Draft proposal for a Pilot Registry for Political Science (PREPS)

EGAP VIII Steering Group[1]

October 3, 2012

ABSTRACT

There is growing interest in establishing a registry for research in political science as a method to mitigate problems of reporting and publication bias. However while the goals of a registry are clear, the details of an appropriate design are not. In particular an effective registry system would need to be based on a shared set of principles, be simple to use for researchers seeking to register designs, involve journals in a way that that helps provide incentives for researchers while not increasing the workload of journals unduly, and ensure that registration norms do not stifle innovation. We describe one set of arrangements to meet these goals, which we hope to serve as a starting point for a discussion to establish a system in the field.

Contents

[1] Michael Findley, Macartan Humphreys, and Jeremy Weinstein.

1 Overview

We outline a scheme for a pilot registry for research designs in political science. We hope that this document will serve as a starting point for journals and researchers to agree on a model to be piloted in the coming years.

The primary goal of a political science registry is to enhance transparency with respect to performance and reporting (De Angelis, et al 2003), which has many scientific and ethical benefits. This effort thus supplements other initiatives, including the movement to make replication data available, by encouraging scholars to avoid data fishing. The results of individual studies should then be less biased and overall publication bias on political science topics should decline.

After a long incubation period, research registration is now the norm in Medicine. Registries have been supported or mandated by a host of individuals, organizations, and even governments (Dickersin and Rennie 2003). But registration did not become conventional until the leading medical journals endorsed registration. Efforts are now also underway in Economics and international development practice to establish trial registries. The principle of registration is now also endorsed by the APSA experimental section which has created a website space for unmoderated posting of designs; since early 2011 EGAP has also been accepting registered designs at http://e-gap.org/design-registration.

However despite the growing interest we expect that establishing registration norms in political science will not be easy, in particular in light of real concerns that, if not implemented correctly, a registry could be onerous and simply stifle creativity and innovation. In light of this there is a need to identify basic principles regarding what should be registered, how constrained research should be by past registration, and what kind of recognition should registered research receive. As we detail below, we believe the task of instituting a registry will rely primarily on investigators to register and on journal editors whose recognition of registration can provide incentives and structure. Research funders may also play an important role in encouraging design registration.

We propose establishing a system to be led by journals to introduce a *pilot* registry, which investigators can elect to use. To this end we encourage journal editors to consider some concrete steps to advance the process of registration:

a. Agree on a common set of principles of a registration **system** with specified roles for authors, journals, registry staff and others (See section 3)
b. Agree on a core set of minimal mandatory **fields** for minimal registration and guidelines for determining level of registry **compliance** (See section 4)
c. **Publicize** participation in the pilot to authors and readers

2 Short Motivation

Both scholarly and ethical concerns motivate a registration process in political science. We discuss a few of them briefly here, but refer those interested in a more extensive treatment to the works cited.

Scholarship: Publication bias is documented in many fields, including political science (e.g., Gerber and Malhotra 2008; Gerber, et al. 2010; Glewwe and Kremer 2006; Dickersin 1990; Ioannidis 1998). While the extent of current data fishing is likely unknowable, existing publication practices provide little ability for

scholars to understand it better. The review process allows referees to probe this to a minimal extent. Requirements to make replication data available also help considerably, but authors still have substantial latitude about which analysis data they provide. For example, explanatory variables not significantly associated with an outcome may be excluded from the replication data set if authors did not use the variable in their analysis, making it difficult to know which variables the authors used and did not use in winnowing down the set of statistically significant relationships.

The fundamental problem of reporting and publication bias is that if the decision to report and publish is a function of results, and if the results of a study have a random component, then reported results will not be right on average. By separating the reporting decision from the outcomes of analysis, registration may provide strong incentives to avoid such biases. Simes (1986) offers evidence that using a trial registry should produce estimates free of publication bias.

Ethics. There are also ethical reasons to take reporting and publication biases seriously. Subjects often participate in research precisely because they believe the research will affect knowledge and associated policy (Dickersin and Rennie 2003; 517). If researchers do not report the results of their studies, or distort them by mining their data, then researchers are not acting in good faith with respect to their subjects. Moreover insofar as political science research informs decision-making on policy (as it does in spheres such as security and international development), publication bias can result not only in flawed research, but also in harm to populations more widely (Dickersin and Rennie 2003; Rasmussen 2011).

In recent years, scholars have begun to call for a registration process in the social sciences (Bose 2010; Duflo et al. 2007; Humphreys et al. 2012; Rasmussen et al. 2011) that will increase transparency and lead to less publication bias.

Besides the public benefits from registration there may also be professional benefits to individual researchers:

1. **Meta-analysis:** The systematic data generated through a registration process can be used to conduct meta-analyses to help establish external validity and situate findings within a population of results (Banerjee and Duflo 2009; Cohen and Easterly 2009).
2. **Credibility:** Registration may serve to improve the credibility of the research in the eyes of reviewers and editors. Indeed, a registration process may signal (appropriately) that the investigators were transparent and responsible at every step of the research process. This transparency may provide justification to authors for preserving the integrity of the study, even when skeptical reviewers ask for non-trivial modifications.
3. **Documentation:** The registration could also be a means of providing evidence to department and evaluation committees of research productivity, especially, perhaps, when it ultimately does not get published because of insignificant findings.
4. **Project Integrity:** Registration could also provide investigators with a basis for arguing against referee comments that take the research astray from the intended purpose.
5. **Idea Protection:** A design registration may provide early confirmation and an official registry citation that could help insulate investigators from being scooped on various design ideas.

3 Description of Proposed Registry

3.1 Key Principles of Design

We propose a pilot registry scheme with the following eleven features:

1. **Compatible with other initiatives:** Other initiatives to establish a registry for social sciences are ongoing including the RIDIE initiative led by 3IE and an initiative being piloted by AEA / J-PAL. While the economics and development registries are likely to have some common features with a political science registry, their practical needs and precise reporting standards will have key differences. Thus, rather than seeking for a common registry across fields, interoperability between registries will be important. The registry should seek to have, at a minimum, common mandatory fields with other initiatives and ideally a common interface.
2. **Voluntary and Nonbinding:** During the pilot period, registration would be **voluntary and nonbinding**. The decision whether to make registration mandatory for some kinds of research is one that we recommend be made in light of what is learned from the pilot.
3. **Focus on Prospective Research:** The registry will be open to **experimental** and **observational** work that employs **prospective** designs – i.e., designs for which *outcomes* have not yet been realized.[2]
4. **Open with respect to Subject Matter**. Registration will be open to any subfield in political science, public policy, and associated subfields. Importantly, the goal of the registry is not to govern the types of questions people ask but rather the approach for ensuring proper disclosure of methods.
5. **Core mandatory elements plus flexible forms.** Registry fields in the pilot phase will be flexible enough to gather a set of required information without imposing onerous costs on authors while also providing the opportunity to input additional optional information. In addition to flexibility upon initial submission, investigators will be able to update their registration as the research develops (a versioning system will track updates).
6. **Submissions Subject to Review**: Upon registering a design, registry staff would review it, and admit the design to the registry conditional on satisfying basic criteria. The initial review would **not screen based on quality of ideas** or methods, but rather on whether investigators are sufficiently clear and comprehensive no matter what they propose to do.
7. **Publicly Searchable.** The registry will normally allow public access to registered research designs.
8. **Sunset Provision for Private Registration**. Although the norm should be publicly available registered designs, in some cases private registration for a limited time might be appropriate. We propose that private registration (with registered designs viewed only by staff, funders, editors, and reviewers) be permitted subject to a three-year limit, which may be negotiated for exceptional cases. All authors would supply very basic public data (title, author), but could choose to keep detailed information private.
9. **Deviations Acknowledged**. If final analyses differ from designs then researchers should highlight the deviations.
10. **Summary Results Recorded within Registry for Published and Unpublished analyses.** The registry should seek to maintain a record of results from research (as reported to the registry by researchers) *whether or not these results are published*. We propose a 3-year lag between registration and the reporting of results to the registry, in the event results do not appear before that time in print

[2] In principle, a much larger set of studies could be included, but if outcomes have already been realized then it is difficult to be sure that a registered design is not itself subject to a selection bias. The registry could nonetheless be designed to allow anyone to post a design, even if doing purely observational work with retrospective data.

somewhere. Some incentive structure may need to be developed to encourage reporting of otherwise unpublished work.

11. **Recognized by Journals:** The registry should be recognized by a core set of journals that can seek a statement from the registry indicating whether journal submissions are registry compliant or not. Journals should indicate in publications whether a submission has been certified as compliant.

3.2 The Registration –Publication Cycle

We envision a publication cycle as shown in Figure 1. In this cycle, registration takes place prior to the implementation of the research project (or, prior to the realization of outcomes). The fact of registration is shared with journals at the submission stage. If manuscripts reach the stage of provisional acceptance Journals taking part in the pilot can seek a report of registry compliance from the registry. At this stage researchers should submit to the registry their report of findings and compliance. Journals then make final publication decisions and signal the degree of registry compliance with the publication of the article.

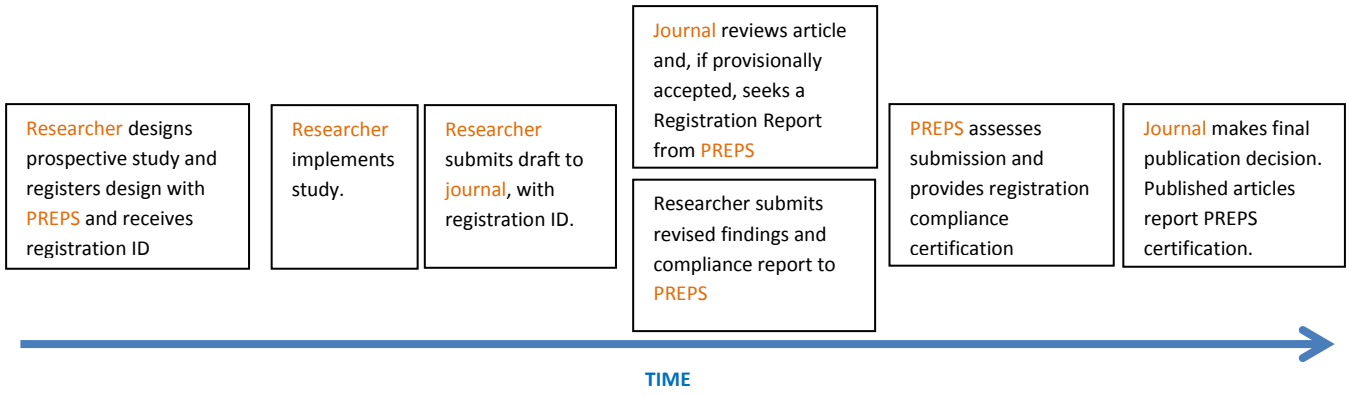

Fig 1: The Registration-Publication Cycle

3.3 Implications for authors

Under this design, authors would need to take several steps to register their designs and follow up on registration throughout the presentation and publication process.

1. **Registration Timing**: Registration should take place *before* outcomes are realized. Because circumstances could change during the course of the study, investigators can update their registration. A versioning system will allow tracking any changes that occur.
2. **Registration Scope**: At the time of initial registration, the minimal requirement is a full specification of the primary analysis. Best practice would be specification of primary and secondary analyses in the form of a mock report (see discussion in Humphreys et al 2012).
3. **Writing About Registration**: At the time final papers or reports are written, investigators should reference the registered design, indicate which analyses were registered and which not, describe all deviations from the registered design, and if possible provide results without deviations somewhere in the report or a supporting appendix. Authors should reference the registered design in the opening acknowledgments, in each table/figure caption explicitly noting whether the conclusion in question was registered, and throughout the text as appropriate. Further they should provide access to the registration as part of a supporting information appendix and by link to the experimental registry.

4. **Submission to Journal**: At time of submission to a journal, investigators should highlight in a cover note that the design was registered and provide relevant information (design title, registration id, etc.) for editors and referees to access the registration.
5. **Prepublication**: If submissions are provisionally accepted, authors should provide a short report of findings and deviations to the registry to facilitate compilation of the registry compliance certification.
6. **Reporting without publication:** In the event that research does not get completed or published researchers should communicate the results of the research to the registry within 3 years.

3.4 Implications for journal editors

Despite considerable attention and numerous false starts, due to industry resistance, funding, lack of awareness and enforcement, it took 30-40 years before registries caught hold in medicine (Dickersin and Rennie 2003). The success of registries in recent years is largely due to the efforts of journal editors, who endorsed their support for registration in print. More importantly, leading journals began requiring registration as a prerequisite for publication (De Angelis et al. 2002). Making registration a requirement in political science is not feasible at this stage. But public support from journals would help advance the agenda. We expect that the costs of such support would not be onerous.

1. **Submission to Journal**: At time of submission, editors indicate to reviewers that the design was registered and provide information about how to access the design. Ideally, referees would be directed to the registry where registry staff can provide information to reviewers. Authors may decide to include the registration as an appendix or journal editors may decide to include the registration documentation as part of the standard review process.
2. **Decision on Manuscript**: At decision time, editors request from the registry a report of ***registry compliance***. Editors would need to provide a copy of the manuscript and would need to allow some time for registry staff to assess compliance. Compliance would be non-binding in the sense that a published paper could be fully non-compliant. Thus, the registry in no way ties the hands of editors in the decisions they make. Ideally, editors would view compliance as an important criterion in the decision about the manuscript. Moreover, it may be helpful if editors take appropriate account of reasonable requests (from authors) to ignore referee comments that ask authors to deviate substantially from the registered design (or to report such requests as supporting information rather than in the main text).
3. **Report compliance level.** At time of publication: editors indicate the type of registry compliance on first page. Suggested scheme:
    a. Registration Compliant ★★★
    b. Registration Compliant with minor deviations ★★☆
    c. Registration Compliant with major deviations ★☆☆
    d. Not Registration Compliant ☆☆☆
    e. Other

    The ideas underlying “compliance” and “deviation” designations will need to be fleshed out so that it is clear what they do and do not mean. The core idea however is that research is compliant if designs are registered and write-ups reflect when research is consistent with or departs from registered designs. Thus “compliant” does not mean that there were no deviations from original plans. Deviations may occur for justifiable reasons and do not imply a failure on the part of investigators to comply with principles of registration; when this arises work would be classified as compliant with

minor or major deviations. As a legal matter the reported level of compliance as published by the journal could reference the determination of the registry and not be seen as a determination of the journal in question.

4. **Publish Journal Position on Registration.** In addition to these logistical and decision-making matters, the registration process would be enhanced considerably if editors publish a statement encouraging registration (on web and, if possible, in print) and make registration details prominent in published work. An editor's note could be included on the cover page of the article, for example. Alternatively, editors could also publish a list of registered studies separately, similar to the published list of manuscript referees.

3.5 Implications for registry staff

The structure has implications for registry staff at three points:

1. **Initial Review**: On submission of registered design, staff reviews registered protocol and indicates whether to accept or not. If accepted, registry staff may still offer suggestions for improving the registered design. Failing to accept does not, of course, imply that designs cannot be updated. If rejected, authors will be notified why and should update the design for resubmission with the basic registry standards. Importantly, the initial review will not assess the quality of ideas or of methods, but rather emphasize comprehensiveness and clarity by which compliance can be assessed later.
2. **Certification:** On request of participating journals, staff provides a report of registration and registry compliance. The basic scheme would include an assessment of the degree of deviation such as that shown in Section 3.4.[3] We expect that this assessment will check that authors reporting of results is consistent with registered designs and is indicated clearly when it is not. *We do not expect that this will involve actual replication of results by registry staff.*
3. **Follow up:** Registry staff should follow up with authors of all registered studies within a fixed period (3 years) to gather basic data on results of research, whether published or not.

3.6 Implications for others

Throughout the pilot period, the registry should be publicized in multiple ways. We anticipate that the collective efforts of EGAP, other political methodology groups, and journals will be needed to attain sufficient publicity within the five year time period. Other options may also be useful including providing information to those teaching surveys/experiments courses in PhD programs.

4 Proposed List of Core Fields for Registry

The registry fields will be designed to gather some basic required information and then offer maximum flexibility. Along with required fields, a set of optional fields would follow asking for more detail. Further, space would be available for detailed analysis plans or mock reports to follow the basic registry form. Other registries report that on average people spend about 30-45 minutes filling out the complete registration.

In the registry form, standard fields and responses will be supplied (perhaps as drop-down menus) to facilitate registry searches as well as comparability with other registries. The registry form for *initial designs* would be structured roughly as follows:

[3] [Possibly?] If authors intend for material to be published or disclosed in some other way, authors could themselves request a registry compliance report. Registry staff would supply the report to the relevant parties as needed.

Box 1: Proposed design registration fields [Non Mandatory Items in Grey]

I. Investigator Information and Contact
   a. Name; Position; Degree; Contact Info; Co-investigators Names *[Required]*
   b. Co-investigators contact info *[Optional]*

II. Title and Background Information
   a. Study Title *[Required]*
   b. Brief Summary
   c. Detailed Summary *[Optional]*
   d. Acronyms; Keywords *[Optional]*

III. Type of study:
   a. Expected date of completion of analysis *[Required]*
   b. Geographical coverage [Required]
   c. Prospective or not *[Automatic]*

IV. Human Subjects Review
   a. Approval status; approval number; board name/location / None *[Required]*
   b. Approval information for collaborators' IRBs/Ethics Boards *[Optional]*

V. Sponsors/Collaborators
   a. Sponsor, institutional collaborator, or funder (Funder ID: NSF ID / Other ID / None)
   b. Who will implement the project? (Researcher/Funder) *[Required]*
   c. Restrictions on presentation and/or publication
   d. Remuneration to researcher (Optional)

VI. Study Design.

VII. **1. Indicate Number of Conclusions**

2. For each conclusion:
   a. **Conclusion Type**: Hypothesis Test / Estimation of Quantity / Other *[Required]*
   b. **Statement of conclusion** (eg Hypotheses, Assessment of estimated quantity)
   c. **Dependent Variable**: Conceptual and operational; None if none; *[If applicable]*
   d. **Independent Variable**: Conceptual and operational; None if none. *[If applicable]*
   e. **Subjects** (Description)
   f. **Data Source**
   g. **Unit of analysis (**Description; Recruitment/Sampling Procedures**)**
   h. **Number of Units**
   i. **Location**
   j. **Estimation strategy / test strategy**: Dropdown list of test types / analysis methods / other
   k. **Subgroup:** List any subgroups analysis *[If applicable]*
   l. Was a power analysis conducted? If so, on which population *[Optional]*
   m. **Other secondary analyses *[If applicable]***
   n. **Controls:** List any controls *[If applicable]*
   o. **Experimental or not?**
   p. **Randomization strategy**; *[For experimental work; Optional]*
   q. **Code** for analysis *[Optional]*
   r. Illustration of Conclusions *[Optional]*

3. Dissemination expectations
   a. Indicate which conclusions form part of each expected publication.
   b. Does author expect research to be peer-reviewed [Optional]

4. Status
   a. Not yet recruiting; recruiting; ongoing; completed; suspended; terminated; *[Optional]*
   b. Initial registration or updated registration; Date of modification *[automatic]*

5. Miscellaneous Information
   a. System produces for user a unique ID and a citation to the registration design (ID included in citation)
   b. Additional information such as websites or research materials *[Optional]*
   c. Details on sampling. *[Optional]*
   d. Detailed analysis plan or Mock Report. This must be done in a way that makes the compliance check manageable for registry staff to check compliance *[Optional]*

The registry form for the *results and compliance report* might be structured roughly as follows. Alternatively, researchers may simply elect to upload a copy of the published report that details how all aspects of the registration were addressed.

Box 2: Results and compliance report fields [Optional Items in Grey]

I. Investigator and Registration Information
- a. ID of registered design (would self-populate many other fields) *[Required]*
- b. Name, Title, Etc *[Required; should be self-populated but can be updated]*

II. Sponsors / Collaborators [Options to Update the Sponsor Fields]

III. Status
- a. Completed and published / Completed and in process of publishing /Completed but abandoned (and why? Tried to publish, zero impact), Not completed *[Required]*
- b. Reasons not completed:
  - i. Funding / Outcomes not realized / Abandonment for Professional Reasons / Abandonment for Personal Reasons / *[Required]*
  - ii. Which data and portions were completed before abandoning [Optional]

IV. Results of Trial:
- a. Report results here for all interventions and outcomes *[Required]*

V. Study flow & Deviations
- a. Significant events affecting the design: research recruitment; assignment; interventions; time periods;
  - i. Were any changes significant to be considered deviations from the registration? *[Required]*
  - ii. Describe deviations; Hypotheses affected; Outcome measures affected *[Required]*

VI. Replication data
- a. Will data be posted; where posted; status of posting *[Optional]*
- b. Post report / paper of the findings *[Optional]*

5 Proposed Staffing and Governance of Registry

To begin, the registry would need one full time manager, one half time IT specialist, one half time graduate student, and oversight from a governing board. As the registry develops, registry staff should become more professionalized. The manager should provide biannual reports to the governing board with recommendations for modifications to the design, which should be adopted or amended by the governing board by a simple majority.

6 Other possible functions of Registry

In addition to the core functions described above, registry staff could provide other services, especially in early stages. These might include:

1. Populate and maintain a database of **disciplinary *priors*** regarding outcomes of prospective research.
2. Construct a **JEL-style** coding scheme for political science research, so that the prospective designs and associated results could be categorized and compared with the larger body of research.
3. **Survey** scholars to understand why some use the registry and others do not. Related, a survey of some sort could provide the basis for an outreach effort to educate scholars about the registry.

4. **Research.** Engage in research to better assess patterns of bias. This could be done for example by collecting **data** on unpublished results in targeted areas, collecting data on frequency with which funded research contradicts funder's mission, and so on.
5. **Tracking Access/Impact**: The registry could keep track of numbers of downloads as well as where the designs are downloaded as a way of capturing how much designs are followed.

7 What we hope to learn from the pilot

The ultimate purpose of the pilot is to determine the right way to design a full-scale registry.

1. Likely volume of projects to be registered and caseload for PREPS for certification.
2. What level of detail regarding designs should be required and what should be optional.
3. Whether in practice researchers depart substantially from designs and whether a binding scheme would be operable or onerous.
4. What is an effective way to indicate which parts of a research report are or are not registry-compliant.
5. Whether and how journal review processes need to be modified in light of the incentives for final research to be consistent with registered designs.

8 Other considerations

1. **Where to house the registry.** Agreement would need to be reached between key stakeholders on where the registry would be housed. Possibilities might include (not necessarily in this order):
    1. APSA
    2. APSA experiments section
    3. POLMETH
    4. EGAP
    5. An independent institution

A key consideration is to ensure that the registry is seen as being open across subject and approaches. This might rule out EGAP as an appropriate host given our focus on experimental work, and to a large extent, on international development.

2. **Including retrospective studies**: While we propose a first focus on studies with a prospective design, in principle such an approach could apply to retrospective studies also although in this case there are risks that hypotheses are generated in light of prior knowledge of patterns in the data. Appropriate protocols may however be developed for such cases.

3. **Capped Pilot**: If demands on the time and resources of registry staff are too great during the pilot the registry may randomly sample which studies to admit to the registry and support.

4. **Engaging research funders in aligning incentives:** Research funders could be asked to recognize the registry and specifically to (a) publicize support (b) request documentation of registration alongside first progress report.

5. **Fora**: We might consider a moderated wiki area for each design, which would provide a way for other scholars to comment on existing designs or results. A private comment field could also be added.

9 References

Banerjee, Abhijit. 2009. "The Experimental Approach to Development Economics." *Annual Review of Economics* 1: 151-178.

Bose, Ron. 2010. "CONSORT Extensions for Development Effectiveness: Guidelines for the Reporting of Randomised Control Trials of Social and Economic Policy Interventions in Developing Countries." *Journal of Development Effectiveness* 2(1): 173-186.

Cohen, Jessica, and William Easterly (Eds.). 2009. *What Works in Development? Thinking Big and Thinking Small.* Washington, D.C.: Brookings Prss.

De Angelis, Catherine, Jeffrey M. Drazen, Frank A. Frizelle, Charlotte Haug, John Hoey, Richard Horton, Sheldon Kotzin, Christine Laine, Ana Marusic, A . John P.M. Overbeke, Torben V. Schroeder, Hal C. Sox, Martin B. Van Der Weyden. 2004. "Clinical Trial Registration: A Statement from the International Committee of Medical Journal Editors." *The New England Journal of Medicine* 351(12): 1250-1251.

Dickersin, Kay. 1990. "The Existence of Publication Bias and Risk Factors for its Occurrence." *Journal of the American Medical Association* 263(10): 1385-1389.

Dickersin, Kay, and Drummond Rennie. 2003. "Registering Clinical Trials." *Journal of the American Medical Association* 290(4): 516-523.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using Randomization in Development Economics Research: A Toolkit." *Handbook of Development Economics* 4: 3895-3962.

Findley, Michael G., Daniel L. Nielson, and Jason C. Sharman. 2012. "Causes of Non-Compliance with International Law: Evidence from a Field Experiment on Financial Transparency." Unpublished Manuscript.

Gerber, Alan S., and Neil Malhotra. 2008. "Do Statistical Reporting Standards Affect What is Published? Publication Bias in Two Leading Political Science Journals." *Quarterly Journal of Political Science 3: 313-326.*

Gerber, Alan S., Neil Malhotra, Conor M. Dowling, and David Doherty. 2010. "Publication Bias in Two Political Behavior Literatures." *American Politics Research* 38(4): 591-613.

Glewwe, P., and Michael Kremer. 2006. "Schools, Teachers, and Education Outcomes in Developing Countries." *Handbook of Economics Education* 2: 945-1017.

Humphreys, Macartan, Raul Sanchez de la Sierra, and Peter van der Windt. 2012. "Fishing." *Political Analysis* (Forthcoming.)

Ioannidis, J. 1998. "Effect of the Statistical Significance of Results on the Time to Completion and Publication of Randomized Efficacy Trials." *Journal of the American Medical Association* 279(4): 281-286.

Rasmussen, Ole Dahl, Nikolaj Malchow-Moller, and Thomas Barnebeck Andersen. 2011. "Walking the Talk: The Need for a Trial Registry for Development Interventions." Unpublished Manuscript: University of Southern Denmark.

Simes, Robert J. 1986. "Publication Bias: The Case for An International Registry of Clinical Trials." *Journal of Clinical Oncology* 4: 1529-1541.

10 Annex

10.1 Illustration of schemes for Indicating Registration Status of Articles

1 Text

Article Title Article Title Article Title Article Title

Author Name 1, Affiliation
Author Name 2, Affiliation

Article Registration status: COMPLIANT

Abstract Abstract

2: Symbols – these may require a footnote explaining meaning of symbols

★★

Article Title Article Title Article Title Article Title

Author Name 1, Affiliation
Author Name 2, Affiliation

Abstract Abstract

3 Text that also indicates absence of registration

Article Title Article Title Article Title Article Title

Author Name 1, Affiliation
Author Name 2, Affiliation

Article Registration status: NONE

Abstract Abstract

10.2 Illustration of registered design for the test of a causal hypothesis

Registration Protocol Number:	1207ABCD
Investigator Information:	Reggi Star, Assoc Prof, SU. Email: r@star.edu
Title and Background Information	
Study Title:	Crime and Punishment
Brief Summary:	We seek to examine the level of political engagement in Rajasthan and whether information about criminality of MLAs leads to lower levels of political engagement.
Expected date of completion of analysis:	June 2014
Human Subjects Review Status:	Approved, Star IRB #1024
Sponsors/Collaborators	
Funder ID:	NSF 100024
Reporting Restrictions:	None
Study Design.	
Number of conclusions	2

Conclusion 1	
Conclusion Type	☐Test ☒Estimation
☒ Estimation Module (C1)	
Informal Description of Estimand:	We seek to estimate voter turnout overall as well as in rich and poor areas.
Subjects:	All voters in Rajasthan.
Unit of analysis:	Constituency.
Number of units:	200.
Location	Rajasthan, India.
Estimation Strategy:	Turnout will be measured based on official records provided by X. Rich and poor areas are defined based on whether government income score X for village lies above or below median. Estimate is calculated with 95% confidence intervals generated from sample variance.
Design:	☒ Observational ☐ Experimental
Sample estimation code:	☐ Not Provided ☒ Provided

☒Code Module (Optional):

```
# Simulate data
set.seed(1); C =  rep(c(F,T),100); X = (1:200)>100;  Y = .9*runif(200)+.1*(!X)+C
mci = function(Z) {ci<- t.test(Z)$conf.int; paste(round(mean(Z),2), "(", round(ci[1],2) , ",",
        round(ci[2],2),")", sep="")}
mci(Y); mci(Y[C]); mci(Y[!C])
```

Conclusion Illustration:	☐ Not Provided ☒ Provided

☒Illustration Module (Optional):

	All	Poor	Rich
Mean (ci)	M (l,u)	M (l,u)	M (l,u)

Conclusion 2	
Conclusion Type	☒Test ☐Estimation
☒Test Module (C2)	
Primary Hypothesis:	Constituencies in which information on criminality is circulated have lower turnout that constituencies do not.
Secondary Hypothesis:	The hypothesis holds in both poor and rich subgroups.
Dependent Variable:	Share of registered voters that vote in the 2014 MLA elections, as provided by Rajasthan electoral commission.
Independent Variables:	In treated areas a radio campaign reports lists criminal indictments as reported in candidate affidavits.
Subjects:	All voters in Rajasthan.
Unit of analysis:	Constituency.
Number of units:	200.
Location	Rajasthan, India.
Primary test description:	One sided test of difference in means between treated and control constituencies using significance level .05, with p

	values generated using randomization inference.
Subgroup analysis:	**Separate analysis for rich and poor areas.**
Controls:	**No controls**
Conclusion Type:	☐ Observational ☒ Experimental

☒Experiment Design Submodule:

Description of randomization procedure:

Simple random assignment of 100 units to treatment. Assignment code is in R.

☒ Assignment Code Submodule (Optional):

```
# Assignment Replication Code in R
set.seed(1); sample(200,100)
```

Test replication code: ☐ Not Provided ☒ Provided

☒Replication Code Submodule (Optional):

```
# Replication Code in R

# Simulate data
set.seed(1); C =  rep(c(F,T),100); X = (1:200)>100;  Y = .9*runif(200)+.1*(!X)+C

# Simulate calculation of ATEs:
c(mean(Y[X])- mean(Y[!X]), mean(Y[X&C])- mean(Y[!X&C]), mean(Y[X&!C])- mean(Y[!X&!C]))

# Simulate calculation of p values (using randomization inference)
p = function(X,Y) {mean(sapply(1:1000, function(i) {Z<-sample(X); mean(Y[Z])-
       mean(Y[!Z])}) <= (mean(Y[X])- mean(Y[!X])))}
c(p(X,Y), p(X[C],Y[C]), p(X[!C],Y[!C]))
```

Conclusion Illustration: ☐ Not Provided ☒ Provided

☒Illustration Module (Optional):

	All	Poor	Rich
Treated	Y (N)	Y (N)	Y (N)
Control	Y (N)	Y (N)	Y (N)
Difference	D (p)	D (p)	D (p)

Publication Plans	**We expect to publish conclusions 1 and 2 as part of the same analysis.**