Improving transparency in observational social science research: A pre-analysis plan approach

Fiona Burlig*

September 25, 2015

Abstract

Modern social science research is increasingly data-driven, rigorous, and policy-relevant as a result of a variety of important methodological innovations. This same research is at risk of being devalued because of the increased evidence of the prevalence of data mining, publication bias, and other problematic research practices and norms. In order to combat these problematic norms, there is growing interest in transparency practices in the social sciences. At present, the bulk of this work is centered around randomized controlled trials. These experiments, however, constitute only a small fraction of all social science research. As a result, in order to improve research transparency in the social sciences, it is crucial to develop ways of improving research practices in non-experimental work as well. In this paper, I propose a method for using pre-analysis plans in observational research. I suggest three scenarios in which pre-analysis plans can be credibly applied to non-experimental research designs: studies using future events or data that has not yet been collected; studies using confidential data; and studies using data that are not available to the public without purchase. I provide guidelines for reliably demonstrating that researchers have not had access to data before submitting these plans. I then outline suggested contents for observational pre-analysis plans, and highlight areas where these plans should deviate from pre-analysis plans designed for experimental research.

JEL codes: A11, B41, C13

Preliminary and incomplete. Do not cite without permission.

*University of California, Berkeley: Department of Agricultural and Resource Economics and Energy Institute at Haas. Email: fiona.burlig@berkeley.edu. This project has benefited greatly from conversations with Dave Rapson and Ted Miguel. I am thankful to Patrick Baylis and Garret Christensen for useful feedback. All remaining errors are my own. In an attempt to practice what I preach about research transparency, I am making this paper available on the Open Science Framework here. I am generously supported by the National Science Foundation’s Graduate Research Fellowship Program.
1 Introduction

Modern social science is the beneficiary of several decades of important methodological innovations. Our research is increasingly rigorous and policy-relevant. Alongside these gains, there is a growing understanding that transparency is critical to performing good research in the social sciences. Several recent studies have empirically demonstrated the need for improvements in social science practices. Fraud has been documented in social psychology (Simonsohn (2013)) and in political science (Broockman, Kalla, and Aronow (2015)) but there is reason for concern beyond these particularly egregious violations of research ethics. Recent research suggests that by employing data mining, the practice of choosing and subsequently reporting outcome variables and/or empirical specifications based on what “works” and returns statistically significant results after testing tens or hundreds of dependent variables and estimating equations, researchers could use the same datasets and experiments to tell wildly different stories (Simmons, Nelson, and Simonsohn (2011), Casey, Glennerster, and Miguel (2012), Humphreys, de la Sierra, and van der Windt (2013)). Indeed, across disciplines, there is evidence of strange patterns of statistical significance in published papers - there are more studies with p-values just under 0.05 than we expect (Ioannidis (2005), Gerber and Malhotra (2008), Masicampo and Lalande (2012), Brodeur et al. (forthcoming), Franco, Malhotra, and Simonovits (2014)). These patterns are suggestive of researcher manipulation, the “file drawer problem” (Rosenthal (1979)), and/or of publication bias, and raise concerns about the validity of social science research.

In the wake of this evidence, there is a growing body of work that advances transparency practices in the social sciences (Miguel et al. (2014), Nosek et al. (2015)). There are a variety of tools that have been developed for transparent social science research; Christensen and Soderberg (2015) provides an overview of a variety of best practices in this area. One of the most popular approaches for increasing credibility in social science research is the pre-analysis plan, wherein researchers pre-specify analysis that they will do before having access to the data that they will use (Casey, Glennerster, and Miguel (2012), Finkelstein et al. (2012), Humphreys, de la Sierra, and van der Windt (2013)). This allows researchers to credibly tie our hands against data and specification mining. There has been an explosion of pre-registration - since its inception in 2013, the American Economic Association’s randomized controlled trial registry has grown to include 447 studies. This registry reflects the current state of pre-registration in economics: it only accepts pre-analysis plans for randomized controlled trials (RCTs). Up to this point, the vast majority of pre-registered research in economics has been for experimental designs. There are good reasons for this: researchers can easily register a pre-analysis plan for an experiment before the data collection phase, ensuring that data access has not been possible; when running experiments, researchers know exactly what data they will collect, making pre-specification much easier; and pre-analysis plans are a tool that social scientists have borrowed from the medical literature, where randomized controlled trials are the standard.

The advances in pre-registration for RCTs in the social sciences are promising, but it is important to recognize that even today, only a small fraction of published papers in economics are experiments: in 2010, only 3% of papers in top journals were field RCTs (Card, DellaVigna, and Malmendier (2011)). Even if RCTs made up the majority of publications in top journals, we might worry that pre-registration of experiments is not extremely valuable (Coffman and Niederle (2015), Olken (2015)), in part because there are strong norms around the presentation of results from papers using randomized controlled trials, so researchers inherently have less freedom than we might expect. Thus far, there have been very few pre-analysis plans registered in economics for observational studies.

---

1 Angrist and Pischke (2010) provides a discussion of some of these advances.
2 A recent blog post by Simonsohn suggests that several of these studies are even underestimating the degree of “p-hacking” present in their various literatures (Simonsohn (2015)).
3 It is important to note that we need not assume that researchers have nefarious intentions to be worried about these types of problems.
4 The AEA registry can be found online here: https://www.socialscienceregistry.org/; I last accessed the registry to check the number of studies on September 11, 2015. Disappointingly, only 105 of these studies are accompanied by pre-analysis plans.
5 The GoBiFo experiment detailed in Casey, Glennerster, and Miguel (2012) is a counterpoint to this - even if we expect a comparison
One famous example is Neumark (1999, 2001).\textsuperscript{6} Dal-Ré et al. (2014) advocate for the use of pre-registration for observational research, but do not suggest ways to overcome the obvious credibility problems inherent in doing this. Miguel et al. (2014) acknowledges these issues, and stresses the importance of developing transparency methods for observational work that uses existing data. To the best of my knowledge, Burlig (2015) provides the only example of non-prospective pre-specified research in economics.\textsuperscript{7}

If we want to improve transparency in social science research, it is critical that we find ways to add credibility to non-experimental work. In this paper, I discuss ways to use pre-analysis plans in observational research. The remainder of this paper is organized as follows: Section 2 provides an overview of pre-analysis plans. Section 3 discusses strategies for credibly pre-registering observational research. Section 4 provides a road-map for the contents of pre-analysis plans for observational research. Section 5 describes what comes after pre-analysis plans. Section 6 concludes.

2 What is a pre-analysis plan?

A pre-analysis plan is, at its core, a tool for producing research that transparently tests hypotheses in keeping with the scientific method.\textsuperscript{8} In practice, pre-analysis plans are documents that prospectively outline planned analyses before a research has an opportunity to explore the data that will be used. Importantly, pre-analysis plans are only beneficial from a transparency perspective when they are submitted before the researcher has seen the final analysis data. Submitting a pre-analysis plan to, for example, specify the use of a publicly available dataset from a previous paper for new analysis is hard to justify on these grounds, because a researcher submitting such a pre-analysis plan cannot credibly defend not having seen the data before pre-registering. As a result, the timing of a pre-analysis plan relative to data access is crucial. In order to ensure that these plans are not edited after the research acquires said data, they are submitted to and stored by a third-party registry, where each plan is given a time stamp.\textsuperscript{9} At present, pre-registration is common practice in medical trials.\textsuperscript{10} There is growing interest in pre-registration of experimental work by social scientists, but the practice remains limited.

2.1 Current norms

There are no established rules in the social sciences for exactly what a pre-analysis plan must contain.\textsuperscript{11} That said, there exist several useful overviews of suggested pre-analysis plan contents, all written with RCTs in mind.

\textsuperscript{6}Levine (2001) wrote the accompanying editor’s letter. Campolieti, Gunderson, and Riddell (2006) is a later paper in the same journal which uses the same empirical specification that was pre-specified by Neumark - though these researchers were unable to truly pre-specify their analysis, because their data were publicly available at the time the research was conducted.

\textsuperscript{7}At the time of the first draft of this document, Burlig (2015) has not yet been submitted to a registry. It is work in progress, and will hopefully be submitted shortly.

\textsuperscript{8}The scientific method is an ongoing process for evaluating hypotheses. It usually begins with observations about the world, developing hypotheses about these observations, creating testable predictions from these hypotheses, then gathering data or running experiments to carry out these tests, and then refining these hypotheses. Pre-analysis plans aid social science researchers in this process by ensuring that researchers begin with hypotheses and then test them, rather than simultaneously performing data analysis and generating or adjusting hypotheses.

\textsuperscript{9}Prominent registries in the social sciences include the American Economic Association’s registry: \url{http://www.socialscienceregistry.org}, the Experiments in Governance and Politics registry: \url{http://egap.org/design-registration/}, the Registry for International Development Impact Evaluations: \url{http://ridie.3ieimpact.org/}, and the Open Science Framework: \url{https://osf.io/}. In order to preserve researcher intellectual property rights, most registries offer the option to keep plans “gated”, or private for some length of time after they are submitted.

\textsuperscript{10}Many top medical journals require pre-registration for publication.

\textsuperscript{11}The FDA’s guidelines for medical studies can be found \url{here}. 

2
Olken (2015) recommends including a primary outcome variable (or variables, with discussion of corrections for multiple inference); clear variable definitions that describe the creation of the variables from the raw data; rules for how the researcher will decide to exclude certain observations and handle missing data; empirical specifications (including covariates); heterogeneity analyses; and other issues of data monitoring. Glennerster and Takavarashi (2013) recommends describing the outcome measures, including a separation between main and secondary outcomes; families that will be used for mean effects analysis; a description of subgroup analysis; the expected direction for one-sided statistical tests; and the main empirical specification. McKenzie (2012) recommends including a discussion of the sample; a description of the data sources; the hypotheses to be tested; a description of variable construction; the estimating equation; plans for dealing with multiple hypothesis testing; plans for coping with survey attrition; addressing variables with limited variation; and the model to be tested.

Several overall themes emerge. It is important for pre-analysis plans to include the outcome variables (and clear definitions of these variables) that will be used in the analysis, and better still for the outcome variables to be declared primary or secondary. Pre-analysis plans should also include the estimating equation to be used; proposed heterogeneity analyses; and (where applicable) multiple testing corrections. Ideally, these plans will also include methods for dealing with outliers, missing data, and other problematic observations. In what follows, I discuss the various aspects of pre-analysis plans in the context of observational research.

3 Credible pre-registration for observational research

For a pre-analysis plan to be useful, it is critical that a researcher can demonstrate that she has not had access to the final dataset before the pre-analysis plan is registered. This is the major barrier to using pre-analysis plans with observational data. In an RCT, a researcher can pre-register analyses while the experiment is in the field, or even up until the final survey data is entered electronically. With observational data, credibly demonstrating that a researcher did not have access to data before writing her pre-analysis plan is much more difficult. In most observational work, the use of pre-analysis plans will remain limited, but there are three key cases where observational research lends itself to pre-registration: prospective research on future events or datasets; research using confidential data; and research using data that is available for purchase. As with pre-registration of RCTs, the common theme is that there must be restrictions in place that allow the researcher to credibly deny having had access to the data before submitting her pre-analysis plan.

3.1 Prospective data

First, it is easy to imagine credibly pre-registering observational research using future events. In fact, the oldest known pre-analysis plan in economics did just this (Neumark (1999, 2001)). As a modern example, imagine I am interested in the effects of 2016 electoral outcomes on American carbon policy.12 As long as I submit my pre-analysis plan before the election, I can very credibly say that I had not seen the outcome data before submitting the pre-analysis plan - since it did not even exist at the time of registration. This credibility context is very similar to the RCT case.13 In this setting, a researcher could even collect her own data, giving her control over variables that she would not have in other observational pre-specification scenarios. A closely related scenario is one in which a researcher pre-registers observational research using data that will be published (or collected) in the future, even though the data generating process occurred before the pre-analysis plan was submitted. Suppose I am interested in

---

12One could imagine doing this with a regression discontinuity design with vote margin as the running variable, and trying to estimate the causal effect of democratic governors on the likelihood of compliance with the EPA’s Clean Power Plan via a mass-based standard.

13Note, however, that there are still pre-specification issues in the later sections that look slightly different in the observational context than in the RCT context.
the effect of Hurricane Katrina on migration outcomes, and suppose further that there was an American migration survey that occurred in 2006, but whose publication is scheduled for 2016. In this case, as long as I submit my pre-analysis plan before the dataset is released, I can ensure my credibility as though I were pre-specifying an RCT. Many datasets lend themselves to prospective pre-specification. For example, the American Census Bureau publishes data release schedules; the World Bank often announces when datasets are in the field; and IPUMS provides upcoming data release dates. There are many more agencies around the world that announce their current data collection and future release activities. Many of these agencies also provide information about what will be included in these datasets, making prospective research an attractive choice for the observational researcher.

### 3.2 Confidential data

Another scenario in which researchers can use pre-analysis plans in an observational setting is that of confidential data. Social science researchers are increasingly using confidential data from corporations and governments under non-disclosure agreements. In these cases, researchers can also credibly pre-specified analyses. The easiest way to demonstrate that a researcher has not seen confidential data before submitting a pre-analysis plan is to submit said plan, and then to include as an addendum a copy of a signed and dated non-disclosure agreement (dated after the pre-registration, of course), or a formal communication from the owner of the confidential data that acknowledges the date from which the researcher has access to the data. This path to pre-registration in the observational setting will likely impose the least burden on researchers: it is almost always the case that in order to access such confidential data, researchers must submit project proposals to the organization that controls the data. These proposals are essentially the outline of pre-analysis plans, so while there may be additional work involved in meeting the (pseudo) requirements of a pre-analysis plan, researchers entering confidential data sharing agreements have already made some headway on this front. One could imagine employing this method when using confidential census microdata (see Isen, Rossin-Slater, and Walker (2013) and Currie et al. (2015) for examples of observational research using this type of data); when using data from a healthcare provider, utility, or other corporation (see Mas and Moretti (2009), Handel (2013), and Ito (2014)) for papers using these types of data; or any other type of data controlled by a non-disclosure agreement. Importantly, these data sources must be controlled by a third party - it is hard to credibly document that a researcher has not accessed her own non-anonymized (and therefore confidential) survey data.

### 3.3 Purchased data

Finally, a researcher can pre-specify an observational analysis when the datasets they are using are not available to the public without purchase. In this case, a researcher can pre-register her analysis before buying the data that she will use, and should append the pre-analysis plan with a copy of the data purchase receipt, dated after

---

14 This might seem like an unreasonably long time delay in the United States, though it is not unlikely in the developing world.

15 The 2012 Economic Census schedule, for example, can be found [here](#).

16 See the data collection timeline for the Living Standards Measurement Survey - Integrated Surveys on Agriculture, as of June 24, 2011, [here](#).

17 These can be found [here](#).


19 In principle, any time where there is a significant barrier to accessing the necessary data, researchers can pre-specify observational research. The researcher must, however, be able to credibly demonstrate that she has not had access to the data. Simply having to fill out an online form prior to downloading data, as is required to access World Bank Living Standards Measurement Surveys, for example, would not be sufficient, because these data are easy to access.
the original plan was submitted.\textsuperscript{20} In these cases, it is best to also provide evidence that other researchers have not shared these data with you before the pre-analysis plan was submitted, since a truly nefarious researcher could simply mine a friend’s copy of these data before pre-specifying her own analysis. Wherever possible, including a copy of a the grant budget that provides for the data purchase is also useful for demonstrating that the data were purchased using official funds, especially if the data are expensive and unlikely to be purchased out of pocket.

Importantly, it is not possible to credibly pre-register all observational research projects. That said, many researchers find themselves in one of the above three scenarios, so beginning to employ pre-registration in conjunction with observational research seems very possible.

4 What belongs in an observational pre-analysis plan?

Armed with a credible way of pre-registering a research design, we can now turn to the pre-analysis plan itself. In most respects, a pre-analysis plan for an observational study will look very similar to that of an RCT. There are several key differences between observational and non-observational pre-analysis plans, especially when the observational design uses data that are not collected by the researcher herself. Below, I discuss the key features of pre-analysis plans, and describe how each of these features should or should not deviate from the standard experimental pre-analysis plan when used for observational research.

4.1 Hypotheses

In every pre-analysis plan, it is important to clearly outline the hypotheses that a researcher will test. As Casey, Glennerster, and Miguel (2012) point out, it is helpful to distinguish between main and alternative hypotheses in a pre-analysis plan. As social scientists, we are in the business of testing theories about human behavior. In an ideal world, pre-analysis plans will describe the model they are testing, as well as the implications of the rejection of or failure to reject certain hypotheses for these models. This is a high bar, and should certainly not be a requirement for every pre-analysis plan, but a careful and pre-specified theoretical framework can help a researcher to guard against accusations of using results to rationalize a model ex-post, and furthermore to advance the body of knowledge in social science. In a similar vein, it is also beneficial to pre-specify tests for mechanisms. In keeping with the conditionality discussion above, a researcher can specify a main model, and then pre-specify different sets of tests to explain various results. A description of how these tests relates to the main hypotheses is useful. Pre-specifying mechanisms is not specific to observational research. In general, pre-specified hypotheses need not look any different in observational research than in pre-analysis plans for RCTs.

4.2 Data

As with an experimental pre-analysis plan, providing a thorough description of the data that will be used is an important part of an observational pre-analysis plan. In both cases, the pre-analysis plan should include a discussion of how the outcome and treatment variables are constructed, and furthermore, which outcome variables will be used to test each hypothesis. In an observational pre-analysis plan where a researcher will not collect her own data, this can be challenging, but it is critical for motivating the chosen research design. Wherever possible,\textsuperscript{20} In Burlig (2015), I use a variant of this method, whereby I purchased the data before submitting the pre-analysis plan due to logistical constraints, but had the dataset shipped directly to my home department’s administrative offices without accessing it. I will append my pre-analysis plan with a signed affidavit from the department administrator certifying that the data were released to me on a given date after the pre-analysis plan was submitted.
an observational pre-analysis plan should include a discussion of the variables that will be in the dataset. These
plans should also include information on how the data was (or is to be) collected, and crucially, a discussion of
how the researcher has been kept from seeing the data before the pre-analysis plan has been filed. Appendices to
observational pre-analysis plans with supporting documentation for credibility purposes are useful.

4.3 Research designs and empirics

Without the benefit of true randomization, observational studies rely on research designs to provide credible
estimates of the causal effect of a given “treatment” on outcomes of interest. Researchers using quasi-experimental
approaches have a great deal of freedom over which design to use, how to implement these designs, and which
alternative specifications and robustness checks to implement. Because of the vast menu of choices available to
researchers doing observational work, pre-specification is especially valuable in the non-experimental context. A
typical experimental pre-analysis plan pre-specifies (some variation on)

\[ Y_i = \alpha + \beta \text{Treat}_i + \varepsilon_i \]

where \( Y_i \) is the outcome of interest for unit \( i \), \( \text{Treat}_i \) is a treatment indicator, equal to one if unit \( i \) is treated
and zero otherwise, and \( \varepsilon_i \) is an idiosyncratic error term. \(^{21}\) There is a much greater diversity of observational
studies, so a proper research design section of an observational pre-analysis plan should be more detailed. I
recommend including information about the primary research design; building in conditional statements; discussing
heterogeneity analyses; outlining robustness checks; and specifying statistical inference in a research design section
of an observational pre-analysis plan.

4.3.1 Primary research design

In particular, an observational pre-analysis plan should be sure to describe the research design in question
in detail. It should include a discussion of the chosen identification strategy, in which the researcher argues that
this design exploits exogenous variation in the right-hand-side variable. The researcher should make sure to pre-
specify all important decisions that go into creating a proper research design. As an example, a researcher using
a nearest-neighbor matching estimator should specify which variables are being used in the match. A researcher
using an instrumental variables approach should clearly define the instrument. If applicable, researchers should be
clear about which units in the dataset constitute the “treatment” group, and which constitute the “control” group.
Though pre-specifying these choices ties the researcher’s hands, the benefits are large and important: readers of
the eventual work can take the results at face value, rather than worrying about specification or data mining.

Researchers should also pre-specify tests necessary to ensure the credibility of their research design. In the
RCT context, it is useful to pre-specify which variables are used for a baseline covariate balance table (Olken
(2015)). In a difference-in-difference context, it is useful to pre-specify tests for parallel trends. In regression
discontinuity designs, covariate smoothness tests and McCrary tests should be pre-specified. Wherever possible,
these should include conditional statements: if one of these tests is “failed”, does the researcher have recourse? If
so, she should pre-specify the appropriate changes to the analysis.

It is critical that the pre-analysis plan also includes a description of the estimator that will be used (this is
ideally presented in the form of an estimating equation). The pre-analysis plan should indicate which estimating
equation represents the preferred specification. If the research design is not “naive” OLS, it may be useful to pre-
specify both the (likely biased) OLS estimator and the preferred estimator, making sure to designate the preferred

\(^{21}\)We can also imagine time-varying treatment or treatment at a cluster level, etc.
estimator as such. Researchers with multiple methods for estimation should be very clear about which method is being used to test each hypothesis. Furthermore, a description of how we should interpret the results is a useful inclusion. Pre-analysis plans should clearly indicate which, if any, control variables will be used in the main analysis.

4.3.2 Conditionality

Pre-analysis plans improve the transparency and credibility of social science research by forcing scholars to commit to a set of analyses before seeing the data. Importantly, however, we can use conditional statements in these pre-analysis plans in useful ways. Conditionality can be beneficial for pre-analysis plans for experiments as well as for observational work.

Olken (2015) and McKenzie (2012) suggest one type of conditionality that can be pre-specified: the decision about what to do with problematic observations. How to clean data is an important decision, and if two researchers starting with the same raw dataset make different cleaning choices, they can end up with different outcomes. Pre-specification can help with this problem. A research can say up front, “I will drop observations if their monthly income values are over $100,000.” This builds in some conditionality (if no observations fall into this category, none get dropped) while also adding credibility by pre-specifying a decision rule.

Some researchers have even suggested the use of “data-adaptive” methods, whereby the analysis is updated according to what is found in the data. A simple example of this is applicable both to experimental and observational research designs: researchers can pre-specify which variables in their dataset to use in a baseline balance test between the treatment group and the control group. We can then go further and additionally pre-specify that if any variable turns out to have a statistically significant imbalance, we will include that variable as a control in all subsequent analysis. This again affords a researcher some flexibility to deal with possible problems of bias while maintaining the credibility that comes with pre-specification. In a more complex example, the Sustainable East Africa Research in Community Health (SEARCH) RCT aimed at reducing the prevalence of HIV/AIDS in Africa used a pre-specified machine learning algorithm to determine which control variables to include in the final analysis in order to increase statistical precision (Balzer et al. (2015)). This was an RCT-based example, but one can imagine doing something similar in an observational setting. A cheeky reader might point out that taken to the extreme, the conditionality argument would suggest a decision rule of “report the results that are statistically significant,” thus rendering the pre-analysis plan moot. In fact, this can be a legitimate (and transparent) decision rule, provided that the researcher also pre-specifies the correct adjustment for multiple inference and carries these out in her tests for statistical significance.

Conditionality becomes especially important in the case of observational work where a researcher is not collecting her own data. Consider, for example, the case of Burlig (2015). In this pre-analysis plan, I specify a quasi-experimental research design using secondary data from India. In order to carry out my research design completely, I need to be able to match village-level data to an outside dataset. It is likely that this will be possible, but require a location directory to de-anonymize villages that is not always included with the data. I cannot say ex-ante, without having seen the data, whether this match will be possible, so instead, I build in conditionality into my pre-analysis plan. If the match turns out not to be possible, I describe how I will change my analysis to accommodate this challenge. Alternatively, suppose that a variable that was in the first two waves of a dataset does not appear in the third (for which the pre-analysis plan was written). We can build in conditionality to adjust the analysis accordingly, which allows us some pre-specified flexibility. Importantly, this means that if a given condition on a dataset fails, we can still do credible and transparent research - the alternative would be to deviate completely from the pre-analysis plan. Some extra effort put into thinking about what types of challenges that might arise with a given observational dataset can help protect researcher credibility if the data disappoints.
4.3.3 Heterogeneity

We are often interested in heterogeneous effects and mechanisms. As with an experimental pre-analysis plan, researchers should pre-specify analyses that are supplemental to the main analysis. Examples of this might include (but need not be limited to) subsetting the sample population by baseline covariates and re-running the analysis, including interaction terms in regression analysis, or performing a test using only the later years of the data. Pre-analysis plans should include descriptions of the estimation procedure for heterogeneity analysis and mechanism tests. Pre-specification of these tests is especially valuable when accompanied with a discussion of what results might mean. This type of pre-specified additional analysis could easily find a place in experimental research as well.

4.3.4 Robustness checks

In order to be fully transparent about her results, a researcher may decide to pre-specify robustness checks or variants on the main tests. Robustness checks should be clearly separated from the main analysis. As with the heterogeneity and mechanism tests above, these checks should be thoroughly described and accompanied with estimating equations wherever possible. Discussions of what results on a particular test mean can be useful as well. Note that the pre-specification of robustness checks is not particular to observational designs.

4.4 Inference

Statistical inference is a key part of pre-analysis plans, both for observational and experimental designs. Each specified hypothesis test should have an accompanying inference procedure. This has the potential to increase the statistical power available to a researcher: using a pre-analysis plan allows for the credible use of one-sided statistical tests. Beyond the type of test used, researchers should also pre-specify how they will be calculating their standard errors. If they will be using Huber-White robust standard errors, Conley HAC standard errors, or other error adjustments such as clustering, they should be clear about all of the relevant details of these methods. A pre-analysis plan that specifies using a two-way clustered standard error estimator, for example, should be sure to specify the clustering variables. As above, it is perfectly reasonable to specify multiple standard error adjustment methods, and present the results of each.

Pre-analysis plans that involve testing multiple related hypotheses should include a description of the multiple testing adjustment(s) that will be used. Common methods for multiple inference include the False Discovery Rate and the Family-wise Error Rate correction. Another option for researchers with many related outcomes is to aggregate these outcomes into indices, which reduces the number of hypotheses tests that are performed. Both Casey, Glennerster, and Miguel (2012) and Finkelstein et al. (2012) discuss using these multiple inference methods in pre-analysis plans. A major benefit of pre-specification is that it allows researchers to credibly conduct multiple inference. In the absence of a pre-analysis plan, it is impossible to know whether the final multiple inference corrected analyses in a paper actually adjust for all of the tests a researcher has performed. In addition, pre-specification allows researchers to credibly group outcomes before seeing the effects of these choices on statistical significance levels.

5 Registration and ex post analysis

After writing a pre-analysis plan, researchers need to make sure that they submit it to a registry before accessing the data. When necessary, addendums verifying the date of data access should be added. At present, the
American Economic Association registry does not accept pre-registration of observational work, but the Open Science Framework (OSF), Registry for International Development Impact Evaluations (RIDIE), and the Experiments in Governance and Politics (EGAP) registry all do. The RIDIE registry is aimed at international development research; and the EGAP registry is designed to host research in political science. Registering with the OSF is available for any type of work; thus far, psychology is the most common type of pre-registered research on the OSF. It is important that pre-analysis plans conform to the requirements of a researcher’s chosen registry. When choosing a registry, researchers will want to be aware of the registry’s policies towards gating (keeping pre-analysis plans private) and submission of additional documents.\footnote{The current guidelines for the EGAP registry can be found online. RIDIE and OSF do not provide strict guidelines, but do have an online FAQ sections.}

The next step is to do the actual data analysis, and to report the results. As pre-registration of observational research is at the cutting edge of the research frontier, the norms around the use of pre-registration are still evolving. I cannot predict exactly what will happen, and how these plans will be used in the future, but there are several points worth noting about actually conducting research that has been pre-specified. Research with an accompanying pre-analysis plan should make it easy for readers to find this pre-analysis plan. It is important that everything that was pre-specified in the plan appears in the paper or the appendix to the paper that is eventually produced. Equally important, it is quite likely that pre-analysis plans will omit things or include mistakes. When this happens, researchers should present all of the pre-specified analyses, and should also include the forgotten or corrected analyses as well. In order to ensure transparency, a researcher simply needs to clearly indicate which results were not pre-specified.\footnote{My preferred method of doing this is from Finkelstein et al. (2011) - they assign a special character (ˆ) to non-pre-specified results, and mention this at the beginning of the text, allowing readers to search (visually or electronically) for these symbols. In the final version of the text (Finkelstein et al. (2012)), all of the non-pre-registered tests can be found in the appendix.} All of this holds equally for experimental and non-experimental pre-specified research.

6 Conclusion

Social science has made great strides over the last few decades. Our research is more data-driven and increasingly methodologically sound. It is critical that we accompany these improvements with credible, transparent research that is born out of the principles of the scientific method. The vast majority of research transparency innovations in the social sciences have been directed towards and designed for randomized controlled trials. These advances are important, and are already increasing the credibility of experimental social science. At the same time, experimental research remains a relatively small fraction of total social science output. In order to improve the overall transparency of social science research, it is crucial that we bring these advances to non-experimental work as well.

In this paper, I propose a method for using pre-analysis plans in observational research. I suggest three scenarios in which pre-analysis plans can be credibly applied to non-experimental research designs. I provide guidelines for reliably demonstrating that researchers conducting non-experimental research have not had access to the analysis dataset before submitting their pre-analysis plan. I then outline what observational pre-analysis plans should include, and describe in detail where these plans should differ from their experimental counterparts. These plans cannot be used in every situation, but it is my strong belief that there are a variety of settings in modern social science research in which they can be beneficial. I am optimistic that these plans will help to improve the credibility of non-experimental work. I hope and expect that we can and should continue to develop innovative methods for increasing the transparency of observational social science research.
References


