking.

However, there are several ways data providers can balance the tradeoffs between exclusivity, organizational risks, and research transparency. As mentioned above, requiring or encouraging partners to preregister their hypotheses with a detailed preanalysis plan is the simplest way to promote the goals of open science, with or without broadening data access. Many data providers already require partners to submit proposals that contain all or most of the content of a preregistration and preanalysis plan in order to obtain permission to use administrative data to answer a specific research question (e.g., University of Houston Education Research Center, 2023), so this would not require substantial additional labor in those cases.

Another solution is for data providers to create protocols to facilitate streamlined access to replication files following a study's completion. If scholars know their work is likely to be replicated, this promotes transparency and accountability that is particularly important in cases where preregistration is not possible. However, relying upon ex-post replication as an accountability mechanism is contingent upon other researchers' willingness to conduct replication studies that may not yield new insights.

A third way data providers can promote credibility, replicability, and transparency in quasi-experimental evaluations is by broadening access to restricted-use data in real-time. This is an especially effective strategy when many scholars can reasonably anticipate a natural experiment whose impact assessment is of general interest (e.g., the introduction of a new benefit, the changing of a rule or threshold, or the staggered rollout of a policy). Even by providing data access to just two teams of researchers, data providers can create a self-reinforcing accountability structure that incentivizes caution, diligence, and robustness.

A more ambitious version of this "horse-race" model of knowledge generation is known as the "many analysts" or "crowdsourced science" approach, wherein multiple teams of researchers (sometimes dozens) are given the same data and work independently to answer the same question, with or without methodological constraints, and publish the results collaboratively (Silberzahn et al., 2018; Huntington-Klein et al., 2024). Evaluations using "horse-race" or "many analysts" approaches can also be preregistered and allow researchers to explore the distributional characteristics of estimated effects—and the factors that influence effect-size estimates—alongside the average estimated effects. To facilitate these types of collaborative studies, data providers could designate a lead research team to coordinate across groups of researchers or stipulate guidelines for coordination as a condition of data access. Funders and journals, whose roles are discussed in further detail below, can also play a role in facilitating many analysts projects by supporting and rewarding this type of collaboration as well as funding or publishing replication studies.

Another challenge of implementing open science practices in quasi-experimental research comes from uncertainty regarding the availability, measurement, or structure of key variables. This uncertainty makes it difficult for researchers to confidently prespecify the details of planned analyses based on preexisting administrative data. While researchers can and should deviate from their preanalysis plans when necessary, data providers can enable the adoption of preregistration and preanalysis plans in quasi-experimental research through strategic decisions regarding the data provision process. Typically, data providers may offer data dictionaries or detailed descriptions of datasets, and while these are valuable resources, they are often insufficient to identify important features and idiosyncrasies of large administrative datasets.

We suggest three practices that data providers can adopt to make quasi-experimental preregistration less difficult and more credible: 1) providing access to simulated datasets, 2) offering partitioned datasets that may omit key outcome variables or time periods, and 3) certifying when researchers gained access to restricted-use data. Practices (1) and (2) allow researchers to familiarize themselves with the structure and character of key variables and identify viable empirical strategies without being influenced by what those decisions imply about their quasi-experimental impact estimates in the real data (i.e., without unintentionally or intentionally HARKing). Practice (3) is a simple step data providers can take to increase the credibility of quasi-experimental preregistration.

4.2 FUNDERS

Government agencies, private foundations, non-profit organizations, and corporations contribute billions of dollars to support social science and education research each year (Gibbons, 2023). Funders have tremendous scope to influence the focus and methodology of research, both directly, through the projects they support, and indirectly, through the incentives they create (Hess & Henig, 2015; Feuer, 2016; Sands, 2023; Reikosky, 2024). Relatedly, funders have an opportunity to promote open science practices through their grantmaking requirements. Federal agencies represent the majority of research investment that flows into institutions of higher education (Gibbons, 2023); they can use this power of the purse to shift research norms. Already, federal agencies like the U.S. Department of Education and U.S. Department of Health and Human Services, the major funders of education research (via the Institute for Education Science) and medical and public health research (via the National Institutes of Health) require or encourage open science practices in experimental research (National Institutes of Health, 2016; Institute of Education Sciences, 2022).

While large governmental or institutional funders could develop policies to encourage the adoption of similar practices in quasi-experimental research they fund, even the most outspoken advocates for increasing the credibility and transparency of funded research have yet to establish frameworks to support open science practices like preregistration or replication in quasi-experimental research. This is likely because it is far more complicated to assess what is reasonable to require in quasi-experimental research than in experimental research. Wide variation in the character of quasi-experiments makes it difficult to settle on universal guidelines, but mounting evidence suggests that researcher degrees of freedom threaten the replicability and credibility of quasi-experimental scholarship (Huntington-Klein et al., 2021; 2024; Breznau et al., 2022; Holzmeister et al., 2024; Wuttke et al., 2024). Funders can help nudge quasi-experimental researchers toward adopting the open science practices that are compatible with their research design by treating these practices as the default and requiring clear explanations for deviations from open science best practices.

While some open science practices can be implemented with or without additional financial resources, others are expensive. A small subset of open access journals do not charge authors article processing fees to publish their work (i.e., the so-called "diamond open access journals"), but most do (Fuchs & Sandoval, 2013). Since open access articles are not directly financed by journal subscriptions, these fees must cover the costs of producing and managing the journal. In journals that charge article processing fees, publishing in an open access journal costs roughly five times more than publishing in a paywalled journal (~$2000 vs ~$400, on average), but article processing charges in the most expensive open access journals can exceed $10,000 per article (Grossman & Brembs, 2021; Limaye, 2022; Borrego, 2023).

Funders can help the scholars they support to democratize access to their research by encouraging the publication of preprints and providing resources to cover the high cost of publishing in open access journals or paying any required fees to make a specific article open-access. Additionally, funders can consider directly supporting open access journals that do not charge article processing fees to authors to supplement the volunteer networks and learned societies that support most diamond journals or sponsoring journals that want to transition to a fee-free model.

4.3 JOURNAL EDITORS AND PEER REVIEWERS

As the primary access point for academic scholarship, journals have substantial influence on the proliferation of open science practices along multiple dimensions. Journals can help readers identify scholarship that adheres to open science principles to help readers calibrate confidence in scholarship. Journal editors and peer reviewers can reward authors who adopt open science practices and hold authors accountable to implementing them with fidelity. Editorial boards and publishers can democratize access to knowledge by adopting fee schedules that acknowledge large gaps in resources by country, institution, or career-stage.

Increasingly, peer-reviewed journals are embedding structures to communicate whether an article has certain open science features (Kidwell et al., 2016). Many journals that publish quantitative research ask authors to provide a statement as to whether they are willing and able to publicly post data and code at the point of article submission, and this is published alongside accepted articles. Other journals include icons or badges that appear in searches or on an article's landing page to indicate that the study's data has been published in a repository or that its materials are published online. In otherwise paywalled journals, a badge might indicate which articles are open access as a result of authors paying for their study to not be behind a paywall.

Additionally, a small but growing subset of journals—like the *Journal for Research on Educational Effectiveness* or the *American Journal of Political Science*—offer open science badges to demonstrate, for example, whether a study has been preregistered or is a replication study. Badges can strongly and quickly influence researchers' willingness to share data: one year after the Journal of Psychological Science introduced data sharing badges in 2014, the proportion of articles with open data increased five-fold (Kidwell et al., 2016; Munafò et al., 2017).

In addition to creating ways to signal the adoption of open science practices, journals' peer review process can be a powerful mechanism to reward and maintain the integrity of these practices. For example, journal editors and peer reviewers should subject the methodological choices of quasi-experimental work that is not preregistered to greater scrutiny (e.g., by continuing to great emphasis on the robustness of the results to alternative choices) while also holding authors of preregistered quasi-experimental studies accountable to following their preanalysis plans and documenting the reasons for any deviations. Editors and peer reviewers can help develop and strengthen new disciplinary norms around preregistration and preanalysis plans to fulfill the promise of open science practices as a tool to increase trust in quasi-experimental (and experimental) research and reduce the prevalence of illusory results. Researchers and reviewers stand to benefit from a simplified review process that deemphasizes exhaustive robustness checks and reduces publishing frictions.

Furthermore, editorial boards and publishers should consider whether and how their journal's practices align with Guidelines for Transparency and Openness Promotion (TOP) in Journal Policies and Practices.[3] TOP guidelines provide a framework to guide journal policies and navigate challenges related to data citation; data, materials, and code transparency; study design and analysis; preregistration; and replication. Journals that are not already open-access can consider adopting processes by which scholars can pay to make their specific articles open access, even if most articles remain behind paywalls. Journals that require article processing fees to cover the costs of publishing can consider sliding scale or fee-free publication for scholars from low-income countries as well as graduate students and others who may have limited access to financial resources. Journals that do not accept registered reports can consider whether and how to integrate this option into their existing peer review process for experimental and quasi-experiment research. Fortunately, there are dozens of highly respected publications across the sciences that journal editors can look to for guidance in adopting or strengthening their open science practices.

4.4 REGISTRIES AND REPOSITORIES

The rise of public research registries and data repositories are central to the adoption and proliferation of open science research practices. As more and more researchers preregister studies and post replication packages online, registries and repositories play a larger public-facing role than ever before.

Table 1 summarizes the features of several prominent social science research registries, describing the characteristics that make each registry more or less accommodating to quasi-experimental research. While some registries (e.g., REES) provide resources that are tailored to specific quasi-experimental research methods, others (e.g., OSF, RIDIE, asPredicted) have default templates that can be easily adapted. However, some (e.g., AEA RCT Registry) actively exclude quasi-experimental research.

[3] https://www.cos.io/initiatives/top-guidelines

Table 1: Features of Common Social Science Research Registries

Feature ↓ Repository→	AEA	As Predicted	OSF	REES	RIDIE
Allows QE studies to be registered	No	Yes	Yes	Yes	Yes
Flexible pre-registration template	No	Yes	Yes	Yes	Yes
QE methods can be selected w/in default template	No	No	Kind of	Yes	Yes
Specific QE preregistration templates	No	No	No	Yes	No
Meta-data/tags to identify QE registrations	No	No	Kind of	Yes	Yes
Specific content or geographic limitations	No	No	No	Yes	Yes

Notes: QE = quasi-experimental. AEA RCT Registry: https://www.socialscienceregistry.org/; AsPredicted: https://aspredicted.org/; OSF Registries: https://osf.io/registries; Registry of Efficacy and Effectiveness Studies (REES): https://sreereg.icpsr.umich.edu/sreereg/; Registry for International Development Impact Evaluations (RIDIE): https://ridie.3ieimpact.org/.

We suggest the following steps repositories can take to lower barriers to quasi-experimental preregistration:

1. Allow quasi-experimental studies to be preregistered;
2. Provide adaptable or open-ended templates that are compatible with quasi-experimental research designs
3. Provide templates for specific quasi-experimental designs
4. Add quasi-experimental features to sortable meta-data, badges, etc.

As research norms evolve, registries can support the proliferation of open science practices in quasi-experimental research by making explicit space for quasi-experimental studies in their platforms. Without any substantive changes, registries can add a field to their default registration template to allow quasi-experimental researchers to select the specific quasi-experimental method(s) their evaluation employs.

However, some features of quasi-experimental studies do not translate straightforwardly to the language of experimentalism that dominates most registries. While most registries do allow for quasi-experimental studies to be preregistered, researchers exploring open science practices may find the generic templates misaligned with the unique demands of their work, leading to frustration and early abandonment of these efforts. Creating specific preregistration templates for common quasi-experimental strategies, like those found in RIDIE, is a simple step repositories can take to make their platforms more conducive to preregistering quasi-experimental studies. This sends an important signal to scholars that open science research practices like preregistration are appropriate for their work and may nudge those on the margin of adopting open science practices to do so.

Additionally, accommodating quasi-experimental research designs expands the utility of registries to new audiences, and allows registries to tag quasi-experimental registrations in searches (e.g., to facilitate meta-analyses, replications, or reviews). Data repositories (e.g, Harvard Dataverse; ICPSR; OSF; ResearchBox; Scientific Data) are generally already accessible to quasi-experimental researchers, but may benefit from similar steps to optimize their user interfaces to make quasi-experimental materials easier to publish, find, review, and use.

5. Discussion

Open science principles promote honesty, equity, and accuracy in the social sciences. In experimental research, the replication crisis prompted a paradigm shift in how experimental research is conducted, and most fears about open science stifling research have proven to be unfounded. In quasi-experimental work, open science principles may be even more important. The econometrics of experimental analyses are generally predetermined by the experimental design—experiments are attractive precisely because they reduce otherwise difficult causal questions to testing for differences in means between two groups. In quasi-experimental work, there are many more dimensions of researcher choice, many ex-post reasonable approaches to constructing samples, defining variables, and constructing models, and therefore, many more opportunities for implicit or explicit biases to influence results. Enduring shifts in disciplinary norms will require coordinated commitments to open science principles among all stakeholders in quasi-experimental research. From researchers to data providers to funders to journal editors and peer reviewers, everyone has a role to play in promoting transparency and enhancing the credibility of quantitative social science.

6. References

Angrist, J. D., & Pischke, J. S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, *24*(2), 3-30.

Biden, J.R. (2024). Remarks by President Biden in State of the Union Address. [Speech Transcript]. Retrieved November 12, 2024 from https://www.whitehouse.gov/briefing-room/speeches-remarks/2024/03/08/remarks-by-president-biden-in-state-of-the-union-address-3/.

Bill & Melinda Gates Foundation. Evaluation Policy. Retrieved November 12, 2024 from https://www.gatesfoundation.org/about/policies-and-resources/evaluation-policy.

Borrego, Á. (2023). Article processing charges for open access journal publishing: A review. *Learned Publishing*, *36*(3), 359-378.

Breznau, N., Rinke, E. M., Wuttke, A., Nguyen, H. H., Adem, M., Adriaans, J., ... & Van Assche, J. (2022). Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. *Proceedings of the National Academy of Sciences*, *119*(44), e2203150119.

Brodeur, A., Cook, N. M., Hartley, J. S., & Heyes, A. (2024). Do Preregistration and Preanalysis Plans Reduce p-Hacking and Publication Bias? Evidence from 15,992 Test Statistics and Suggestions for Improvement. *Journal of Political Economy Microeconomics*, *2*(3), 527-561.

Card, D. (1990). The impact of the Mariel boatlift on the Miami labor market. *Ilr Review*, *43*(2), 245-257.

Chambers, C. D., Dienes, Z., McIntosh, R. D., Rotshtein, P., & Willmes, K. (2015). Registered reports: realigning incentives in scientific publishing. *Cortex*, *66*(3), A1-A2.

Chan, M., & Heller, B. H. (2023). When Pell Today Doesn't Mean Pell Tomorrow: Evaluating Aid Programs With Dynamic Eligibility. EdWorkingPaper no.23-851.

Christensen, G., Wang, Z., Levy Paluck, E., Swanson, N., Birke, D., Miguel, E., & Littman, R. (2020). Open science practices are on the rise: The state of social science (3S) survey.

Council of Economic Advisors, Executive Office of the President. (2014). *Economic Report of the President*. Washington, DC: Government Printing Office.

Council of Economic Advisors, Executive Office of the President. (2018). *Economic Report of the President*. Washington, DC: Government Printing Office.

Council of Economic Advisors, Executive Office of the President. (2022). *Economic Report of* *the President*. Washington, DC: Government Printing Office.

Currie, J., Kleven, H., & Zwiers, E. (2020). Technology and big data are changing economics: Mining text to track methods. *AEA Papers and Proceedings*, *110*, 42-48.

Feuer, M. J. (2016). *The Rising Price of Objectivity: Philanthropy, Government, and the Future of Education Research*. Harvard Education Press. 8 Story Street First Floor, Cambridge, MA 02138.

Field, S. M., Wagenmakers, E. J., Kiers, H. A., Hoekstra, R., Ernst, A. F., & van Ravenzwaaij, D. (2020). The effect of preregistration on trust in empirical research findings: Results of a registered report. Royal Society Open Science, 7(4), 181351.

Fleming, J. I., Wilson, S. E., Hart, S. A., Therrien, W. J., & Cook, B. G. (2021). Open accessibility in education research: Enhancing the credibility, equity, impact, and efficiency of research. *Educational Psychologist*, *56*(2), 110-121.

Fuchs, C., & Sandoval, M. (2013). The diamond model of open access publishing: Why policy makers, scholars, universities, libraries, labour unions and the publishing world need to take non-commercial, non-profit open access serious. *TripleC: Communication, capitalism & critique*, *11*(2), 428-443.

Gehlbach, H., & Robinson, C. D. (2018). Mitigating illusory results through preregistration in education. *Journal of Research on Educational Effectiveness*, *11*(2), 296-315.

Gehlbach, H., & Robinson, C. D. (2021). From old school to open science: The implications of new research norms for educational psychology and beyond. *Educational Psychologist*, *56*(2), 79-89.

Gibbons, M.T. (2023). *R&D Expenditures at U.S. Universities Increased by $8 Billion in FY 2022.* National Science Foundation, National Center for Science and Engineering Statistics. NSF 24-307. Retrieved October 30, 2024 from https://ncses.nsf.gov/pubs/nsf24307.

Goldsmith-Pinkham, P. (2024). Tracking the Credibility Revolution across Fields. arXiv preprint arXiv:2405.20604.

Grossmann, A., & Brembs, B. (2021). Current market rates for scholarly publishing services. *F1000Research*, *10*.

Hardwicke, T. E., & Wagenmakers, E. J. (2023). Reducing bias, increasing transparency and calibrating confidence with preregistration. *Nature Human Behaviour*, *7*(1), 15-26.

Heller, B. H. (2024a). GED® College Readiness Benchmarks and Post-Secondary Success. EdWorkingPaper no.24-914.

Heller, B.H. (2024b). High School Equivalency Credentialing and Post-Secondary Success. OSF Registries. Retrieved November 13, 2024 from: https://osf.io/72hm6.

Hess, F. M., & Henig, J. R. (Eds.). (2015). *The new education philanthropy: Politics, policy, and reform*. Cambridge, MA: Harvard Education Press.

Holzmeister, F., Johannesson, M., Böhm, R., Dreber, A., Huber, J., & Kirchler, M. (2024). Heterogeneity in effect size estimates. *Proceedings of the National Academy of Sciences*, 121(32), e2403490121.

Huntington-Klein, N., Arenas, A., Beam, E., Bertoni, M., Bloem, J. R., Burli, P., ... & Stopnitzky, Y. (2021). The influence of hidden researcher decisions in applied microeconomics. *Economic Inquiry*, *59*(3), 944-960.

Huntington-Klein, N. and 145 others. (2024). The Sources of Researcher Variation in Economics. Unpublished Manuscript.

Imbens, G. W. (2024). Causal inference in the social sciences. *Annual Review of Statistics and Its Application*, *11*.

Institute of Education Sciences. (2022). Standards for Excellence in Education Research. Retrieved November 13, 2024 from: https://ies.ed.gov/seer/.

Kaplan R.M. & Irvin V.L. (2015) Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time. PLoS ONE 10(8): e0132382. Retrieved November 13, 2024 from: https://doi.org/10.1371/journal.pone.0132382

Kennedy, J. E. (2024). Addressing researcher fraud: retrospective, real-time, and preventive strategies–including legal points and data management that prevents fraud. Frontiers in Research Metrics and Analytics, 9, 1397649.

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L. S., ... & Nosek, B. A. (2016). Badges to acknowledge open practices: A simple, low-cost, effective method for increasing transparency. PLoS biology, 14(5), e1002456.

Klonsky, E. D. (2024). Campbell's Law Explains the Replication Crisis: Pre-Registration Badges Are History Repeating. *Assessment*, 10731911241253430.

Lakens, D. (2024). When and How to Deviate from a Preregistration. *Collabra: Psychology*, *10*(1).

LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review*, *76*(4), 604-620.

Lee, M., Loeb, S., & Robinson, C. D. (2024a). Effects of High-Impact

Tutoring on Student Attendance: Evidence from the OSSE HIT Initiative in the District of Columbia. EdWorkingPaper. Annenberg Institute at Brown University.

Lee, M., Loeb, S., & Robinson, C. D. (2024b). Year 2 of Effects of High-Impact Tutoring on Student Attendance: Evidence from the OSSE HIT Initiative in the District of Columbia. Retrieved from osf.io/uw36q

Limaye, A. M. (2022). Article Processing Charges may not be sustainable for academic researchers. *MIT Science Policy Review 3*, 17-20.

Logg, J. M., & Dorison, C. A. (2021). Pre-registration: Weighing costs and benefits for researchers. *Organizational Behavior and Human Decision Processes*, *167*, 18-27.

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur?. *Perspectives on Psychological Science*, *7*(6), 537-542.

Makel, M. C., & Plucker, J. A. (2014). Facts Are More Important Than Novelty: Replication in the Education Sciences. *Educational Researcher*, *43*(6), 304-316.

Marx, B. M., & Turner, L. J. (2018). Borrowing trouble? Human capital investment with opt-in costs and implications for the effectiveness of grant aid. *American Economic Journal: Applied Economics*, *10*(2), 163-201.

McAdams, D. P., & McLean, K. C. (2013). Narrative identity. *Current directions in psychological science*, *22*(3), 233-238.

McShane, B. B., Tackett, J. L., Böckenholt, U., & Gelman, A. (2019). Large-Scale Replication Projects in Contemporary Psychological Research. *The American Statistician*, *73*(sup1), 99–105. https://doi.org/10.1080/00031305.2018.1505655

Mellor, D. (2021). Improving norms in research culture to incentivize transparency and rigor. *Educational Psychologist*, *56*(2), 122–131. https://doi-org.stanford.idm.oclc.org/10.1080/00461520.2021.1902329

Miguel, E. (2021). Evidence on research transparency in economics. *Journal of Economic Perspectives*, *35*(3), 193-214.

Mueller-Langer, F., Fecher, B., Harhoff, D., & Wagner, G. G. (2019). Replication studies in economics—How many and which papers are chosen for replication, and why?. *Research Policy*, *48*(1), 62-83.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie du Sert, N., ... & Ioannidis, J. (2017). A manifesto for reproducible science. *Nature human behaviour*, *1*(1), 1-9.

National Institutes of Health. (2016). NIH Policy on the Dissemination of NIH-Funded Clinical

Trial Information. Retrieved November 12, 2024 from: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-149.html.

Neal, Z. P. (2022). A quick guide to sharing research data & materials. https://doi.org/10.31219/osf.io/9mu7r

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. Proceedings of the National Academy of Sciences, 115(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114

Obama, B.H. (2014). Remarks of President Barack Obama – State of the Union Address As Delivered. [Speech Transcript]. Retrieved November 12, 2024 from: https://obamawhitehouse.archives.gov/the-press-office/2016/01/12/remarks-president-barack-obama-%E2%80%93-prepared-delivery-state-union-address.

Özler, B. (2019). Registering studies when all you want is a little more credibility. *Development Impact.* World Bank Blogs. Retrieved on August 30, 2024 from: https://blogs.worldbank.org/en/impactevaluations/registering-studies-when-all-you-want-little-more-credibility

Park, R. S. E., & Scott-Clayton, J. (2018). The impact of Pell Grant eligibility on community college students' financial aid packages, labor supply, and academic outcomes. *Educational Evaluation and Policy Analysis*, *40*(4), 557-585.

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., Farley, A., West, J., & Haustein, S. (2018). The state of OA: A large-scale analysis of the prevalence and impact of open access articles. *PeerJ Journal Peerj*, *6*, e4375.

Reich, J. (2021). Preregistration and registered reports. *Educational Psychologist*, *56*(2), 101-109.

Reikosky, N. (2024). For (Y) our Future: Plutocracy and the Vocationalization of Education (Doctoral dissertation, University of Pennsylvania).

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological bulletin, 86*(3), 638.

Sands, S. R. (2023). Institutional Change and the Rise of an Ecosystem Model in Education Philanthropy. *Educational Policy*, *37*(6), 1511-1544.

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., ... & Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science, 1*(3), 337-356.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological science*, *22*(11), 1359-1366.

University of Houston Education Research Center. (2023). Policies & Procedures: General Information. Retrieved November 15, 2024 from: https://uh.edu/education/research/institutes-centers/erc/proposal-preparation-and-submission/uherc-general-information_rev1_feb2023.pdf.

van den Akker, O. R., van Assen, M. A., Bakker, M., Elsherif, M., Wong, T. K., & Wicherts, J. M. (2024). Preregistration in practice: A comparison of preregistered and non-preregistered studies in psychology. *Behavior Research Methods*, *56*(6), 5424-5433.

The White House. Economic Report of the President. Retrieved November 12, 2024 from: https://www.whitehouse.gov/cea/economic-report-of-the-president/.

Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions.(I). Psychological review, 8(3), 247.

Wuttke, A., Freitag, K., Kiemes, L., Biester, L., Binder, P., Buitkamp, B., … Samarskiy, A. (2024). Observing many students using difference-in-differences designs on the same data and hypothesis reveals a hidden universe of uncertainty. SocArXiv Papers. Retrieved November 15, 2024 from: https://doi.org/10.31235/osf.io/j7nc8.