Journal of Economic Perspectives—Volume 35, Number 3—Summer 2021—Pages 193–214

# Evidence on Research Transparency in Economics

Edward Miguel

Open science and research transparency (terms I'll use interchangeably) are advanced when scientific claims are independently verifiable, including through the promotion of free and open sharing of the process of conducting research, and when the content and findings generated during research are objects that readers can check for themselves. A decade ago, "research transparency" and "open science" were not on the radar screen of most economists or other social scientists. However, a new scholarly movement has coalesced around bringing new open-science practices, tools, and norms into the mainstream. Prominent social science organizations have taken the field, including the Center for Open Science (cos.io), the Society for the Improvement of Psychological Science (improvingpsych.org), and the Berkeley Initiative for Transparency in the Social Sciences (bitss.org). The goal of this article is to lay out the emerging evidence on the adoption of these approaches in three specific areas—open data, pre-registration and pre-analysis plans, and journal policies—and more tentatively, to begin to assess their impacts on the quality and credibility of economics research.

In a broad normative perspective, the open science movement seeks to bring the research practices and culture of economics in line with a classical "scientific ethos" of open inquiry that goes back centuries. In one prominent discussion, Merton (1942) laid out the four so-called "Mertonian norms" of scientific inquiry: universalism, communality, disinterestedness, and organized skepticism. In the

■ *Edward Miguel is the Oxfam Professor of Environmental and Resource Economics, University of California Berkeley, Berkeley, California. He is also a Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts. His email address is emiguel@berkeley.edu.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.35.3.193.

social sciences, one aspect of these norms is the basic ability of others in the scholarly community to reproduce published findings, and thus, to understand fully how they were produced and what alternative analyses might be possible. In economics, relatively few papers—in the range of one-third to one-half of published articles (Chang and Li 2015; Gertler, Galiani, and Romero 2018)—achieve even the basic goal of computational reproducibility, which involves using the publicly shared materials to produce the results in the paper. Much of this has to do with the quality of the replication data and code, which are often incomplete or poorly documented.

In addition, the movement for open science and research transparency seeks to reduce the extent to which investigator bias or other biases can creep into both research practices and publication decisions. There is ample evidence of pervasive problems in empirical economics research (and in related quantitative fields like political science and social psychology). I won't present a comprehensive account of these concerns here—for more detail, readers can refer to Christensen and Miguel (2018) and the Christensen, Freese, and Miguel (2019) book—but it is worth offering some concrete examples.

One common pattern is that if a study produces a "null result" that is not different from zero at a conventional level of statistical significance, typically 5 percent, it is less likely to be written up by a social science researcher and less likely to be published if it is written up (for example, Franco, Malhotra, and Simonovitz 2014). High proportions of null results "disappearing" have been documented in well-designed studies in economics (Andrews and Kasy 2019) and other fields (Turner et al. 2008; Simonsohn, Nelson, and Simons 2014). When null results disappear from public view, these unwritten findings are effectively lost to the broader research community. The result is that looking at published research may lead to misleading conclusions regarding topics of intellectual and public policy importance (Ioannidis 2005). The disappearance of nulls also wastes research funding and human resources by producing efforts that are never published and leads to duplicated efforts when other scholars carry out work that (unbeknownst to them) was already tried earlier.

This issue of the disappearing nulls is closely related to *publication bias*, and the related concern about selective presentation of results. If authors believe that results need to attain an arbitrary level of statistical significance to be meaningful or publishable, they will have an incentive to manipulate their research accordingly, in what is sometimes called *p*-hacking or phishing. Empirical economists have traditionally engaged in lengthy periods of largely unstructured data-mining, in which potentially thousands of statistical tests were run, but they then only report their handful of "preferred" estimates in the final manuscript (Leamer 1983). Such choices lead almost unavoidably to cherry-picking of results to obtain p-values below 0.05 and inflated statistical significance levels. Brodeur et al. (2016) and Brodeur, Cook, and Heyes (2020) demonstrate in leading economics outlets, including top-five journals, that there are substantial spikes in empirical estimates barely above the statistical significance level of 5 percent, with apparent "canyons" just below this value—a clear sign that the published research was pre-selected by authors or journals to meet this standard. Gerber and Malhotra (2008a, 2008b) show a similar pattern of empirical results in leading political science and sociology journals.

In the discussion below, I present evidence indicating that economics is in a period of rapid transition toward new transparency norms in the areas of open data, preregistration and pre-analysis plans, and journal policies.[1] I will argue that there are indications of at least some social benefits from these practices. There is also growing reason to believe that critics' worst fears regarding their potential costs—like onerous adoption costs or the stifling of creativity—have not been realized. I close by arguing that further cultural change is needed to reinforce and sustain the changes that are already underway in economics.

## Open Data

When I was a graduate student in the 1990s, obtaining the underlying data and analysis scripts for a published paper was typically either challenging or impossible. But dramatic changes in technology—especially the rise of the internet over the past 25 years—and in policies of academic journals and professional associations have led rapid shifts in prevailing practices.

In economics, one catalyst for these changes was the data-sharing policy adopted by the American Economic Association (AEA) in 2005, which came in response to growing evidence that many, if not most, published empirical analysis in economics could not be readily reproduced (Dewald, Thursby, and Anderson 1986; Bernanke 2004). The policy led to an almost immediate increase in the posting of data and analysis code for the *American Economic Review* (Christensen, Dafoe et al. 2019). Other leading general interest journals and field journals followed suit, with many adopting the AEA policy verbatim (Christensen and Miguel 2018). This has led to a dramatic increase in access to data for published research in our discipline in a relatively short span of time.

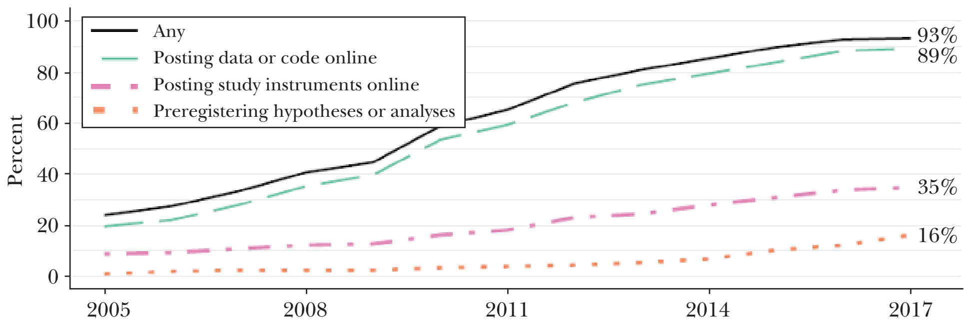
Figure 1 illustrates the rise of data sharing in economics, based on an attempt to obtain a representative survey sample of scholars who had published in top-ten economics journals during 2014–2016 (as well as in other social science fields, although here we discuss economics).[2] We achieved a respectable response rate of nearly 50 percent.[3] While relatively few of the economists in our sample had shared data circa 2005 (just before the AEA policy was adopted), by 2017 nearly 90 percent

[1] Another open science practice that has gained traction in economics is *disclosure*, including the 2012 AEA policy mandating that authors disclose personal, professional, and financial conflicts of interest (https://www.aeaweb.org/journals/policies/disclosure-policy). While I briefly mention the importance of disclosure below, I do not focus on it here: for additional discussion, see Miguel et al. (2014). A further research transparency practice that has long been important in other fields, especially medical research (through the CONSORT guidelines, Begg et al. 1996), is the establishment of author *reporting guidelines*. While promising, this practice has not yet caught on in economics, and again, I do not discuss it here.

[2] Christensen, Wang et al. (2019) present data across disciplines. Patterns over time are similar in economics and political science. Psychology is ahead of economics in the adoption of preregistration but behind economics in data sharing, while sociology has the slowest adoption on all measures. Differences in research methods by field (for example, the prevalence of experimental versus non-experimental studies, and quantitative versus qualitative approaches) appear to explain many of the gaps.

[3] We also restrict the analysis in the figure to scholars who had received their PhD before 2009 (N = 204), meaning they had been active researchers long enough for the time series of their research practices to be meaningful. We also surveyed PhD students, but do not report their behaviors in the figure.

*Figure 1*
**Research Transparency Practices Are Rising in Economics**




*Source:* From Swanson ⓡ al. (2020, Figure 1, panel A).
*Note:* The chart shows for a given year the proportion of published authors who report having first completed an open science practice in that year or previously. The solid black line shows the proportion of published authors who had completed any open science practice by that year. The dashed green line shows the proportion who had posted data or code online by that year. The dash-dotted purple line shows the proportion of published authors who had posted study instruments online by that year. The dotted orange line shows the proportion who had preregistered an analysis or hypothesis by that year. Posting study instruments online is the response to the question, *"Approximately when was the first time you publicly posted study instruments online?"* Posting data or code online is the response to the question, *"Approximately when was the first time you publicly posted data or code online?"* Preregistering hypotheses or analyses is the response to the question, *"Approximately when was the first time you preregistered hypotheses or analyses in advance of a study?"* The sample is restricted to published authors who completed their PhDs by 2009 (N = 204).

had publicly shared data at least once. There was a similar trend, although with lower levels, for sharing study instruments such as field surveys or lab protocols; there is less adoption of preregistration and pre-analysis plans (which we discuss in the next section) although it too is rising over time.

Many economics journals now directly host data and code on their own website, while there is also growing popularity of internet data repositories, including well-known outlets such as the Harvard Dataverse and Inter-university Consortium for Political and Social Research (ICPSR). These have been so rapidly successful that it is easy to forget what an important innovation the professional curation, storage, and management of research data and code has been.

Yet journal data-sharing policies are not a panacea. The threat of replication alone may provide only weak incentives for academic integrity from scholars who do not anticipate their study will garner extensive interest or citations. Replication materials for published papers are too often poorly documented and disorganized. Partly in response to such concerns, in 2019 the American Economic Association adopted a new and more ambitious set of data sharing standards: the updated AEA Data and Code Availability Policy can be found at https://www.aeaweb.org/journals/policies/data-code. Some key additions are the requirement that data and code be submitted to AEA journals before final paper acceptance and that it be posted on a data repository (openICPSR, rather than at a journal website). There is also an expanded role for the AEA Data Editor and associated staff, who carry out pre-publication verification of analytical results whenever feasible (Vilhuber 2020), where the Data Editor's role is to assess computational reproducibility, not to judge the appropriateness of the underlying econometric choices. Across AEA journals, the Data Editor's team carried out 216 pre-publication assessments across 138 papers between July and November 2019, and none had "fundamental flaws" (Vilhuber, Turrito, and Welch 2020), meaning any issues found were communicated to the authors and resolved before publication.

It is worth noting several limitations of the AEA data requirements and of open data policies in general. First, not all data can be accessed by the AEA team—for instance, if it is proprietary or subject to government confidentiality restrictions. In these cases, one option is for verification of results by a third-party replicator who has access to the data used in the author's paper; for details, see https://aeadataeditor.github.io/aea-de-guidance/protocol-3rd-party-replication.html. Second, in most cases the data that is shared through the AEA process is not the detailed micro-data, but rather, an aggregated and processed file. Posting the raw underlying data (wherever possible) would generate additional social value for the broader research community, and moves to encourage this would be a useful direction for future open data reforms by the AEA and other economics associations and journals.

However, by conditioning final paper acceptance on including relatively high-quality replication materials across the range of AEA journals, the Association has raised the bar for the field and brings economics close to what is considered "best practice" across other scientific disciplines as captured in the Transparency and Openness Promotion (TOP) Guidelines (Nosek et al. 2015). It seems likely to me that other economics journals will eventually follow suit, as they did after the adoption of the 2005 AEA guidelines. According to the TOP guidelines, which capture open data as well as other practices discussed below, the AEA journals currently rank as the most compliant with open science standards (among the 50 most cited economics journals), with TOP scores similar to leading general scientific journals like *Nature* and *Science* (Bogdanoski and Stillman 2021).

An impressive 97 percent of economists express support for data sharing in the Swanson ⓡ al. (2020) survey data—although respondent beliefs about their colleagues' support for research transparency practices is consistently below respondents' stated support. While most economists believe the rise in data sharing has been worth it, it is still valuable to assess costs, private benefits, and social benefits.

On the cost side, the time it takes to assemble replication materials for a forthcoming article could take anywhere from several hours to a few weeks of work, based on my own personal experience and anecdotally. The variation across projects is related to the size and complexity of the underlying dataset, of course, but is also greatly affected by whether the data, code, and documentation materials are put together along the way as a project is being conducted, or if one needs to assemble them after analysis has been completed. Creating documented data and code materials from a project completed years earlier can be particularly time-consuming and difficult. However, the shift to a new norm of (nearly) universal sharing of data and code means that economists today know they will need to share these materials with other scholars going forward (for publication in a prestigious AEA journal, say), giving them every incentive to comment generously on their code, label variables clearly, write README files, and generally keep materials organized along the way.

There is some quantitative evidence on the time costs of preparing data and code materials. Since 2016, Innovations for Poverty Action (IPA), a development economics research organization, has funded staff who are tasked with preparing replication materials for field data collection projects that they had supported, and they recorded the time it took (IPA 2020).[4] Across 65 project datasets, the average amount of time to prepare replication materials for public sharing was 31.5 hours, with an interquartile range of 10.0 to 40.5 hours (and a 10th to 90th percentile range of 5.8 to 80.2 hours). This is non-trivial for most projects: still, remember that this estimate of preparation time applies to field experiments that often require multiple years of work on collecting data, so it remains a very small share of overall project work time.

A frequently discussed concern is that enhanced data- and code-sharing requirements will be particularly onerous for scholars who lack the resources to hire a dedicated research assistant, thus exacerbating existing inequalities among researchers. For scholars at an early career stage or who are not working in resource-rich institutions, including many in low- and middle-income countries, these 31.5 hours of work (on average) to assemble data and code for posting will need to be carried out on nights and weekends, given already heavy faculty teaching loads and administrative responsibilities. A promising solution could be for more research funders to dedicate resources to making data and code publicly available, such as efforts recently carried out by IPA, the Berkeley Initiative for Transparency in the Social Sciences (BITSS), and the Jameel Poverty Action Lab (J-PAL). Beyond providing a fairer playing field for all economists, expanded funding for dataset and code preparation would help align private and social incentives for the creation of these research public goods.

The most immediate private benefit that I and many other scholars have personally experienced from new open data norms is the fact that our own research data is better organized *for ourselves* and thus easier to reuse for other analyses and papers as a result of the effort that is put into improved documentation (like the README files and other replication materials). Many scholars (myself included) have often procrastinated on assembling data documentation materials and doing the final grunt work needed to get a dataset ready for sharing with other scholars. Journal policies that force one to do this to get your paper to final acceptance and publication do focus the mind.

Another private benefit from public data sharing is the possibility that it will lead to further related work by others, and thus to greater citations and influence. The likelihood that data-sharing will generate citations has been enhanced by the policies of the AEA, other journals, and nearly all data repositories to provide digital object identifiers (DOIs) for posted research datasets. A number of scholars have shown that data sharing at the article level is correlated with higher citations for that

[4] I am grateful to Hasina Badani of the IPA Research Transparency, Data Governance and Ethics Team for generously sharing this data.

paper (for instance, Piwowar, Day, and Fridsma 2007; Piwowar and Vision 2013), although there remain obvious omitted variable concerns associated with the nonrandomness of the data-sharing decision.

In Christensen, Dafoe et al. (2019), we attempt to address the possible selection into data-sharing using the 2005 AEA data-sharing policy as a natural experiment. In particular, we compare papers published in the *AER* versus the *Quarterly Journal of Economics (QJE)* for four years before and after the 2005 policy change. The availability of data and code for *AER* articles increased quickly, while rates of data availability at the *QJE* (which did not adopt a comparable data policy until 2016) remained low in our study period. In addition, average article citations (through November 2017) rose roughly 50 percent for articles published in the *AER* after the policy change. These results should be viewed as provocative, rather than definitive, given the sample of two journals. Yet the possibility that posting data and code generates higher citations would create strong private incentives to support open data and may help to explain why open data has quickly become a strongly held norm in economics.

One possible social benefit is that open data may enable other scholars to uncover research fraud more readily: for example, discoveries of fraud in political science and social psychology were enabled by journal open data policies. In one vivid example, Broockman, Kalla, and Aranow (2015) downloaded replication data and code from the website of *Science* and discovered statistical anomalies, including too little variation in key measures, which they correctly concluded were consistent with the data having been generated by a random number generator rather than collected in the field. The strength of open data norms in economics—which emerged several years before other social sciences (Christensen, Wang et al. 2019)—could partially explain why there have been fewer high-profile instances of research fraud in our discipline in recent years.

Perhaps the most widely discussed potential social benefit of open data is the opportunity it provides for other scholars to gain a deeper understanding of the research and to build from it. For example, a reanalysis can consider the robustness of findings: that is, do the findings change substantially with modest changes to the specification or research approach? Replication can apply the same research method to a different dataset. More broadly, earlier results can be extended by looking at the results of variations in the underlying model, data over a longer time period, and so on. In this way, embracing research transparency can also be a step toward a fairer and more inclusive scholarly community. I believe that making replication of empirical analysis the norm will have major scientific benefits for economics in the long-run. Indeed, as Maximilian Kasy explains in his paper in this symposium, findings of later replications can allow other scholars to quantify the extent of publication bias in economics, together with associated econometric approaches to correct for it.

However, researchers' growing ability to access data and code from previous studies has led to some controversy (for discussion, see Christensen and Miguel 2018). Gertler, Galiani, and Romero (2018) note that there may be "overturn bias," in which reanalysis and replications that contradict an originally published paper are seen as more publishable. It will be important to address this incentive, in part by making sure that "successful" replications are also published, not just those that claim to debunk earlier findings. Other replicating authors may also be motivated by personal or monetary conflicts of interest—for instance, if their work is funded by a research sponsor with a financial stake in the answer to the question at hand, such as a pharmaceutical firm or energy company—making strong conflict-of-interest disclosure requirements even more essential.

But ultimately, in my view, many of the concerns around replications from greater openness of data and code are growing pains due to the fact that we are in a transitional phase between an earlier era (like my grad school days) when data was rarely available and replications seldom carried out, and a not-too-distant future when reproducibility of results and other data checks like those in the AEA journals will have become *de rigueur* for both researchers and journals.

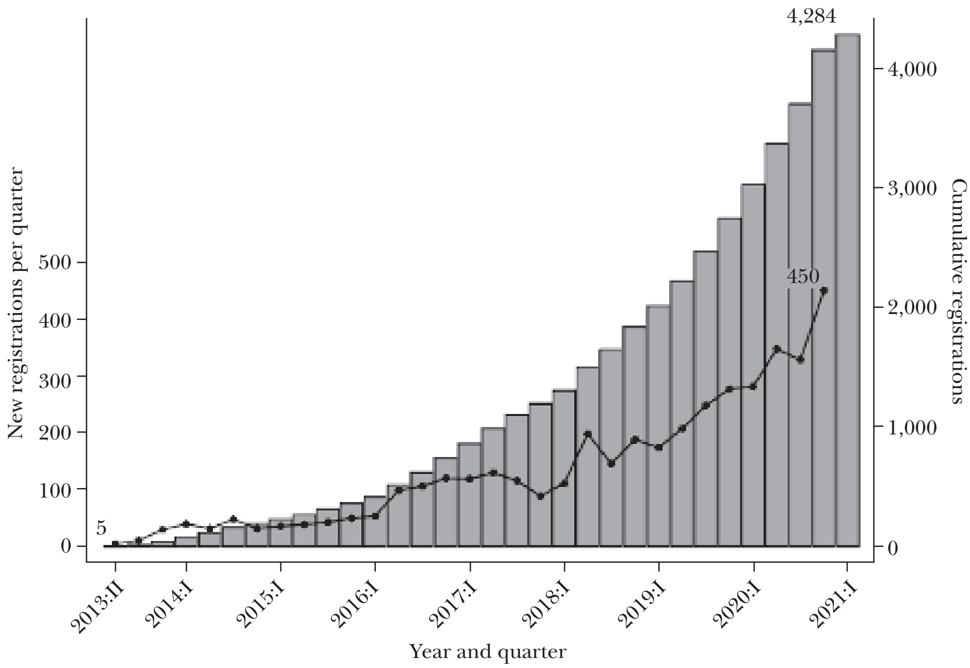
Stepping back, perhaps the most important lingering concern about the expansion of data-sharing requirements is the potential for reduced incentives to collect new data. As Christensen Freese, and Miguel (2019) note: "Data are the lifeblood of empirical science, and it would be a perverse consequence of a data-sharing policy if it reduced the amount of important data collected." There is clearly a need to balance these incentives for the generation of new data versus the social gains of greater sharing of such data, and to do so approaches like temporary "data embargoes" (similar in spirit to technology patents) could be useful. Continuing the current norm of time-limited monopoly rights over the use of data that scholars have generated themselves could be essential to incentivize researchers to carry out ambitious future data collection projects. More thought and debate are still needed regarding how to strike the right balance between these competing concerns to craft the most effective data-sharing policies in economics; in doing so, it will be useful to learn from the experiences of other scientific fields (Hill, Stein, and Williams 2020).

## Preregistration and Pre-analysis Plans

Among the open science innovations that have taken place in economics over the last 15 years, the creation of a study registry and growing use of pre-analysis plans is arguably the biggest break with previous research practices. Since its founding in 2013, the AEA Randomized Control Trial Registry has seen exponential growth; by January 2021, over 4,200 studies were registered, as shown in Figure 2. The registry asks for basic study characteristics, like the where, when, and what of the data, approval of an Institutional Review Board, and a few other items. Since 2017, 45 percent of newly registered prospective studies have also posted a pre-analysis plan, with additional more detailed step-by-step description of just how the analysis will be carried out.[5] Similar changes are underway in other social sciences: in political science, the Experiments in Government and Politics (EGAP) registry is widely

[5] This statistic is based on publicly available data downloaded from the AEA registry (https://www.socialscienceregistry.org/) on January 31, 2021; see https://doi.org/10.7910/DVN/FUO7FC.

*Figure 2*
**Studies Posted over Time, American Economic Association Randomized Controlled Trial Registry**




*Source:* This figure was produced by Garret Christensen, Edward Miguel, and Sarah Stillman, and is in the public domain at https://doi.org/10.7910/DVN/FUO7FC. Cumulative and new registrations of studies (by quarter) on the AEA Registry for Randomized Controlled Trials. Data downloaded on January 31, 2021 from https://www.socialscienceregistry.org/. The quarterly figure is not shown for Quarter 1 of 2021 (since data is only available to date for the first month of that quarter).

used; in psychology, most scholars register either on the Open Science Framework (OSF) or on AsPredicted.

Views towards preregistration and pre-analysis plans are generally positive in economics, but with some doubts. The Swanson ⓡ al. (2020) survey data indicate that slightly over half of economists support these practices (with many expressing indifference); in development economics, the subfield where study registration and pre-analysis plans first took off, stated support is far higher at 80 percent. My goal in this piece is not to rehash the ongoing debates about potential benefits of adopting preregistration, and whether they justify the up-front costs. For an overview of these debates, the reader can turn to Olken (2015) and Coffman and Niederle (2015): for additional views, recent starting points are Christensen, Freese, and Miguel (2019); Duflo et al. (2020); and Abrams, Libgober, and List (2020). Rather I will briefly sketch the parameters of the existing debate and then devote my attention to newly emerging evidence on the real-world practice and implementation.

The case for preregistration and pre-analysis plans comes in a few flavors. First, a registry creates a "paper trail," which can help scholars working in an area to learn about each other's ongoing work. Second, preregistration and pre-analysis plans generate accountability: the rest of the research community (including journal referees) can see which questions the scholars initially intended to ask and this can help reduce publication bias by increasing the reporting of all results, including null results. The option on the AEA registry to keep pre-analysis plans temporarily private (before a paper with the results is released, for instance) reduces concerns that other scholars will troll the registry to "scoop" particularly innovative ideas. Third, a pre-analysis plan can reduce pressure on researchers to emphasize a certain subset of results that may be favored by government officials, research funders, or even colleagues. Finally, an underappreciated benefit of preregistration and a pre-analysis, in my view, is that it improves the quality of the research by pushing scholars to think more carefully about their design and data beforehand. I return to this point below.

There are also potential costs. First, Olken (2015) mentions the time costs, which in turn will depend on the level of detail needed beyond the basic study characteristics demanded in the AEA registry (as discussed in Duflo et al. 2020). Second, authors fear that pre-imposed constraints on their analytical work may produce an end-product that is less creative and interesting—and possibly less publishable. However, this second concern seems overstated to me. A norm has quickly emerged in economics that allows—and even encourages—authors to present additional analyses to whatever was prespecified, with the caveat being that authors must transparently report what was and wasn't in their plan. Indeed, the first two papers published in economics that employed a pre-analysis plan (to my knowledge), namely, Finkelstein et al. (2012) and Casey, Glennerster, and Miguel (2012), both describe why they felt it was also necessary to publish some analyses that went beyond their pre-analysis plans, and they clearly label these results as such.

While we do not yet know for sure what registration will do in economics in the long-run given how recently the AEA registry was set up, we can learn from the experience of other fields. In particular, the rise of randomized control trials in economics was preceded by the growth of medical trials, and the creation of the AEA registry in economics was directly inspired by ClinicalTrials.gov, which was set up in 2000.

Several benefits have been documented in clinical trial research from having a registry. First, it has become possible to assess how published papers deviate from original plans. A number of studies have audited these deviations in medical studies (such as Mathieu et al. 2009), something that could easily be adopted in economics to immediately provide a greater level of accountability and make sure fewer results disappear.

Second, the creation of a clinical trials registry in medical research appears to lead to more reported null results. In Figure 3, reproduced from Kaplan and Irvin (2015), each dot represents a study on nutritional supplements funded by the same funding arm of the National Institutes of Health; the fact that all were chosen for funding provides some degree of study comparability and quality control. The vertical line marks the founding of ClinicalTrials.gov in 2000. The pre-post research design here is obviously not ideal, but the pattern is striking. Before the registry, the majority of published results were statistically significant and showed benefits, many

*Figure 3*
**Relative Risk of Treatment by Publication Year**


*Source:* Reproduced from Kaplan and Irvin (2015), Figure 1 (Creative Commons Attribution, CC BY, license).
*Note:* Data are from large NHLBI trials on pharmaceutical and dietary supplement interventions. Positive trials are indicated by the plus signs while trials showing harm are indicated by a diagonal line within a circle. Prior to 2000 when trials were not registered in clinical trials.gov, there was substantial variability in outcome. Following the imposition of the requirement that trials preregister in clinical trials.gov, the relative risk on primary outcomes showed considerably less variability around 1.0.

with large effect estimates. After ClinicalTrials.gov was set up and medical journals began requiring study registration as a publication requirement (De Angelis et al. 2004), far more null results appeared in the literature, and in fact, hardly any significant positive results showed up.

In the decades before 2000, there were repeated scandals in medical research involving clinical trials funded by self-interested pharmaceutical companies, often accompanied by some evidence that "null" trial results that would have hurt these firms' bottom lines systematically went unreported (Turner et al. 2008). The existence of the trial registry combined with journal requirements to preregister makes this much harder to do, making the clinical trial literature more credible.

Could similar benefits emerge in economics? In economics (and political science), the most detailed evidence to date on the real-world use and impacts of preregistration and pre-analysis plans comes from two papers by Ofosu and Posner (2020a, 2020b). Ofosu and Posner (2020a) examine all working papers released by the National Bureau of Economic Research from 2011 to 2018 and searched for all that used experimental (field and lab) research methods—because these methods are most likely to preregister and write pre-analysis plans. They then search among these working papers for those that also mention a pre-analysis plan. In all, 8.4 percent of experimental working papers during this period mention the existence of an associated pre-analysis plan, with rates rising over time. Ofosu and Posner then determine which of these papers were ultimately published (and where), and through web searches gather total citation counts on Google Scholar as of 2019. They ask whether experimental papers in economics that used a pre-analysis plan have different publication and citation trajectories than those that did not. Of course, adoption of pre-analysis plans was not randomly allocated, but they argue that their focus on the subfield of papers using experimental methods and the fact that all are written by NBER affiliates means they are not comparing apples to oranges. That said, the authors emphasize that the results should be treated as "suggestive" and as a "snapshot."

Ofosu and Posner (2020a) find that the overall likelihood of being published is somewhat lower for studies with a pre-analysis plan (44 percent) compared to those without one (at 54 percent, though this difference is not significant at traditional levels). However, studies with pre-analysis plans are more than twice as likely to have been published in "top-five" economics general interest journals than others (27 percent versus 12 percent).[6] Studies associated with a pre-analysis plan also have 60 percent higher citations by 2019. The authors do not provide a definitive answer for why studies with pre-analysis plans receive more citations and are published in more prestigious journals. One possibility is that perhaps stronger researchers tended to adopt pre-analysis plans sooner or did so for their most promising projects. Another possibility is that studies with pre-analysis plans that obtain a null finding might find it easier to be accepted for journal publication: for example, the first two pre-analysis plan papers in 2012 both contained results that could be seen as "disappointing" or went against some scholars' priors but were still both published in a leading journal (Casey, Glennerster, and Miguel 2012; Finkelstein et al. 2012). Finally, perhaps the process of writing a pre-analysis plan improves the research, leading to stronger papers that are easier to publish in leading venues.

Their second study (Ofosu and Posner 2020b) builds on a novel survey conducted among scholars in economics and political science who belonged to networks specializing in experimental research—and were thus likely to have registered pre-analysis plans—regarding their experiences, practices, and beliefs. They

[6] The authors define the top-five journals in the usual way: *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*.

also review the content in a representative subset of 195 registered pre-analysis plans. The survey has some limitations. It has a relatively low response rate among those contacted (at 23 percent). Also, it focuses on pre-analysis plans written through 2016, which places this data in the early days of preregistration. Norms may have evolved considerably since then. Still, their data remains among the best available sources of quantitative information (to my knowledge) on real-world use of pre-analysis plans.

For example, the modal time to write a pre-analysis plan is two to four weeks of work time among the survey respondents, a figure that resonates with my own experience. Most survey respondents also mention that this time is not all additive, because it is much faster to move directly into analysis mode if you have already spent weeks carefully laying out the regressions that will be run and thinking through how to avoid certain pitfalls. In fact, 33 percent of respondents indicate "that these [time] savings were equal to or greater than the time spent to draft the PAP in the first place"(Ofosu and Posner 2020b). That said, a pre-analysis plan may impose larger time costs on some scholars, perhaps because some research is intrinsically more complex, or because some researchers tend to write quite detailed pre-analysis plans (myself included) while others focus on a tighter subset of analytical issues (as discussed in Duflo et al. 2020).

The survey evidence also suggests potential quality benefits to writing pre-analysis plans: "An overwhelming majority (8 in 10) said that drafting a PAP caused them to discover things about their project that led to refinements in their research protocols and/or data analysis plans" (Ofosu and Posner 2020b). Ofosu and Posner (2020b) advocate finding ways to harness this potential advantage of the pre-analysis plans by getting early feedback on the research plan before registering it. Indeed, pre-analysis plans are already starting to be incorporated as a normal research product to present in some venues, including the Working Group in African Political Economy (WGAPE) meetings.[7]

Finally, Ofosu and Posner (2020b) assess registered pre-analysis plans along four dimensions: "specifying a clear hypothesis; specifying the primary dependent and independent/treatment variable(s) sufficiently clearly so as to prevent post hoc adjustments; and spelling out the precise statistical model to be tested." Here the record is mixed. In their sample, 90 percent of pre-analysis plans state a clear hypothesis and 80 percent contain at least three of the four elements. However, many of the resulting papers report results that were not in the original pre-analysis plan without always clearly labeling them as such. It remains possible that this situation has improved since their data from 2016, but updated research could document how the use of pre-analysis plans has evolved over time.

Abrams, Libgober, and List (2020) carry out a related audit of the pre-analysis plans listed on the AEA registry. They point out that norms regarding registration vary considerably even across experimental fields, with high rates among economists conducting field experiments but far lower levels among those carrying out lab experiments. They also provide a set of useful reform proposals, including possibly mandating registration before projects are carried out, greater incentives

[7] Dan Posner and I have co-organized WGAPE meetings together with several colleagues since 2002.

to post results generated by the research, and posting materials from Institutional Review Boards.

I cannot claim to have a final answer on whether the benefits of pre-analysis plans exceed their costs, although it seems clear that the more dire predictions from the early days of the AEA registry regarding onerous time costs and stifled creativity have not been borne out. When Ofosu and Posner (2020b) ask directly, 64 percent of scholars respond that "[writing a PAP] takes a considerable amount of time, but it is worth it," while 6 percent write that "It doesn't take much time, so the cost is low," meaning that 70 percent of researchers actively working in this area are largely positive about the benefit to cost ratio. This lines up with the 80 percent of development economists (surveyed in Swanson ⓡ al 2020) who support preregistration.

My sense as a co-author, colleague referee, and adviser is that there is still considerable variation in the style of pre-analysis plans that economists are writing: some are more detailed, others less, some contain more literature review or conceptual discussion, some don't, and so on. My own view is even a relatively sparse pre-analysis plan that lays out the primary outcomes, the core analysis, and main regression specifications remains useful in addressing the most extreme forms of selective reporting and data mining as well as publication bias. Other leading economics associations, including the European Economic Association and the Econometric Society, have moved partially in the same direction regarding registration, and "encourage authors of papers that use RCTs [randomized control trials] to register their experiments" but do not (yet) mandate it.[8]

Preregistration to date has largely been utilized in fields that employ experimental methods, including applied microeconomics fields (especially development economics) and experimental economics. Preregistration and pre-analysis plans have made far less headway in structural econometric work, including in industrial organization, international trade, and macroeconomics. Preregistration appears to be more challenging to implement in structural work, where underlying theoretical models are often more complicated and their construction and estimation involves myriad judgement calls that may be challenging to anticipate and specify in advance—and also more difficult for outside observers to discern. The resulting increase in researcher degrees of freedom likely makes it harder to detect biased reporting. One immediate way forward in these fields—albeit partial—would be for at least some steps of the research process to be prespecified, for instance, the value of particular parameters (like the intertemporal discount rate) to be used in quantitative exercises, or the specific dataset to be analyzed. In the absence of preregistration, a broader discussion is needed in these fields regarding whether there are alternatives that could enhance transparency and similarly constrain *p*-hacking, lest we witness a growing methodological breach across economics subfields over time.

[8] These policies (as of May 31, 2020) are at https://www.eeassoc.org/index.php?site=JEEA&page=42 and https://www.econometricsociety.org/publications/econometrica/information-authors/instructions-submitting-articles, respectively.

## Journal Policies and Practices

Journal policies and practices are influential in setting norms in any scientific field. Here, I will assess two policy changes related to open science issues that have recently been implemented at high-profile economics journals: *pre-results review* and *editorial statements*. Behind both policies is the notion that economic research should be judged by authors and journals based on whether the project was worth undertaking in the first place.

Specifically, the idea behind pre-results review is that referees and editors should ideally judge the quality of a research paper based on its design, data, and the importance of the underlying question, rather than being influenced by whether the results are surprising, well-suited for a press release, statistically significant, or confirm (or contradict) prevailing theory. This approach has become more common in other social science fields, especially psychology and cognitive science, where papers published using this approach are often called "registered reports."

One immediate objection to pre-results review might be that scholars lack the capability to evaluate submitted articles without seeing the results. However, scholars evaluate research proposals that lack results all the time: for instance, when sitting on National Science Foundation or National Institute of Health panels that review grant proposals, when deciding which graduate student travel awards to fund, or when serving on a dissertation prospectus committee. Growing familiarity with pre-analysis plans also facilitates pre-results review.

With pre-results review, an empirical article goes through two stages of review During the first stage, authors submit a "proposal," usually similar to a pre-analysis plan, though with more emphasis on the existing literature and discussion of the project's conceptual or theoretical contributions. Referees review this proposal, and the editor may engage in some back-and-forth with the submitting author, similar to the revise-and-resubmit process in a regular article submission. If the editor decides that the study is valuable and meets the journal's quality bar, it is awarded an "in-principle acceptance," similar to a conditional acceptance. The authors then analyze their data, write up results, and submit the full paper for stage two review.

In the second stage, the full paper is submitted with results, interpretation, and any extensions beyond the original plan (which are acceptable as long as they are clearly delineated). The key idea behind pre-results review is that the journal has committed to publishing the paper as long as the results are presented credibly, the interpretation is reasonable, and there were no major data problems along the way (which would drop the paper below the journal's standard for publication). For instance, if you tried to carry out a study in a country that then experienced a civil war or a natural disaster and you were unable to collect most data, the editor might decide the in-principle acceptance was no longer valid. But if the endline data looks to be of sufficiently high quality and the interpretation given to results is sensible, then the journal is committing to publishing the final paper even if the results are not statistically significant, challenge conventional wisdom, are surprising, or do not seem to "hang together" with a single unambiguous theoretical interpretation.

Virtually no social science journals used pre-results review in 2013, but the numbers have risen quickly with approximately 100 journals accepting "registered reports" in 2018 and 277 journals today (Hardwicke and Ioannidis 2018).[9] In economics, the most prominent example of pre-results review is in the *Journal of Development Economics* (under editors Andrew Foster and Dean Karlan and with support from BITSS), starting in May 2018 (Foster et al. 2019). The *JDE* was a natural venue for a pilot given the already widespread use of pre-analysis plans in development economics, and to my knowledge, it is the first economics journal to adopt pre-results review as a standard article submission format.[10] As of January 2021, roughly two-and-a-half years in, the *JDE* had received 90 submissions for pre-results review, with a rising rate over time. Of these, 18 have received in-principle acceptance and three have been accepted in stage 2 and are now forthcoming in the journal, while the others are either undergoing stage-2 review, still assembling their data, carrying out analysis, or writing up the paper.

As part of the pre-results review adoption process at the *Journal of Development Economics*, a BITSS staff member (Aleksandar Bogdanoski) carried out phone interviews with 12 submitting authors to gain a qualitative sense of how pre-results review is being perceived (Foster et al. 2019). The interviews indicate that, despite being slightly different than regular articles, the refereeing process for pre-results review submissions went smoothly overall and no major red flags were raised, in part perhaps because detailed explanatory materials had been prepared in advance (for authors and referees), as well as a suggested template for the proposals. The most commonly cited benefit, by far, was that writing the proposal for peer review forced authors to think through their research design more carefully, and feedback from the referees at that early stage helped to further improve it. Another pattern was that junior scholars—particularly those who are going on the job market or up for tenure promotion—appreciate the ability to obtain an in-principle acceptance for a project that has not yet been completed (BITSS 2020).

Open questions remain regarding how pre-results review might work in subfields other than development economics. One other economics journal, *Experimental Economics*, has launched a pre-results pilot. The recent rise of alternative article formats in economics, inspired by the short format approach pioneered by *American Economic Review: Insights*, may facilitate acceptance of other novel approaches like pre-results review.

A distinct and lighter-touch change in journal policy are editorial statements to make a particular issue salient. In 2015, the editors of eight leading health economics journals issued an editorial statement emphasizing the importance of publishing

[9] This data is as of February 21, 2021. Up to date information on the adoption of pre-results review and registered reports can be found at https://www.cos.io/rr.

[10] The earliest pre-results analysis and pre-results review in economics (to my knowledge) is Neumark (2001), based on a one-off attempt to implement a pre-results review process at one economics journal in the 1990s. In 1996, there were heated minimum wage debates between Card and Krueger and Neumark and colleagues. According to my Berkeley colleague David Levine, who was the editor of *Industrial Relations* at the time, the late Alan Krueger had the idea in 1996 for various participants in the minimum-wage literature to pre-specify their analysis before the next federal wage increase, and as editor of *Industrial Relations*, Levine would commit to publish results (Levine 2001). Levine believes the idea originated from Danny Kahnemann, who in the 1980s and 1990s developed what he called "adversarial collaboration" with colleagues who disagreed with him, and with whom he worked together to design lab experiments and write articles. Christensen, Freese, and Miguel (2019) contains a more detailed discussion.

null results. They sent letters to referees reminding them to judge papers based on design and quality, not on whether the results are statistically significant.

Such a statement may seem like a small step, but it clearly encouraged a shift of norms. Blanco-Perez and Brodeur (2020) compare the share of published null results during 2014–2018 in the eight health economics journals to two applied microeconomics journals with no similar editorial statement. Figure 4 presents their data from the pre-period, the period when the editorial statement was implemented, and the post-period. The light gray line captures the share of papers presenting statistically significant results (at the 5 percent level) in the control journals and the dark gray line captures this proportion in the journals affected by the editorial statement. Pre-statement, roughly 50 percent of articles have null results in both the control and treatment journals, but there is a sharp rise in the publication of null results after the 2015 statement, with the share of null results increasing by 18 percentage points. This is due to some combination of changes in both editor and referee behavior; Blanco-Perez and Brodeur (2020) do not find meaningful changes in the characteristics of the papers submitted by authors to these journals over the study period.

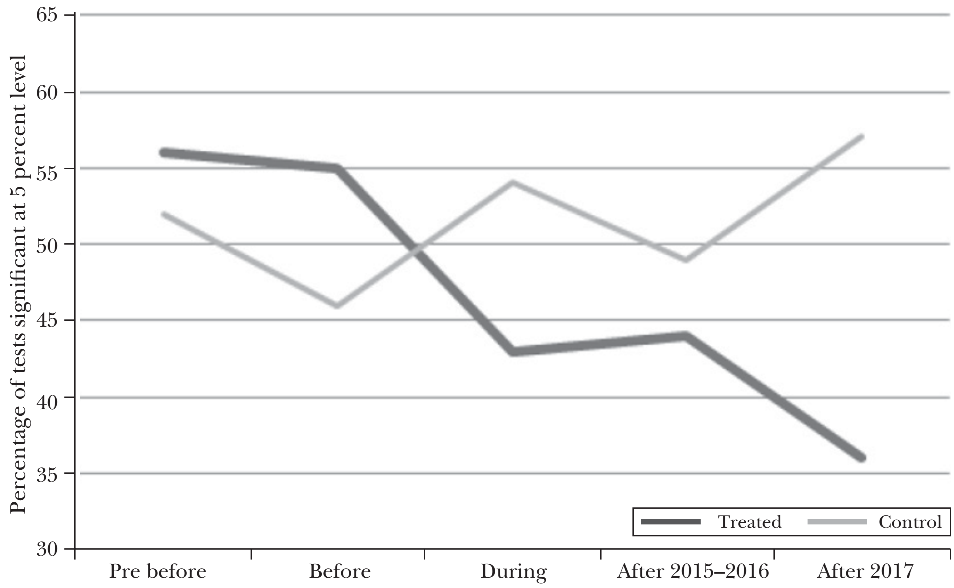
Of course, one can raise questions concerning the possibility of other changes that were occurring in these journals, or in the field of health economics, over time. Yet this evidence suggests that even simple and low-cost actions by editors might help promote changes in norms, even for something as deeply engrained as the bias in favor of publishing significant results. It seems worth considering similar editorial statements (with associated referee reminder letters) by other economics journals on the subject of null results, and perhaps on other open-science issues as well.

## Looking Forward

The past two decades have seen rapid changes in policies and practices to promote open science in economics. Policies that were largely foreign to the discipline of economics when I was in graduate school—sharing of data and code, study registration and pre-analysis plans, and conflict-of-interest disclosure statements—are now routine parts of economists' workflows. Opening up the research process in economics promises to make our research more credible while also potentially promoting a more inclusive scholarly community. However, while some of the underlying problems of publication bias, specification search, and tendentious reporting may have receded, they have not yet gone away (Andrews and Kasy 2019; Brodeur, Cook, and Heyes 2020). In this essay, I have already mentioned some promising areas to enhance research transparency in economics. Here, I mention a few more.

In the area of pre-analysis plans, Laitin et al.'s (2020) plan to Report All Results Efficiently (RARE) proposes to make it standard practice for authors to post all results related to their pre-analysis plans on public study registries, in a so-called "pre-analysis plans report," even if those finding never make their way into a published paper (a proposal related to some ideas developed in parallel in Duflo et al. 2020 and Abrams, Libgober, and List 2020). This step would allow searches of study registries to yield far more complete evidence on work that has been carried

*Figure 4*
**Journal Editorial Statements and the Publication of Significant Results**




*Source:* From Blanco-Perez and Brodeur (2020, Figure 3).
*Note:* Treated journals include *Journal of Health Economics*, *European Journal of Health Economics*, *Health Economics*, *Health Economics Review*, and *International Journal of Health Economics and Management*. Control journals include *Journal of Public Economics* and *Labour Economics*. Percentage of tests significant at the 5 percent level by categories. "Pre Before the editorial" category includes papers that were published one year before the category "Before." "Before the editorial" category includes papers that were submitted and published before the statement on negative findings. "During the editorial" category includes papers that were submitted before the statement on negative findings, but published after. The "After the editorial" categories include papers submitted and published (respectively in 2015–2016 and 2017) after the statement on negative findings. Reproduced with permission from Abel Brodeur.

out on a certain topic to date, leading to improved meta-analysis as well as more informed choices for scholars launching new projects.

Another set of steps would seek to integrate preregistration approaches into some non-experimental research. For example, it might be possible to preregister studies of observational data in a way where it is possible to verify that the pre-analysis plan truly preceded the data analysis (for a discussion of this issue in medical research, see Dal-Ré et al. 2014). One can imagine a preregistration approach for studies that will be conducted after a particular event has occurred (such as an election or data release) or more generally before scholars have been granted access to restricted data (Burlig 2018; Christensen, Freese, and Miguel 2019). Ofosu and Posner (2020b) find that roughly 4 percent of pre-analysis plans that they reviewed were for observational data: in fact, among some studies discussed earlier, both Blanco-Perez and Brodeur (2020) about changes in journal editorial policies and Christensen, Dafoe et al. (2019) about impact of data-sharing were preregistered observational studies. The path to realistically utilizing preregistration for a substantial share of observational nonprospective studies is uncertain, but remains a critical direction for future debate and innovation.

A cluster of other work is actively enriching preregistration in various ways, including by studies that compare effects of treatments with expert forecasts (Della Vigna and Pope 2018; Della Vigna, Pope, and Vivalt 2019), preregistering plans for split-sample analysis (Fafchamps and Labonne 2016; Anderson and Magruder 2017); or using a pre-analysis plan to guide the application of machine learning tools (Ludwig, Mullainathan, and Spiess 2019).

New ideas are also emerging about how to make reproducibility work better in economics: Lars Vilhuber (the Data Editor for the American Economic Association) is leading an effort in collaboration with the Berkeley Initiative for Transparency in the Social Sciences with the aim of Accelerating Computational Reproducibility in Economics (ACRE https://www.socialsciencereproduction.org/). The goal is a crowd-sourced platform to assemble and organize replication activities (which are often carried out today as graduate course assignments) in a systematic way, so that it can become possible to move away from black-or-white takes on whether a finding "replicates," and to illuminate the nuances involved in verifying empirical results. There is also a concrete proposal for how to bring more research transparency tools into public policy, termed Open Policy Analysis (Hoces de la Guardia, Grant, and Miguel 2018), which involves taking a specific policy analysis (say, Congressional Budget Office analysis of the effects of the minimum wage) and then fully specifying how the result was reached in an open-source online document that any member of the public can access.

Even as these open science tools expand in scope and influence, I think more work will also need to be done to change the culture and the mindset of the economics research community. In my opinion, economists should encourage ourselves, our colleagues, and our students to work on important problems without worrying so much about whether the results turn out to be immediately exciting: after all, if scholars are collecting good data and applying thoughtful methods while working on an important problem, even null results are meaningful. We should stress that all research conducted in this way contributes to the broader social goal of generating facts and learning about the world.

■ *I am grateful to Sarah Stillman and Simon Zhu for excellent research assistance. This article would not have been possible without my earlier collaborations with Aleks Bogdanoski, Kate Casey, Garret Christensen, Josh Cohen, Allan Dafoe, Andrew Foster, Jeremy Freese, Rachel Glennerster, Sean Grant, Fernando Hoces de la Guardia, Katie Hoeberling, Dean Karlan, David Laitin, Temina Madon, Don Moore, Betsy Levy Paluck, Andrew Rose and others, much of which is cited here. I also benefited from conversations with other colleagues at the Berkeley Initiative for Transparency in the Social Sciences, including Carson Christiano, Guillaume Kroll, Kelsey Mulcahy, Jen Sturdy, and Alex Wais. Gordon Hanson, Enrico Moretti, Timothy Taylor, and Heidi Williams provided useful suggestions. Some of the arguments here also feature in my 2019 NBER Summer Institute Methods Lecture: https://www.nber.org/econometrics_minicourse_2019/. All errors are my own.*

## References


**Abrams, Eliot, Jonathan Libgober, and John List.** 2020. "Research Registries: Facts, Myths, and Possible Improvements." Artefactual Field Experiments Working Paper 00703.

**Anderson, Michael L., and Jeremy Magruder.** 2017. "Split-Sample Strategies for Avoiding False Discoveries." NBER Working Paper 23544.

**Andrews, Isaiah, and Maximilian Kasy.** 2019. "Identification of and Correction for Publication Bias." *American Economic Review* 109 (8): 2766–94.

**Begg, Colin, Midred Cho, Susan Eastwood, Richard Horton, David Moher, Ingram Olkin, Roy Pitkin, et al.** 1996. "Improving the Quality of Reporting of Randomized Controlled Trials: The CONSORT Statement." *Journal of the American Medical Association* 276 (8): 637–39.

**Berkeley Initiative for Transparency in the Social Sciences (BITSS). (**2020) "Exit Interviews with Authors of Stage 1 Proposals at the *Journal of Development Economics*." Unpublished data (accessed June 4, 2020).

**Bernanke, Ben S.** 2004. "Editorial Statement." *American Economic Review* 94 (1): 404.

**Blanco-Perez, Cristina, and Abel Brodeur.** 2020. "Publication Bias and Editorial Statement on Negative Findings." *Economic Journal* 130 (629): 1226–47.

**Bogdanoski, Aleksandar, and Sarah Stillman.** 2021. "Transparency and Openness Promotion (TOP) Factor Scores of Leading Economics Journals." Harvard Dataverse. https://doi.org/10.7910/DVN/JINXLP. (accessed February 8, 2021).

**Brodeur, A., N. Cook, and A. Heyes.** 2020. "Methods Matter: P-Hacking and Publication Bias in Causal Analysis in Economics." Unpublished.

**Brodeur, Abel, Mathias Lé, Marc Sangnier, and Yanos Zylberberg.** 2016. "Star Wars: The Empirics Strike Back." *American Economic Journal: Applied Economics* 8 (1): 1–32.

**Broockman, David, Joshua Kalla, and Peter M. Aranow.** 2015. "Irregularities in LaCour (2014)." https://osf.io/preprints/metaarxiv/qy2se/.

**Burlig, Fiona.** 2018. "Improving Transparency in Observational Social Science Research: A Pre-analysis Plan Approach." *Economic Letters* 168: 56–60.

**Casey, Katherine, Rachel Glennerster, and Edward Miguel.** 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-analysis Plan." *Quarterly Journal of Economics* 127 (4): 1755–1812.

**Chang, Andrew C., and Phillip Li.** 2015. "Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say 'Usually Not'." Federal Reserve Board Finance and Economics Discussion Paper 2015-083.

**Christensen, Garret, Allan Dafoe, Edward Miguel, Don A. Moore, Andrew K. Rose.** 2019. "A Study of the Impact of Data Sharing on Article Citations Using Journal Policies as a Natural Experiment." *PLoS ONE* 14 (12).

**Christensen, Garret, Jeremy Freese and Edward Miguel.** 2019. *Transparent and Reproducible Social Science Research: How to Do Open Science.* Oakland, CA: University of California Press.

**Christensen, Garret, and Edward Miguel.** 2018. "Transparency, Reproducibility, and the Credibility of Economics Research." *Journal of Economic Literature* 56 (3): 920–80.

**Christensen, Garret, Zenan Wang, Elizabeth Levy Paluck, Nicholas Swanson, David Birke, Edward Miguel, and Rebecca Littman.** 2019. "Open Science Practices Are on the Rise: The State of Social Science (3S) Survey." *MetArXiv.* doi:10.31222/osf.io/5rksu.

**Coffman, Lucas C., and Muriel Niederle.** 2015. "Pre-analysis Plans Have Limited Upside, Especially Where Replications Are Feasible." *Journal of Economic Perspectives* 29 (3): 81–98.

**Dal-Ré, Rafael, John P. Ioannidis, Michael B. Bracken, Patricia A. Buffler, An-Wen Chan, Eduardo L. Franco, Carlo La Vecchia, and Elisabete Weiderpass.** "Making Prospective Registration of Observational Research a Reality." *Science Translational Medicine* 6 (224).

**De Angelis, Catherine D., Jeffrey M. Drazen, Frank A. Frizelle, Charlotte Haug, John Hoey, Richard Horton, Sheldon Kotzen, Christine Laine, et al.** 2004. "Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors." *New England Journal of Medicine* 351 (12): 1250–51.

**DellaVigna, Stefano, and Devin Pope.** 2018. "Predicting Experimental Results: Who Knows What?" *Journal of Political Economy* 126 (6): 2410–56.

**DellaVigna, Stefano, Devin Pope, and Eva Vivalt.** 2019. "Predict Science to Improve Science." *Science* 366 (6464): 428–29.

**Dewald, William G., Jerry G. Thursby, and Richard G. Anderson.** 1986. "Replication in Empirical Economics: The Journal of Money, Credit and Banking Project." *American Economic Review* 76 (4): 587–603.

**Duflo, Esther, Abhijit Banerjee, Amy Finkelstein, Lawrence F. Katz, Benjamin A. Olken, and Anja Sautmann.** 2020. “In Praise of Moderation: Suggestions for the Scope and Use of Pre-analysis Plans for RCTs in Economics.” NBER Working Paper 26993.

**Fafchamps, Marcel, and Julien Labonne.** 2016. “Using Split Samples to Improve Inference about Causal Effects.” NBER Working Paper 21842.

**Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group.** 2012. “The Oregon Health Insurance Experiment: Evidence from the First Year.” *Quarterly Journal of Economics* 127 (3): 1057–1106.

**Foster, Andrew, Dean Karlan, Edward Miguel and Aleksandar Bogdanoski.** 2019. “Pre-results Review at the Journal of Development Economics: Lessons Learned So Far.” *Development Impact*, July 15. https://blogs.worldbank.org/impactevaluations/pre-results-review-journal-development-economics-lessons-learned-so-far.

**Franco, Annie, Neil Malhotra, and Gabor Simonovits.** 2014. “Publication Bias in the Social Sciences: Unlocking the File Drawer.” *Science* 345 (6203): 1502–05.

**Gerber, Alan, and Neil Malhotra.** 2008a. “Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Political Science Journals.” *Quarterly Journal of Political Science* 3: 313–26.

**Gerber, Alan S., and Neil Malhotra.** 2008b. “Publication Bias in Empirical Sociological Research: Do Arbitrary Significance Levels Distort Published Results?” *Sociological Methods & Research.* 37 (1): 3–30.

**Gertler, Paul, Sebastian Galiani, and Mauricio Romero.** 2018. “How to Make Replication the Norm.” *Nature* 554: 417–19.

**Hardwicke, Tom E., and John P.A. Ioannidis.** 2018. “Mapping the Universe of Registered Reports.” *Nature Human Behaviour* 2: 793–96.

**Hill, Ryan, Carolyn Stein, and Heidi Williams.** 2020. “Internalizing Externalities: Designing Effective Data Policies.” *AEA Papers and Proceedings* 110: 49–54.

**Hoces de la Guardia, Fernando, Sean Grant, and Edward Miguel.** 2018. “A Framework for Open Policy Analysis.” *MetaArXiv.* doi:10.31222/osf.io/jnyqh.

**Innovations for Poverty Action (IPA). (**2020) “Operational metrics of the Research Transparency, Data Governance and Ethics Team,” Unpublished data. (accessed June 2, 2020).

**Ioannidis, John P. A.** 2005. “Why Most Published Research Findings Are False.” *PLoS Medicine* 2 (8).

**Kaplan Robert M., and Veronica L. Irvin.** 2015. “Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time.” *PLoS ONE* 10 (8): e0132382

**Laitin, David, Edward Miguel, Ala Alrababhah, Aleksandar Bogdanoski, Sean Grant, Katherine Hoeberling, Cecilia Mo, Don Moore, Simine Vazire, Jeremy Weinstein, Scott Williamson.** 2020. “Unlocking the File Drawer: A RARE Proposal to Normalize Complete Reporting.” Unpublished.

**Leamer, Edward E.** 1983. “Let’s Take the Con out of Econometrics.” *American Economic Review* 73 (1): 31–43.

**Levine, David I.** 2001. “Editor’s introduction to ‘The Unemployment Effects of Minimum Wages: Evidence from a Prespecified Research Design’.” *Industrial Relations* 40 (2): 161–62.

**Ludwig, Jens, Sendhil Mullainathan, and Jann Spiess.** 2019. “Augmenting Pre-Analysis Plans with Machine Learning.” Unpublished.

**Mathieu, Sylvain, Isabelle Boutron, David Moher, Douglas G. Altman, and Philippe Ravaud.** 2009. “Comparison of Registered and Published Primary Outcomes in Randomized Controlled Trials.” *Journal of the American Medical Association* 302 (9): 977–84.

**Merton, Robert K.** 1942. “A Note on Science and Democracy.” *Journal of Legal and Political Sociology* 1: 115–126.

**Miguel, E., C. Camerer, K. Casey, J. Cohen, K.M. Esterling, A. Gerber, R. Glennerster et al.** 2014. “Promoting Transparency in Social Science Research.” *Science* 343 (6166): 30–31.

**Neumark, David.** 2001. “The Employment Effects of Minimum Wages: Evidence from a Prespecified Research Design the Employment Effects of MinimumWages.” *Industrial Relations* 40 (1): 121–44.

**Nosek, B.A., G. Alter, G.C. Banks, D. Borsboom, S.D. Bowman, S.J. Breckler, S. Buck et al.** 2015. “Promoting an Open Research Culture.” *Science* 348 (6242): 1422–25.

**Ofosu, George K., and Daniel N. Posner.** 2020a. “Do Pre-analysis Plans Hamper Publication?” *AEA Papers and Proceedings* 110: 70–74.

**Ofosu, George K., and Daniel N. Posner.** Forthcoming. “Pre-analysis Plans: An Early Stocktaking.” *Perspectives on Politics.*

**Olken, Benjamin A.** 2015. “Promises and Perils of Pre-analysis Plans.” *Journal of Economic Perspectives* 29 (3): 61–80.

**Piwowar, Heather A., Roger S. Day, and Douglas B. Fridsma.** 2007. “Sharing Detailed Research Data Is

Associated with Increased Citation Rate." *PloS One* 2 (3): e308.
**Piwowar, Heather A., and Todd J. Vision.** 2013. "Data Reuse and the Open Data Citation Advantage." *PeerJ* 1: e175.
**Simonsohn, U., L.D. Nelson, and J.P. Simmons.** 2014. "P-curve: A Key to the File-Drawer." *Journal of Experimental Psychology: General* 143 (2): 534–47.
**Swanson ⓡ, Nicholas, Garret Christensen ⓡ, Rebecca Littman ⓡ, David Birke ⓡ, Edward Miguel ⓡ, Elizabeth Levy Paluck ⓡ, and Zenan Wang ⓡ.** 2020. "Research Transparency Is on the Rise in Economics." *AEA Papers and Proceedings* 110: 61–65.
**Turner, Erick. H., Annette M. Matthews, Eftihia Linardatos, Robert A. Tell, and Robert Rosenthal.** 2008. "Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy." *New England Journal of Medicine* 358 (3): 252–60.
**Vilhuber, Lars.** 2020. "AEA Data and Code Availability Policy." *AEA Papers and Proceedings* 110: 776–78.
**Vilhuber, Lars, James Turrito, and Keesler Welch.** 2020. "Report by the AEA Data Editor." *AEA Papers and Proceedings* 110: 764–75.