

Ulysses' Pact or Ulysses' Raft: Using Pre-Analysis Plans in Experimental and Non-Experimental Research*

Sarah A. Janzen^a and Jeffrey D. Michler^b

^a*Department of Agricultural and Consumer Economics, University of Illinois*

^b*Department of Agricultural and Resource Economics, University of Arizona*

Abstract

Economists have recently adopted pre-analysis plans in response to concerns about robustness and transparency in research. The increased use of registered pre-analysis plans has raised competing concerns that detailed plans are costly to create, overly restrictive, and limit the type of inspiration that stems from exploratory analysis. We consider these competing views of pre-analysis plans, and make a careful distinction between the roles of pre-analysis plans and registries, which provide a record of all planned research. We propose a flexible “packraft” pre-analysis plan approach that offers benefits for a wide variety of experimental and non-experimental applications in applied economics.

JEL Classification: A14, B41, C12, C18, C90, O10, Q00

Keywords: Pre-analysis plan, Pre-registration, Hypothesis Registry, Research Ethics, Transparency

*Authors listed alphabetically by last name; corresponding author is Sarah Janzen. Author emails: sjanzen@illinois.edu and jdmichler@arizona.edu. Paper accepted on November 20, 2020. The authors thank Walter Acpangan, Bolanle Atilola, Sebastian Bascom, and Brian McGreal for research assistance. Anna Josephson, Talip Kilic, Nicholas Magnan, Conner Mullally, David Spielman, and the AEPP editor Craig Gundersen all provided helpful comments. We also are grateful for discussion and comments by participants of the AAEA Annual Meeting International and Africa Section Virtual Track Session 2020. All opinions and errors are our own.

*They bound me hand and foot in the tight ship—
erect at the mast-block, lashed by ropes to the mast.*

— Homer, *The Odyssey*, Book XII

*And now the master mariner steered his craft,
sleep never closing his eyes, forever scanning
the stars.*

— Homer, *The Odyssey*, Book V

The research process is not unlike the journey of Ulysses (Latin translation of the Greek Odysseus) from the plain before Troy to Ithaca, his home. The journey is long and arduous; Ulysses took ten years to get home. The challenges along the way are many and often unanticipated; Ulysses faced monsters, gods, and numerous distractions. Even the final result can be surprising; Ulysses found his home full of suitors vying to displace him. Throughout, Ulysses relied on his *mētis*, his “cunning intelligence,” to deal with each new challenge. As Dougherty (2015) argues, what characterizes Ulysses’ success is his ability to adapt the tools at hand to each new situation. As Ulysses approached the island of the Sirens, who lured sailors to their death with beautiful singing, he made a pact with his sailors. Ulysses would fill the ears of each sailor with wax, so they could not hear the sirens’ song, and the sailors would tie Ulysses to the mast so he could listen to but not act on the Sirens’ temptation. At another point in his journey, Ulysses was stranded on the island of the nymph Calypso, who enchanted him with her singing. In this case, Ulysses improvised a raft out of the materials at hand to escape the island.

These two literary tropes, Ulysses’ pact and Ulysses’ raft, represent two distinct, some might argue opposite, visions of the research process. In the medical and legal profession, a Ulysses pact is a freely made decision to bind oneself to a set of actions in the future (Spellecy, 2003).¹ The past decade has seen the rapid rise of a Ulysses pact in economics - the pre-analysis plan. Proponents of pre-analysis plans contend that pre-commitment to a specific analysis plan is the best way to minimize the probability of false discoveries (Olken, 2015). These proponents highlight how the use of a pre-analysis plan mirrors the scientific method, by hypothesizing first, then collecting data, and finally using the data to confirm or reject the hypothesis (Miguel et al., 2014).

Critics of pre-analysis plans argue that binding oneself to the mast of a ship is foolish. Better to, like Ulysses on Calypso’s island, build a raft using the material at hand, and go exploring (Brewster, 1926). These critics claim that pre-analysis plans are overly restrictive, limiting a researcher’s ability to explore the data and make the type of insights only possible *ex post* (Leamer, 1978). Furthermore, some argue pre-analysis plans are designed to solve a problem common to biomedical trials but

¹In literature, it is more common to refer to the Greek Odysseus than the Latin Ulysses. In contrast, in law and medicine, the contract is commonly referred to as a “Ulysses pact.” For consistency, we use the Latin form of Odysseus throughout the paper.

less common in the social sciences - the incentive to falsify research (Coffman and Niederle, 2015). Why bind oneself to the mast when the sirens aren't present?

In this paper, we explore the application and value of pre-analysis plans for applied economists. Such plans are closely associated with the online registries designed to keep a public record of such plans, although as we point out, registries and pre-analysis plans should be viewed as two distinct tools, each fulfilling a unique purpose. We are not the first to discuss the topic,² though our paper offers three unique contributions. First, we draw attention to the distinct purpose of pre-analysis plans as opposed to registries, a distinction which has often been lost in the economics literature. Second, we highlight many practical benefits of pre-analysis plans that are often overshadowed by the discussion around the proposed gains related to publication bias and researcher ethics. Third, given these practical benefits, we suggest that pre-analysis plans can be usefully applied to a wide variety of experimental and non-experimental applications in applied economics, including the economics of agriculture, international and rural development, resources and the environment, food and consumer issues, and agribusiness.

We begin with a brief history of pre-analysis plans, and online registries for those plans, in the social sciences, followed by a description of the content of a pre-analysis plan. The key insight emerging from these two sections of the paper is that the problems pre-analysis plans are designed to solve differ in important ways from the problem registries of plans are designed to solve. A pre-analysis plan reduces a researcher's degrees of freedom: the ability to select data and methods to ensure "discovery" of significant results. This can take the form of running many statistical tests but only reporting significant tests (data dredging), reanalyzing data to yield a target result (*p*-hacking), or "Hypothesizing After Results are Known," and presenting these as *a priori* hypotheses (HARKing). By comparison, registries, either of pre-analysis plans or research studies undertaken, creates a record of all planned research and is designed to solve the "file drawer problem." Here the issue is that studies that fail to produce a statistically significant result are less likely to be published than those that do produce a statistically significant result. This publication bias creates a distorted picture of the state of knowledge in science. What is important for registries is that they contain the universe of all studies undertaken.

We next review the debate between those who promote the use of pre-analysis plans and those who see the costs of such pacts as outweighing the benefits. Economists have used pre-analysis plans almost exclusively for experimental work, and in the next section we suggest this narrow focus misses out on a number of potential benefits for pre-specifying replication studies as well as purely observational studies. In these sections we illustrate the costs and benefits of developing pre-analysis plans for both experimental and non-experimental research using our recent experience in developing such plans. A theme that emerges from both of our experiences is that the benefits of pre-specification and registration extend beyond reducing data dredging, *p*-hacking, and HARKing;

²See Coffman and Niederle (2015); Olken (2015); Christensen and Miguel (2018); Duflo et al. (2020) for recent discussion offering differing views.

or limiting the file drawer problem. Substantial benefits accrue to the researcher by forcing them to think through and address constraints ahead of time and by setting expectations for research teams.

In the final section, we discuss a middle ground that might be termed Ulysses' packraft. A packraft is a portable inflatable boat first designed by Peter Halkett in the 1840s. The packrafts were used by the Hudson's Bay Company in exploring the Canadian arctic. An explorer will hike with the packraft for use when a river, lake, or ocean blocks their path. Though packrafts did not exist in Homeric Greece, Ulysses might have found one useful in escaping the god Helios or leaving the Phoenicians for his home in Ithaca. The packraft approach would have the researcher develop a basic pre-analysis plan and file it with a public registry. The registered plan documents the researchers' intentions, but unlike a binding pact, researchers are allowed to update the plan as new information arises. Transparency about deviations from the original plan allows for flexibility, while also addressing issues of data dredging, p -hacking, and HARKing. Registration of the plan helps resolve the file draw problem. Unlike previous proposals in economics, use of a Ulysses packraft is not limited to experimental work, but can be used by any economist engaged in new research. Ultimately, the goal of the packraft approach is to allow the modern day Ulysses to prepare a plan, while providing the flexibility to take an alternative route as needed.

A brief history of pre-analysis plans

A pre-analysis plan (PAP), sometimes referred to as a statistical analysis plan, hypothesis registration, or a pre-specified research design, is a document that defines how the researcher plans to analyze a data set prior to any analysis. The goal of such a plan is to form a Ulysses pact in which the researcher is bound to implement the plan, regardless of what the research results might be. The plan typically describes the the data, hypotheses, outcomes of interest, how key variables will be defined, and the statistical models to be estimated. Such plans are filed with an online registry where they are time stamped and can be viewed by other researchers. The plan should be written and filed prior to data analysis, and ideally prior to data collection.

The use of statistical analysis plans has been required by law in the United States for all drug trials³ since November 1997 (Casey et al., 2012). The stated purpose of the law and the registry it created in February 2000 (ClinicalTrials.gov), was to counter the strong incentives of drug companies to falsify research on the efficacy of the drugs they develop. Without a pre-defined plan of analysis, drug companies would have the freedom to data dredge, p -hack, or HARK their way to an "effective" drug.

In 2004, the twelve journals belonging to the International Committee of Medical Journal Editors (ICMJE) announced that public trial registration would be a pre-requisite for publication in

³Other non-drug clinical and medical trials in the U.S. do not require registration of statistical analysis plans. Drug trials outside of the U.S. are also not affected by the regulation.

any of the journals (De Angelis et al., 2004). According to Franco et al. (2014), the new policy sought to combat the file drawer problem. The concern was that researchers, referees, and editors alike get more excited about new results if they are large in magnitude and statistically significant, than when they are small or non-significant. Null results and those that simply confirm previous work are frequently ignored. To combat this, registries create a record of all studies so that, as De Angelis et al. (2004) writes, “anyone [is] able to learn of any trial’s existence and its important characteristics.” Prior to the change in editorial policy, ClinicalTrials.gov averaged 1,879 registrations per year. Since the change in policy, the average number of new registrations made each year at ClinicalTrials.gov is 21,492.⁴

Around the same time, Neumark (2001) was the first to use a pre-analysis plan, what he called a “prespecified research design,” in economics. Neumark’s research question concerned the effect of increases to minimum wage on employment. Neumark (2001) begins by citing ten recent studies that asked the same question and noting two patterns. First, there was wide variation in the reported effects, with some papers reporting large positive impacts and others reporting moderate negative impacts. Second, multiple studies by the same author(s) tended to find the same results, suggesting the existence of “author effects,” in which author(s) publish results that confirm or are consistent with their previous published research. Neumark (2001) is careful to not make any strong claims of researcher misconduct, but raises the possibility that researchers either are engaged in (possibly unintentional) specification search to find results that align with their priors, or are burying results that fail to align with their priors in the proverbial file drawer. To guard against this potential author bias, Neumark developed a pre-analysis plan in which he described his research question, defined the data to be used, the construction of variables, and wrote out the statistical models he would estimate. Neumark (2001) suggests the “evidence resulting from such prespecified research designs should provide estimates of the employment effects of minimum wages that are free of author effects by eliminating specification search.”

Although Neumark’s analysis was published in 2001, pre-analysis plans were not widely discussed or adopted by economists for at least another decade.⁵ That changed in 2012 with two articles published in the *Quarterly Journal of Economics*. One of those articles is Finkelstein et al. (2012), who concisely state, “Our prespecification was designed to minimize issues of data and specification mining and to provide a record of the full set of planned analyses.” The second seminal article was Casey et al. (2012), who utilize a pre-analysis plan “to bind [their] hands against data mining.” The Casey et al. (2012) paper provides a detailed discussion of the value of pre-analysis plans, especially when combined with registration. The authors make the case that pre-analysis plans can be particularly useful when researchers have “wide discretion over what they report”

⁴This calculation does not take into account a number of other public registries that have been created since 2004. The calculation is based on the number of trials listed in the registry (357,017) as of 6 November 2020. The number of pre-announcement registrations (12,153) comes from Humphreys et al. (2013).

⁵One potential reason why Neumark (2001) is overlooked is that he developed a pre-analysis plan for an observational study, not for an experimental one, which is where we have seen the recent discussion focus.

(alluding to data dredging and p -hacking) and when they “face professional incentives to affirm the priors of their academic discipline or the agenda of donors and policy makers.” The authors illustrate this using an RCT evaluating an approach to making institutions more democratic and egalitarian. They point out that there are multiple ways one could define and measure “success” in this context, and show how data dredging or selective presentation could lead to two opposite sets of conclusions. The authors also point out the potential conflict of interest that arises when researchers work quite closely with project implementers who have a vested interest in a positive outcome. By developing a pre-analysis plan, the researchers not only constrained themselves, they also constrained their partners – similar to the constraints imposed on drug companies through the establishment of ClinicalTrials.gov – thereby reducing organizational pressure to come up with favorable results.

In 2012, the American Economic Association (AEA) established the [AEA RCT Registry](#) for lab and field experiments in social sciences. The AEA RCT Registry does not allow for non-experimental studies, but such studies can be registered with the [Open Science Foundation](#) (OSF). The OSF platform is also used to host registries maintained by other organizations. For example, the International Initiative for Impact Evaluations (3ie) maintains the [Registry for International Development Impact Evaluations](#) (RIDIE), and allows both RCTs and quasi-experimental program evaluations, but restricts trials to those undertaken in low and middle income countries.

In economics (Olken, 2015), as well as political science (Humphreys et al., 2013), the need for a registry of pre-analysis plans is frequently tied to precedents in the medical literature. Although a pre-analysis plan and a registry of those plans are often discussed simultaneously, it is helpful to make a clear distinction between the goals of each. A registry creates a record of all planned analyses, regardless of whether findings are ultimately null or statistically significant. In this way, the registry could theoretically accommodate rich meta-analyses that can be used to fill gaps in the literature, directly targeting the file drawer problem.⁶ Alternatively, the pre-analysis plan is a Ulysses pact that restricts *how* the hypotheses are to be tested, and in some cases, how the data will be collected. A pre-analysis plan reduces the ability of a researcher to choose a preferred hypothesis, empirical specification, or data set *ex post*. At least in theory, the result is that a pre-analysis plan increases the probability that a published positive result is true, explicitly combating researcher misconduct. Filing a pre-analysis plan on a registry helps ensure that the researcher will be held to the Ulysses pact, but a pre-analysis plan could just as easily be filed and timestamped on [arXiv.org](#) or with a journal editor. In fact, some journals now accept registered reports, which are reviewed and approved for publication before data analysis has begun.⁷

To assess the extent pre-analysis plans are being used by the profession, we reviewed recent

⁶As of now, there is no mechanism to force authors to update the registry with their results, which complicates this analysis. However, Duflo et al. (2020) argue for establishing norms in economics to update registries, and the AEA registry now has a section for results.

⁷Examples include the *Journal of Development Economics*, the *Journal of Accounting Research*, *The Review of Financial Studies*, and *PLOS One*.

articles published in the flagship journals of the Agricultural and Applied Economics Association (*American Journal of Agricultural Economics*) and the American Economics Association (*American Economic Review*). Although not a comprehensive review of applied economics journals, our small review highlights the extent to which members of these associations are utilizing pre-analysis plans. To complete the review, we searched all articles published after January 2018, including recently accepted articles (up until May 2020) for any of the following search terms: “PAP,” “pre-analysis plan,” “preanalysis plan,” “analysis plan,” “registry,” and “registered report.” The search included supplemental files, particularly online appendices.⁸

Many of the articles that met the initial search criteria referred to a data registry; some referred to a “voter registry” or something entirely irrelevant to our intended search. As such, we conducted a second qualitative review to assess whether articles made a clear reference to a pre-analysis plan. Among the 40 articles that met the initial search criteria, we found four *American Journal of Agricultural Economics* and 17 *American Economic Review* articles with clear references to pre-analysis plans. These articles are listed in Table 1. Other articles that met the initial search criteria but failed to reference an actual pre-analysis plan are listed in Appendix Table ???. It is interesting to note that all four *AJAE* articles listed in Table 1 were published in the early part of 2020. Although adoption of pre-analysis plans seems to be increasing, the adoption of pre-analysis plans is still far from universal. Rather, the use of pre-analysis plans still seems to be the exception rather than the norm.

Components of a pre-analysis plan

There exists considerable heterogeneity in what actually goes into a pre-analysis plans. The length and content of a plan tends to be based on what the researcher views as the severity of the problem pre-analysis plans are trying to solve. Here we describe the components of a basic pre-analysis plan as well as the additional components that can be included for a more comprehensive plan.

Basic requirements of a pre-analysis plan

Coffman and Niederle (2015) along with Duflo et al. (2020) promote the adoption of a concise pre-analysis plan that includes the following components.

1. **Data description:** Describe the process by which data was generated or collected. This includes the source of data, including where exactly it comes from (place and time) but also how it was collected (survey instruments, administrative data, etc.). Information on sampling method should be included. If the study relies on an experiment, information

⁸A complete review of each journal was conducted by two separate research assistants, but it is possible that some relevant articles were missed.

on the randomization method should be detailed enough that a different researcher could confidently replicate the method.

2. **Hypotheses:** Clearly state the research questions and hypotheses to be tested.
3. **Outcome variables:** List the primary outcome variables as well as any secondary outcomes to be tested.
4. **Variable construction:** Precisely define how each outcome and control variable is to be constructed.
5. **Model specification:** Describe the statistical method of analysis. In most cases, the researcher will write out the exact equations to be estimated, list all covariates, and specify an approach to inference (clustering of standard errors, bootstrapping, randomization inference, etc.).

These items are components of most grant proposals so, for those seeking funding, much of the work of writing a basic pre-analysis plan is already being done. One simply needs to file it with a registry.

Additional components of a comprehensive pre-analysis plan

For studies in which the researcher has substantial degrees of freedom to choose one method over another, a more comprehensive pre-analysis plan may be useful. McKenzie (2012), Humphreys et al. (2013), and Olken (2015) have all promoted more comprehensive pre-analysis plans which include some or all of the following additional components.

1. **Motivation:** Provide motivation and context for the study. A review of the literature can be helpful.
2. **Theoretical model and/or theory of change:** If a theoretical model is being tested, include it as part of the pre-analysis plan. A theory of change or causal diagram that justifies the selected hypotheses can also be included.
3. **Multiple inference adjustments:** If multiple hypotheses are being tested, address how multiple inference will be dealt with. There are numerous ways to adjust for either the familywise error rate (FWER) or the false discovery rate (FDR), and the economics profession has yet to settle on a preferred approach, which means the researcher has additional degrees of freedom.
4. **Data cleaning:** Outline basic operating principles for how data cleaning will be conducted, such as winsorizing outliers, correcting for attrition, and adjusting for contamination.

5. **Subgroup analysis:** Describe any subgroup or heterogeneity analysis. Both in medical trials and in economics, there are numerous ways that the researcher can subdivide the data in order to find heterogeneous effects. By pre-specifying what subgroup analysis will occur, the researcher avoids the temptation to data mine for clever or unusual differences across different populations.
6. **Power calculations:** If data allows, include information on sample size and power calculations. Ensuring sufficient power to reject nulls is most beneficial *ex ante*.
7. **Exploratory analysis:** Recent work by Anderson and Magruder (2017) and Fafchamps and Labonne (2017) develop strategies for exploring a partial data set to help inform decisions. One might even choose to simulate a model data set, as in Humphreys et al. (2013), to conduct trial runs of different statistical models. Simulated data would only be used to choose the preferred model *ex ante*, while the true data would be used *ex post*.

Proponents of the comprehensive pre-analysis plan suggest whenever a researcher faces a choice, which could ultimately affect results, they should always pre-specify a preferred approach. By pre-specifying an approach, the temptation to present results that confirm a researchers pre-existing biases (whether known by the researcher or not) is removed.

Critics point out that sometimes the “preferred approach” is not obvious *ex ante*. If no “best practice” exists, then how exactly should a researcher choose a preferred approach? Should they flip a coin, or choose the practice they are most familiar with (possibly biasing the research further)? If a “best practice” consensus emerges after registering the pre-analysis plan, is the author allowed to update their decision? These opponents argue the researcher is as much an artist as a scientist, and should have some degree of control over critical decisions not only at the onset of the research process, but over the course of the full research process as new information is constantly being learned.

These different viewpoints highlight a growing tension over the value of pre-analysis plans in our discipline. The next section unpacks these different arguments further.

Reconciling costs and benefits of pre-analysis plans in economics

Similar to the debate regarding the use of randomized control trials 15 years ago, there is an ongoing debate regarding the use of pre-analysis plans and registries in economics. Much of the debate hinges on what one sees as the problem pre-analysis plans are trying to solve, and how prevalent that problem is within the research community. If the problem that pre-analysis plans address is *p*-hacking, data mining, and HARKing, and that problem is pervasive, than the benefits from forming a Ulysses pact outweigh the cost of sacrificing the ability to explore and improvise on a Ulysses raft. However, if *p*-hacking is not a pervasive problem, or if the primary purpose of

pre-analysis plans is to publicly register hypotheses to avoid the file draw problem, than there is no need to bind oneself to the mast, and the researcher should be able to explore as they see fit.

Proponents of comprehensive pre-analysis plans believe that the dangers of the siren song of p -hacking are extreme. Olken (2015) provides an early review article promoting the advantages of pre-analysis plans. He begins with a hypothetical story of a nefarious researcher trying to mine the data or p -hack their way to a significant result. Having started with this spooky tale of the nefarious researcher, Olken (2015) recommends developing a detailed pre-analysis plan. For him, a pre-analysis plan not only limits researcher degrees of freedom, it also creates several advantages for the constrained researcher. A pre-analysis plan allows the researcher to take advantage of all statistical power by using one-sided tests and avoiding corrections for multiple inference. It also reduces the need to conduct endless robustness checks, since the pre-analysis plan eliminates the fear that the researcher engaged in an extensive specification search. A key argument in favor of pre-analysis plans is that, in a world where replication is unlikely, pre-analysis plans are a tool for ensuring that significant results were not generated through “research initiative,” Glaeser’s (2008) euphemism for malfeasance and fraud. While Olken (2015) and other proponents of comprehensive pre-analysis plans readily acknowledge the costs of developing the plans, they argue that the benefits outweigh the costs.

In juxtaposition to Olken (2015), opponents view pre-analysis plans as an overly onerous attempt to solve what in the social sciences is a relatively minor problem. Here Coffman and Niederle (2015) are representative. For these authors, the benefits of forming a Ulysses pact are not obvious, since the prevalence of p -hacking is not obvious. Using data from Brodeur et al. (2016), Coffman and Niederle (2015) argue that there is little evidence of p -hacking in experimental work. In fact, the type of studies where Brodeur et al. (2016) and Brodeur et al. (2020) find evidence of p -hacking is in non-experimental work that relies on instrumental variables and other methods of causal identification. Coffman and Niederle (2015) point out that those championing the use of pre-analysis plans focus almost exclusively on their use in RCTs, the place where, given available data, they are needed the least. Instead of a comprehensive pre-analysis plan, Coffman and Niederle (2015) make the case for hypothesis registries to solve the file drawer problem, and for more replication studies to reveal p -hacking. While Coffman and Niederle (2015) and other critics of pre-analysis plans acknowledge their value in limiting researcher degrees of freedom, they argue that the costs outweigh the benefits of allowing researchers to explore their data and make new discoveries.

Recent reflections on the use of pre-analysis plans have taken a more pragmatic approach and tend to align with the idea in Coffman and Niederle (2015) that, at least for experimental work, the value of pre-analysis plans lies in their public registration. Duflo et al. (2020) argue that pre-analysis plans should not be overly specific or long. Rather, the plans should chart a course for the research while also allowing the researcher freedom to explore. Duflo et al. (2020), similar to Humphreys et al. (2013), see the role that pre-analysis plans play in communicating what did

and did not work (the file drawer problem) as more valuable than restricting researcher degrees of freedom.

There are several reasons for the seeming emergence of a more moderate approach, what we coin the packraft approach, to the use of pre-analysis plans. One reason is the continued evidence that the problem of p -hacking is not as pervasive as researchers originally feared. Gelman (2013), among others, argue that quantitative social scientists should not be particularly worried about committing Type I errors. Rather, we should be more concerned about Type S errors, getting the sign right, and Type M errors, getting the magnitude right (Gelman and Tuerlinckx, 2000; Ziliak and McCloskey, 2008; Rommel and Weltin, 2020). Pre-analysis plans are less useful in combating these type of errors. A second reason for the emergence of a scaled back approach to what pre-analysis plans should include is just how onerous a comprehensive pre-analysis plan is to develop. As experiments become more complex, the costs of developing a thorough pre-analysis plan grow substantially.

Although pre-analysis plans are usually promoted for the benefits they provide to the profession, there are additional practical benefits for a researcher. First, it is simply a good exercise to carefully think through how the data will be used before collecting or acquiring access to the data. In our experience, the attention to detail and thoughtfulness required to write a good pre-analysis plan, leads to an equally careful and thoughtful final analysis. Second, we have found that pre-analysis plans can be particularly helpful when working as part of a research team. Since it is usually up to a single team member to implement the analysis using statistical programming, it can be helpful to make sure everyone is on the same page – in a very detail-oriented way – regarding the empirical approach before actually executing it. Open communication regarding the plan prevents the need for everyone to needlessly sift through table after table of results. Instead, researchers commit to a method prior to seeing results, which may help research teams as a whole remain impartial. Third, soliciting feedback at the planning stage can prevent high cost mistakes, some of which can be impossible to alter *ex post*, dramatically improving research quality.

The following case study provides an example of the more moderate packraft approach. The approach taken by the authors retains many of the benefits of a traditional pre-analysis plan, including the valuable practical benefits described in the previous paragraph. However, the flexible approach taken reduces the cost of utilizing such a plan.

Case study #1: A packraft pre-analysis plan for an experimental study

Janzen et al. (2020) utilize a pre-analysis plan for a randomized control trial with smallholder grain producers in Kenya. The pre-analysis plan was written after data was collected (in July 2017), but before data was analyzed in any way. The study was registered in the AEA RCT Registry on 29 August 2017, with updates one week later (Hughes et al., 2017).⁹ The pre-analysis plan

⁹The paper was accepted by the *American Journal of Agricultural Economics* after the review presented earlier,

used for Janzen et al. (2020) contains all the basic required components of a pre-analysis plan and many optional details (e.g., motivation, discussion around multiple hypothesis testing, subgroup analysis, power calculations). Because it is a lab-in-the-field experiment with multiple treatments, the pre-analysis plan describes at length the treatments and experimental design.

A unique component of the study was the collection of willingness to pay for insurance data, collected during a multiple price list auction. Rarely do economists collect this sort of data, which accommodates drawing an individual’s complete demand curve. Working with new unfamiliar data raises both the benefits and costs of a comprehensive pre-analysis plan. While limiting p -hacking, the plan also constrains researchers’ ability to explore alternative approaches for using data they are unfamiliar with. In this case, the plan pre-specified two linear regression approaches to be used, while also retaining the option to explore alternative approaches “if a scatter plot shows that demand is highly nonlinear in price.” This approach aligns with the recommendation of Duflo et al. (2020) who state, “when the researchers are not sure of something at the time of writing the pre-analysis plan, they should note it and why.”

There are many times a researcher’s initial ideas are later deemed uninformative, unnecessarily complicated, or even wrong. A packraft approach to pre-analysis plans allows the researcher to update decisions. There are many ways in which the authors decide to alter course, deviating from the initial plans laid out in Hughes et al. (2017). In the spirit of transparency, these deviations and the reasons for doing so are outlined in Online Appendix G of the published manuscript. Alternatively, the researchers could have created what Duflo et al. (2020) call a “populated PAP” with all pre-specified regressions. In some cases, certain pre-specified analyses were redundant or uninformative and would have led to an unnecessarily lengthy, and uninteresting, paper. In these cases, the authors used their judgment to choose the best specification while omitting others (or including them in the online appendices). In other cases, the authors realized retrospectively that some of their pre-specifications overlooked details that, if incorporated, would improve the analysis.

As one example of a deviation from the original plan, Equation 1 in the published paper is used to analyze impacts on learning, instead of the analogous pre-specified equation. In the paper, the authors only consider the effects of playing an insurance game on learning, and do not consider the quality of the insurance product was offered. The reason for this, as articulated in the paper’s Online Appendix G, is that “in retrospect, there is no reason to believe playing either... game would convey different information with respect to the knowledge questions farmers were asked.” The paper’s appendix Table A.5 presents the results from estimating the pre-specified equation. This same initial oversight also affected the researchers’ approach to inference for the same equation. The research design consists of two treatments types in which one randomization (insurance type) was carried out at a cluster (session) level, whereas the second randomization (game type) was done at the individual level. The pre-analysis plan states that “Hypothesis testing for effects identified

and thus not included in Table 1.

by the randomization across sessions will rely on the wild cluster bootstrap...” whereas, “since the game treatment was randomized within sessions, inference on the game effect... can be done out without accounting for clustering.” The pre-specified equation would have required the wild cluster bootstrap, resulting in lower power than the estimated equation presented in the the main manuscript.

In the end, some of the findings reported in the final manuscript were unanticipated. To explain these unanticipated findings, the authors employ a theoretical model from Gagnon-Bartsch and Bushong (2019). When researchers set out to test a theoretical model, it is appropriate to include the theoretical model in the pre-analysis plan. In this case, the findings were unanticipated, the RCT was not designed to test a theoretical model, and the model was not included in the pre-analysis plan. To make this evident in the final paper, the theoretical model is included after the presentation of results, instead of at the beginning of the paper (prior to the empirical model) where a theory section can typically be found.

Frequently, the final analysis in a published paper will differ from the pre-analysis plan as a result of the peer-review process. As an example, the inclusion of a theoretical model to explain the findings in Janzen et al. (2020) was added based on the suggestion of a helpful reviewer. Indeed, the peer-review process is not set up to accommodate adherence to a well-designed plan. A helpful referee will often make useful suggestions, in which case authors have much to gain from plan deviations. Taking the packraft approach to pre-analysis plans is rather sensible in such cases.

Applying pre-analysis plans to non-experimental research

While the use of pre-analysis plans in observational studies is common in the medical sciences, economists have largely used pre-analysis plans only for experimental research.¹⁰ In addition, as noted above, the AEA registry is restricted to experimental studies. Duflo et al. (2020) state the registry was designed to be a social science version of ClinicalTrials.gov, with the goal of providing “a list of the universe of randomized control trials (RCTs).” Yet, ClinicalTrials.gov does not restrict registration to experiments. Of the universe of trials registered at ClinicalTrials.gov, 76,197 are observational studies, or about 21 percent of all entries.¹¹ If the AEA Registry seeks to solve the file drawer problem per De Angelis et al. (2004), then it remains unclear why the registry excludes non-experimental work.

In fact, there is nothing inherent in the basic or even a more comprehensive pre-analysis plan outlined earlier that would exclude its use for studies based on non-experimental data. Even power calculations, which are typically thought of as being only useful for experiments, can be

¹⁰As of this writing, we are aware of only four pre-analysis plans used in observational studies that seek to answer new research questions (i.e., not replications). These are Neumark (2001), Chang et al. (2020a), Neumark and Yen (2020) and Michler et al. (2020). Monogan III (2013) provides an example from political science.

¹¹Numbers as of 6 November 2020.

implemented with non-experimental data. Ensuring sufficient power to reject nulls is every bit as important in observational studies as it is in experimental studies (Brown et al., 2019).

One possible reason for ignoring the use of pre-analysis plans in economic non-experimental studies is because the early literature was primarily focused on their pre-commitment aspect, in contrast to the more recent focus on the file drawer problem. The main challenge with writing a pre-analysis plan for non-experimental studies is the retrospective nature of many such studies (Olken, 2015). In many cases, the researcher would be writing a plan with the data in hand. This provides an opportunity for a nefarious researcher to write a plan while simultaneously analyzing the data. The concern is that there is limited scope for curtailing the p -hacking that the pre-analysis plan was supposed to prevent, rendering the pre-analysis plan no better than the status quo. Indeed, Brodeur et al. (2016) and Brodeur et al. (2020) provide recent evidence that p -hacking is a bigger problem in non-experimental work than experimental research.

This seems to suggest there is value in finding creative ways to credibly document limited data access while writing a pre-analysis plan, in order to curtail concerns regarding p -hacking in observational studies (Burlig, 2018). Neumark (2001) found a way around this challenge. A key component of his plan was the use of government data released at the end of May 1997, allowing him to file the plan with journal editors prior to May 1997.¹² Numerous existing data sources require and record the data request, creating a public record of data acquisition. Even archival access typically requires submitting a request to the archive. In many cases, one could credibly demonstrate that the plan was filed prior data access, as in the case study described in the subsection below. Burlig (2018) discusses additional situations in which researchers can credibly commit to a pre-analysis plans for observational studies.

In fact, the same risk is also present in many experimental studies as well. The Hughes et al. (2017) pre-analysis plan described in the case study above was written after data was collected, but before data was analyzed. This scenario is not uncommon. Under these circumstances, just like with observational studies that cannot document a specific time at which data access was granted, a registered plan is useful for addressing the file drawer problem, but can only prevent data dredging and p -hacking if the researchers can be trusted to have blinded themselves from the data, despite having access to it. Further complicating these trust issues, split-sample methods recently proposed by Anderson and Magruder (2017) and Fafchamps and Labonne (2017) require having at least some data in hand at the time of writing the pre-analysis plan.¹³ If plans written *ex post* are acceptable for experimental economists, then the same should be true for non-experimental studies.

As this paper has emphasized, pre-commitment is not the only benefit of writing a pre-analysis

¹²Neumark and Yen (2020) and Monogan III (2013) provide additional examples. Both rely on a timed data release, so the authors could credibly file their pre-analysis plan prior to ever having access to the data.

¹³Chang et al. (2020a) provide an example of using split sample methods in a non-experimental setting. After collecting the data, but prior to looking at it, they split the sample into an exploratory data set that contained a sub-set of older observations. They used this data for exploratory analysis, filed their pre-analysis plan on Open Science Foundation (OSF), and only then were given access to the full confirmatory data set (Chang et al., 2020b).

plan. If the primary value of a pre-analysis plan is related to the file drawer problem, then as Olken (2015) writes, “in principle, there is no reason” a registered pre-analysis plan cannot be used for an observational study.¹⁴ Because the packraft approach to pre-analysis plans is less restrictive than a traditional plan, we view it as more conducive to non-experimental work. We see two potentially fruitful extensions of pre-analysis plans into non-experimental work.

First, pre-analysis plans can help guide the researcher engaged in analyzing observational data. Case study #2 below provides one example. Another obvious application is to incorporate them into graduate student instruction and curriculum. The vast majority of students in applied economics graduate programs will not employ experimental methods in their masters’ thesis or PhD dissertation, but there is no reason this should preclude them from the benefits of writing a well thought out plan. In fact, it is typical for most PhD programs in Agricultural and Applied Economics to require defense of a prospectus. In practice, the prospectus is rarely prospective in nature; it typically requires presentation of some early results. In presenting a pre-analysis plan instead, the student will carefully contemplate the empirical strategy in a detailed manner. The prospectus becomes an opportunity to focus on the empirical approach – without knowing the results so that the student and committee can maintain impartiality – and accommodates an opportunity for feedback on those methods early on in the research process. The dissertation defense then becomes a time for discussing challenges and modifications, research findings, and implications. Another application is to require the writing of a pre-analysis plan for a hypothetical study in applied graduate-level classes. This could replace the typical final paper. This shifted focus would require students to spend more time thinking about their hypotheses, the data and any methodological assumptions prior to implementing data analysis. Given the profession’s increasing attention to methodological rigor, we think this is time well spent.

Second, pre-analysis plans can be used to guide replications. Replications have frequently been put forward as a substitute for pre-analysis plans. The idea is that a replication can provide the level of confidence in results meant to be conveyed by comprehensive pre-analysis plans (Coffman and Niederle, 2015; Coffman et al., 2017; Duvendack et al., 2017). However, pre-analysis plans can compliment and even facilitate replication. Coffman et al. (2017) suggest one reason replications remain uncommon in economics is because journals are reluctant to publish replications that confirm the initial analysis. This creates an ethical issue in that publication bias incentivizes the replicating researcher to find a way to overturn the findings of the original work. A pre-analysis plan for the replication directly combats this issue. While still uncommon in economics, there is a small group of authors using pre-analysis plans to direct their replication work (Chang and Li, 2017; Chang, 2018; Chang and Li, 2018, 2020). The forthcoming special issue on replications in this journal may provide an opportunity to use pre-analysis plans to guide replication work in applied economics.

¹⁴In sharp contrast to Olken (2015), Duflo et al. (2020) write that “PAPs for non-experimental research, which tends to be retrospective, are rarely advocated for or used (yet) in practice, presumably because they are neither desirable nor, in most cases, practical.”

In the following subsection, we provide an example of how a pre-analysis plan could be used in an observational study. The approach taken by the authors develops a traditional basic pre-analysis plan applied to the combination of several existing observational data sets.

Case study #2: A basic pre-analysis plan for a non-experimental study

Michler et al. (2020) provides an example of how and why one might develop a pre-analysis plan for an observational study. The central research question is, how does measurement error in remote sensing weather data affect estimates of agricultural production? To answer this question, the authors combine geospatial weather data from a variety of remote sensing products with the geo-referenced household survey data from the seven Sub-Saharan African countries that are part of the World Bank Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) initiative. The plan contains all five elements discussed as part of a basic pre-analysis plan.

The authors chose to develop a pre-analysis plan for two reasons. One was to provide a shield against criticism of author bias if the research revealed that some remote sensing products were less accurate than others. To accomplish this, the authors split the research team into two (and pre-specified the teams), one team to be responsible for the remote sensing weather data and the other to be responsible for the analysis. In this way, and similar to medical trials, the analysis team was blinded to which data came from which product. A second reason for the pre-analysis plan was to form a Ulysses pact, limiting the analysis team's ability to engage in data mining or specification search. With this goal in mind, the authors prespecified the remote sensing data products, variable construction, and model specification. The pre-analysis plan was filed with OSF prior to gathering any remote sensing data (Michler et al., 2019).

In order to address concerns about credibly filing the pre-analysis plan prior to the start of analysis, the authors specified how the research team would be divided in two and how data would be shared between the teams. Per World Bank privacy policies, the geo-referenced household data is not available to the public, only the anonymized data. This means that no one outside the World Bank can perfectly match the household data with the geospatial weather data. Furthermore, the sharing of the matched household/geospatial data was done via a secure server which included timestamps. Thus, third parties are able to verify that the data was shared after the filing of the pre-analysis plan.

Some critics have argued that observational studies require more exploratory work than experimental studies, but this turned out not to be an issue. The research team found that having a pre-analysis plan was actually beneficial in circumscribing the possible avenues of exploration. As Gelman (2013) writes, "often it is only when focusing on the write-up that we fully engage with our research questions." One goal of the study was to explore the extent to which estimates of agricultural production depend on systematic measurement error. To ensure that any results were not simple local anomalies, they cast a wide net in terms of what data to include and what

specification to estimate. Developing a pre-analysis plan allowed the authors to discuss ahead of time and decide upon what were likely to be the most fruitful combinations of data/variable/specification. The pre-analysis plan also provided a stopping rule, eliminating the need to test yet one more combination during the writing process.

While implementing the pre-analysis plan, the authors encountered a number of challenges, gaining a number of insights. The first key insight relates to the costs of writing a plan. The plan runs for 25 pages and provides a high level of detail in terms of data sources. Much of this detail was unnecessary and could have easily been replaced with references to both the remote sensing and LSMS-ISA repositories. The initial goal was to pre-write a large amount of the subsequent research paper, including the introduction, data section, and empirical analysis section, as in Finkelstein et al. (2012). Yet, as the analysis played out, the authors realized that much of this material required substantial revision, making the time spent in writing these parts of the plan ill spent. As Duflo et al. (2020) has recently argued, pre-analysis plans should focus on defining the critical elements that govern the analysis, not spend time trying to write the paper prior to the analysis.

The original Michler et al. (2019) pre-analysis plan likely specified too many non-essential details. The over-specificity relates not to regressions or variable construction, but to how the authors promised to deal with outliers, missing values, and attrition. Pre-specified methods for dealing with troublesome data were based on past experience with cleaning large observational data sets. Yet the methods defined in the plan were frequently infeasible or not appropriate when applied across seven countries and 20 rounds of data. In hindsight, instead of defining a Procrustean standard for cleaning all variables, the plan should have specified guiding principles to govern the data cleaning. For example, instead of stating that continuous variables would be winsorized at the 1 and 99 percent, the authors could have stated that outliers would be dealt with through winsorization or imputation until the coefficient of variation for the variable fell within a target range.

None of these challenges are unique to observational data. In our experience, all of these issues arise when attempting to follow a pre-analysis plan for an RCT or lab experiment. While at times these challenges have been frustrating for the authors of Michler et al. (2020), none of them are fatal for the analysis. The packraft approach to a pre-analysis plan allows researchers the freedom to explore, as long as they clearly communicate what elements were prespecified and what elements arose after the fact.

Final thoughts on packraft pre-analysis plans

Much like the debate surrounding RCTs in the 2000s, the recent debate regarding the value of pre-analysis plans has pitted two distinct views of the research process against each other. One side has argued that the sirens' song to p -hack ones way to significance is too strong for economists to resist. Therefore, the ethical economist will form a Ulysses pact, tying themselves to the mast by

using a comprehensive pre-analysis plan and filing that plan, before data collection begins, with a public repository. Only then can the modern Ulysses safely traverse the seas of economic research. The other side has argued that, at least in experimental work, there is little danger from the sirens. Rather, what is needed is the freedom to build a Ulysses raft by searching for the appropriate specification and improvising with the data and materials at hand. Only with this sort of freedom can the modern Ulysses go exploring for new discoveries.

Based on our understanding of the ethical issues surrounding economics research, and our experience with developing and implementing pre-analysis plans in both experimental and non-experimental work, we propose an alternative metaphor, that of Ulysses' packraft. Though packrafts did not exist in Bronze Age Greece, we believe the concept would have applied to Ulysses' *mētis*, his "cunning intelligence," to deal with new challenges.

The packraft approach to developing pre-analysis plans envisions the researcher drawing up a basic pre-analysis plan described in this paper. Similar to Duflo et al. (2020), the plan is detailed enough to steer the researcher through key analytical challenges, while also allowing for the option to alter course later and explore. The plan would then be filed with a public repository like OSF or an expanded AEA Registry that allows for non-experimental work. Public registration of the plan helps resolve the file draw problem by creating full transparency with respect to performance and reporting of any and all economics research. A key component of the packraft approach is that, unlike previous proposals in economics, it is not limited to experimental work. Creating a list of economic studies that is truly universal, and not constrained to RCTs, is key to solving the file drawer problem. Even in the absence of registration, such plans can benefit all applied researchers by improving research quality.

The governing principle in developing packraft pre-analysis plans, registering them, and writing papers based on them should be transparency. As Laitin (2013) writes, "we ought to encourage a disciplinary practice of telling readers at what point in the research process (and for what reason) a particular model emerged, and to keep careful notebooks allowing us to report this accurately." The disciplinary practice of transparency that pre-analysis plans and hypotheses registries seek to foster should not be limited to experimental studies. And, as we have shown, developing a pre-analysis plan for observational research need not have higher costs than experimental work. Instead of pre-specifying exact methods for data cleaning and covariate selection, a basic pre-analysis plan for an observational study can prespecify approaches or principles that allow for adaptation to the data at hand. This is exactly what a packraft calls for - preparation that is flexible enough to adapt to unexpected obstacles.

To encourage adoption of what we call Ulysses' packraft, association journals could consider instituting a policy similar to that instituted by the editors of the ICMJE journals. For any new study that either 1) collects new data or 2) claims to establish a causal relationship, association journals could "require, as a condition of consideration for publication, registration in a public

trials registry” (De Angelis et al., 2004). We are not necessarily advocating for such a policy, only noting that a discussion of the costs and benefits is warranted. Such a policy would go a long way to solving the file draw problem by creating a nearly universal list of tested hypotheses in applied economics. Researchers seeking to test hypotheses and establish causal relationships would file, at a minimum, basic pre-analysis plans ahead of any analysis. Some research, in which economists have substantial researcher degrees of freedom, would necessitate more comprehensive pre-analysis plans (Duflo et al., 2020). Such a policy could still allow unfettered exploration of data when the researcher is simply seeking to document trends, relationships, or correlations.

Another avenue for encouraging adoption is to create space for presenting packraft pre-analysis plans in classrooms, during a regular seminar series, and at conferences. Such opportunities facilitate feedback *ex ante*, when it is needed most. Many donors already require award recipients to present research proposals in front of peers, encouraging constructive feedback early on.

We believe packrafts can be an important addition to the economist’s kit of research tools. At the start of every new project the economic researcher should prepare by equipping oneself with a packraft to help navigate unexpected obstacles.

References

- Anderson, M. L. and J. Magruder (2017). Split-sample strategies for avoiding false discoveries. Working Paper 23544, National Bureau of Economic Research.
- Brewster, F. (1926). The raft of odysseus. *Harvard Studies in Classical Philology* 37, 49–53.
- Brodeur, A., N. Cook, and A. Heyes (2020). Methods matter: p-hacking and publication bias in causal analysis in economics. *American Economic Review* 110(11), 3634–60.
- Brodeur, A., M. Lé, M. Sangnier, and Y. Zylberberg (2016). Star Wars: The empirics strike back. *American Economic Journal: Applied Economics* 8(1), 1–32.
- Brown, J. P., D. M. Lambert, and T. R. Wojan (2019). The effect of the conservation reserve program on rural economics: Deriving a statistical result from a null finding. *American Journal of Agricultural Economics* 101(2), 528–40.
- Burlig, F. (2018). Improving transparency in observational social science research: A pre-analysis plan approach. *Economic Letters* 168, 56–60.
- Casey, K., R. Glennerster, and E. Miguel (2012). Reshaping institutions: Evidence on aid impacts using a preanalysis plan. *Quarterly Journal of Economics* 127(4), 1755–1812.
- Chang, A. C. (2018). A replication recipe: List your ingredients before you start cooking. *Economics: The Open-Access, Open-Assessment E-Journal* 12, 1–8.
- Chang, A. C., L. R. Cohen, A. Glazer, and U. Paul (2020a). Politicians avoid tax increases around elections. Mimeo.
- Chang, A. C., L. R. Cohen, A. Glazer, and U. Paul (2020b). Politics of gasoline and other taxes. Retrieved from osf.io/f2nr9.
- Chang, A. C. and P. Li (2017). A preanalysis plan to replicate sixty economics research papers that worked half of the time. *American Economic Review* 107(5), 60–4.
- Chang, A. C. and P. Li (2018). Measurement error in macroeconomic data and economics research: Data revisions, gross domestic product, and gross domestic income. *Economic Inquiry* 56(3), 1846–69.
- Chang, A. C. and P. Li (2020). Is economics research replicable? sixty published papers from thirteen journals say “often not”. *Critical Finance Review*, forthcoming.
- Christensen, G. and E. Miguel (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Perspectives* 56(3), 920–80.
- Coffman, L. C. and M. Niederle (2015). Pre-analysis plans have limited upside, especially where replications are feasible. *Journal of Economic Perspectives* 29(3), 81–98.
- Coffman, L. C., M. Niederle, and A. J. Wilson (2017). A proposal to organize and promote replications. *American Economic Review* 107(5), 41–5.

- De Angelis, C., J. M. Drazen, F. A. Frizelle, C. Haug, J. Hoey, R. Horton, S. Kotzin, C. Laine, A. Marusic, A. J. P. Overbeke, T. V. Schroeder, H. C. Sox, and M. B. Van Der Weyden (2004). Clinical trial registration: A statement from the international committee of medical journal editors. *New England Journal of Medicine* 351(12), 1250–1.
- Dougherty, C. (2015). Nobody’s home: *Metis*, improvisation and the instability of return in Homer’s *Odyssey*. *Ramus* 44(1-2), 115–40.
- Duflo, E., A. Banerjee, A. Finkelstein, L. F. Katz, B. A. Olken, and A. Sautman (2020). In praise of moderation: Suggestions for the scope and use of pre-analysis plans for rcts in economics. Working Paper 26993, National Bureau of Economic Research.
- Duvendack, M., R. Palmer-Jones, and W. R. Reed (2017). What is meant by “replication” and why does it encounter resistance in economics? *American Economic Review* 107(5), 46–51.
- Fafchamps, M. and J. Labonne (2017). Using split samples to improve inference about causal effects. *Political Analysis* 25(4), 465–82.
- Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, K. 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.
- Franco, A., N. Malhorta, and G. Simonovits (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science* 345, 1502–05.
- Gagnon-Bartsch, T. and B. Bushong (2019). Learning with misattribution of reference dependence. Technical report, Michigan State University.
- Gelman, A. (2013). Preregistration of studies and mock reports. *Political Analysis* 21, 40–1.
- Gelman, A. and F. Tuerlinckx (2000). Type S error rates for classical and Bayesian single and multiple comparison procedures. *Computational Statistics* 15, 373–90.
- Glaeser, E. (2008). Researcher incentives and empirical methods. In A. Caplin and A. Schotter (Eds.), *The Foundations of Positive and Normative Economics*, pp. 300–19. Oxford: Oxford University Press.
- Hughes, K., S. Janzen, N. Magnan, C. Mullally, and J. Oduol (2017). Experimental games to teach farmers about weather index insurance. AEA RCT Registry. September 05. <https://doi.org/10.1257/rct.2401-2.0>.
- Humphreys, M., R. Sanchez de la Sierra, and P. van der Windt (2013). Fishing, commitment, and communication: A proposal for comprehensive nonbinding research registration. *Political Analysis* 21, 1–20.
- Janzen, S., N. Magnan, C. Mullally, S. Shin, I. B. Palmer, J. Oduol, and K. Hughes (2020). Can experiential games and improved risk coverage raise demand for index insurance? evidence from kenya. *American Journal of Agricultural Economics*, forthcoming.
- Laitin, D. D. (2013). Fisheries management. *Political Analysis* 21, 42–7.

- Leamer, E. E. (1978). *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York: Wiley.
- McKenzie, D. (2012). A pre-analysis plan checklist. <http://blogs.worldbank.org/impactevaluations/a-pre-analysis-plan-checklist>. World Bank Development Impact Blog, 28 October 2012.
- Michler, J. D., A. Josephson, T. Kilic, and S. Murray (2019). Empirically estimating the impact of weather on agriculture. OSF Registries. July 1. <https://doi.org/10.17605/osf.io/z3snh>.
- Michler, J. D., A. Josephson, T. Kilic, and S. Murray (2020). Estimating the impact of weather on agriculture. Mimeo, University of Arizona.
- Miguel, E., C. Camerer, K. Casey, J. Cohen, K. Esterling, A. Gerber, R. Glennerster, D. Green, M. Humphreys, G. Imbens, D. Laitin, T. Madon, L. Nelson, B. Nosek, M. Petersen, R. Sedlmayr, J. Simmons, U. Simonsohn, and M. Van der Laan (2014). Promoting transparency in social science research. *Science* 343(6166), 30–1.
- Monogan III, J. E. (2013). A case for registering studies of political outcomes: An application in the 2010 house elections. *Political Analysis* 21, 21–37.
- Neumark, D. (2001). The employment effects of minimum wages: Evidence from a prespecified research design. *Industrial Relations* 40(1), 121–44.
- Neumark, D. and M. Yen (2020). Effects of recent minimum wage policies in california and nationwide: Initial results from a pre-specified analysis plan. IZA Discussion Paper No. 13062.
- Olken, B. A. (2015). Promises and perils of pre-analysis plans. *Journal of Economic Perspectives* 29(3), 61–80.
- Rommel, J. and M. Weltin (2020). Is there a cult of statistical significance in agricultural economics? *Applied Economic Perspectives and Policy*, forthcoming.
- Spellecy, R. (2003). Reviving ulysses contracts. *Kennedy Institute of Ethics Journal* 13(4), 373–92.
- Ziliak, S. T. and D. N. McCloskey (2008). *The Cult of Statistical Significance: How the Standard Error Cost Us Jobs, Justice, and Lives*. Ann Arbor: University of Michigan Press.

Table 1: Recently published economics articles with pre-analysis plans

American Journal of Agricultural Economics		
2020	The Impact of Commercial Rainfall Index Insurance: Experimental Evidence from Ethiopia	Shukri Ahmed, Craig McIntosh, and Alexandros Sarris
2020	Information and Communication Technologies to Provide Agricultural Advice to Smallholder Farmers: Experimental Evidence from Uganda	Bjorn Van Campenhout, David J. Spielman, and Els Lecoutere
2020	Parent’s demand for sugar sweetened beverages for their pre-school children: evidence from a stated-preference experiment	Ou Yang, Peter Sivey, Andrea DeSilva, and Anthony Scott
2020	Producer attitudes toward output price risk: Experimental evidence from the lab and from the field	Marc Bellemare, Yu Na Lee, and David Just
American Economic Review		
2018	Measuring and Bounding Experimenter Demand	Jonathan de Quidt, Johannes Haushofer, and Christopher Roth
2018	Time versus State in Insurance: Experimental Evidence from contract farming in Kenya	Lorenzo Casaburi and Jack Willis
2018	Why do Defaults Affect Behavior? Experimental Evidence from Afghanistan	Joshua Blumenstock, Michael Callen, and Tarek Ghani
2019	Demand and Supply of Infrequent Payments as a Commitment Device: Evidence from Kenya	Lorenzo Casaburi and Rocco Macchavello
2019	Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India	Karthik Muralidharan, Abhijeet Singh and Alejandro Ganimian
2019	Does Diversity Matter for Health? Experimental Evidence from Oakland	Marcella Alsan, Owen Garrick and Grant Graziani
2019	Identification of and Correction for Publication Bias	Isaiah Andrews and Maximilian Kasy
2019	Making Moves Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings	Adnan Khan, Asim Khwaja and Benjamin Olken
2019	Parents’ Beliefs about their Children’s Academic Ability: Implications for Educational Investments	Rebecca Dizon-Ross
2019	Paying for Kidneys? A Randomized Survey and Choice Experiment	Julio Elias, Nicola Lacetera and Mario Macis
2019	The Dynamics of Discrimination: Theory and Evidence	Aislinn Bohren, Alex Imas and Michael Rosenberg
2019	The Impact of Media Censorship: 1984 or Brave New World?	Yuyu Chen and David Yang
2020	Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services	Nava Ashraf, Oriana Bandiera, Edward Davenport and Scott Lee
2020	Outsourcing Education: Experimental Evidence from Liberia	Mauricio Romero, Justin Sandefur, and Wayne Aaron Sandholtz
2020	Targeted Debt Relief and the Origins of Financial Distress: Experimental Evidence from Distressed Credit Card Borrowers	Will Dobbie and Jae Song
2020	The Welfare Effects of Social Media	Hunt Allcott, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow
2020	A Theory of Experimenters: Robustness, Randomization and Balance	Abhijit Banerjee, Sylvain Chassang, Sergio Montero and Erik Snowberg