Welcome to #BITSS2014!

Agenda:

Thursday 9A-12P Training

Thursday 1P-5P Conference

Friday 8:30A-2:30P Research Seminar

Registrations and Pre-Analysis Plans

Making research more transparent and reproducible

Garret Christensen[1]

[1]Berkeley Initiative for Transparency in the Social Sciences
University of California Berkeley
and
Center for Open Science

BITSS Annual Meeting, 2014

Outline

1. Problems
   - Publication Bias
   - P-Hacking
2. Solutions
   - Registration
   - Pre-Analysis Plan
3. Conclusion

Publication Bias

- There is a higher fraction of rejected hypothesis tests in the social sciences than in physical sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- Data on the full set of experiments run with a large survey shows strong results are 40pp more likely to be published, and 60pp more likely to be written up. The file drawer problem is massive. (Franco, Malhotra, Simonovits 2014—see this afternoon)

Publication Bias

- There is a higher fraction of rejected hypothesis tests in the social sciences than in physical sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- Data on the full set of experiments run with a large survey shows strong results are 40pp more likely to be published, and 60pp more likely to be written up. The file drawer problem is massive. (Franco, Malhotra, Simonovits 2014—see this afternoon)

Publication Bias

- There is a higher fraction of rejected hypothesis tests in the social sciences than in physical sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- Data on the full set of experiments run with a large survey shows strong results are 40pp more likely to be published, and 60pp more likely to be written up. The file drawer problem is massive. (Franco, Malhotra, Simonovits 2014—see this afternoon)

Publication Bias

- There is a higher fraction of rejected hypothesis tests in the social sciences than in physical sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- Data on the full set of experiments run with a large survey shows strong results are 40pp more likely to be published, and 60pp more likely to be written up. The file drawer problem is massive. (Franco, Malhotra, Simonovits 2014—see this afternoon)

Publication Bias

- There is a higher fraction of rejected hypothesis tests in the social sciences than in physical sciences (Fanelli 2010).
- Published null results are disappearing over time, in all disciplines (Fanelli 2011).
- This is very unlikely to represent the true state of the universe.
- Data on the full set of experiments run with a large survey shows strong results are 40pp more likely to be published, and 60pp more likely to be written up. The file drawer problem is massive. (Franco, Malhotra, Simonovits 2014—see this afternoon)

Publication Bias

If we only write up/publish significant results, and we have no record of all the insignificant results, we have no way to tell if our 'significant' results are real, or if they're the 5% we should expect due to randomness.

P-Hacking

- Also called fishing, researcher degrees of freedom, data mining, data massaging, or specification searching.
- Definition: flexibility in data analysis allows portrayal of *anything* as below an arbitrary p-value threshhold; significance loses its meaning.
- Not something only evil people do. It can be subconcious—humans are really good at motivated reasoning.

P-Hacking

- Also called fishing, researcher degrees of freedom, data mining, data massaging, or specification searching.
- Definition: flexibility in data analysis allows portrayal of *anything* as below an arbitrary p-value threshhold; significance loses its meaning.
- Not something only evil people do. It can be subconcious—humans are really good at motivated reasoning.

P-Hacking

- Also called fishing, researcher degrees of freedom, data mining, data massaging, or specification searching.
- Definition: flexibility in data analysis allows portrayal of *anything* as below an arbitrary p-value threshhold; significance loses its meaning.
- Not something only evil people do. It can be subconcious—humans are really good at motivated reasoning.

P-Hacking

- Also called fishing, researcher degrees of freedom, data mining, data massaging, or specification searching.
- Definition: flexibility in data analysis allows portrayal of *anything* as below an arbitrary p-value threshhold; significance loses its meaning.
- Not something only evil people do. It can be subconcious—humans are really good at motivated reasoning.

Does this actually happen? (JLP 2011)

1. In a paper, failing to report all of a study's dependent measures
2. Deciding whether to collect more data after looking to see whether the results were significant
3. In a paper, failing to report all of a study's conditions
4. Stopping collecting data earlier than planned because one found the result that one had been looking for
5. In a paper, "rounding off" a *p* value (e.g., reporting that a *p* value of .054 is less than .05)
6. In a paper, selectively reporting studies that "worked"
7. Deciding whether to exclude data after looking at the impact of doing so on the results
8. In a paper, reporting an unexpected finding as having been predicted from the start
9. In a paper, claiming that results are unaffected by demographic variables (e.g., gender) when one is actually unsure (or knows that they do)
10. Falsifying data

1. In a paper, failing to report all of a study's dependent measures	63.4
2. Deciding whether to collect more data after looking to see whether the results were significant	55.9
3. In a paper, failing to report all of a study's conditions	27.7
4. Stopping collecting data earlier than planned because one found the result that one had been looking for	15.6
5. In a paper, "rounding off" a *p* value (e.g., reporting that a *p* value of .054 is less than .05)	22.0
6. In a paper, selectively reporting studies that "worked"	45.8
7. Deciding whether to exclude data after looking at the impact of doing so on the results	38.2
8. In a paper, reporting an unexpected finding as having been predicted from the start	27.0
9. In a paper, claiming that results are unaffected by demographic variables (e.g., gender) when one is actually unsure (or knows that they do)	3.0
10. Falsifying data	0.6

	Admission rate	Defensibility rate
1. In a paper, failing to report all of a study's dependent measures	63.4	1.84 (0.39)
2. Deciding whether to collect more data after looking to see whether the results were significant	55.9	1.79 (0.44)
3. In a paper, failing to report all of a study's conditions	27.7	1.77 (0.49)
4. Stopping collecting data earlier than planned because one found the result that one had been looking for	15.6	1.76 (0.48)
5. In a paper, "rounding off" a *p* value (e.g., reporting that a *p* value of .054 is less than .05)	22.0	1.68 (0.57)
6. In a paper, selectively reporting studies that "worked"	45.8	1.66 (0.53)
7. Deciding whether to exclude data after looking at the impact of doing so on the results	38.2	1.61 (0.59)
8. In a paper, reporting an unexpected finding as having been predicted from the start	27.0	1.50 (0.60)
9. In a paper, claiming that results are unaffected by demographic variables (e.g., gender) when one is actually unsure (or knows that they do)	3.0	1.32 (0.60)
10. Falsifying data	0.6	0.16 (0.38)

P-Hacking

- Do people actually do this? (Previous—John, Loewenstein, Prelec 2011)
- Listening to the Beatles' "When I'm Sixty-Four" makes you younger. (Simmons, Nelson, Simonsohn 2011)
- Inordinately many .049 p-values, and indordinately few .051's. (Brodeur et al 2013)
- Political ideologues literally see in black and white (Nosek, Spies, Motyl 2012)

P-Hacking

- Do people actually do this? (Previous—John, Loewenstein, Prelec 2011)
- Listening to the Beatles' "When I'm Sixty-Four" makes you younger. (Simmons, Nelson, Simonsohn 2011)
- Inordinately many .049 p-values, and indordinately few .051's. (Brodeur et al 2013)
- Political ideologues literally see in black and white (Nosek, Spies, Motyl 2012)

P-Hacking

- Do people actually do this? (Previous—John, Loewenstein, Prelec 2011)
- Listening to the Beatles' "When I'm Sixty-Four" makes you younger. (Simmons, Nelson, Simonsohn 2011)
- Inordinately many .049 p-values, and indordinately few .051's. (Brodeur et al 2013)
- Political ideologues literally see in black and white (Nosek, Spies, Motyl 2012)

P-Hacking

- Do people actually do this? (Previous—John, Loewenstein, Prelec 2011)
- Listening to the Beatles' "When I'm Sixty-Four" makes you younger. (Simmons, Nelson, Simonsohn 2011)
- Inordinately many .049 p-values, and indordinately few .051's. (Brodeur et al 2013)
- Political ideologues literally see in black and white (Nosek, Spies, Motyl 2012)

Solutions

- Publication Bias—Registration
- P-Hacking—Pre-Analysis Plans

Solutions

- Publication Bias—Registration
- P-Hacking—Pre-Analysis Plans

Registrations

- Publicly stating all research you will do, and what hypotheses you will test, prospectively.
- Store this statement in a public registry.
- Near universal adoption in medical RCTs. Top journals won't publish if it's not registered.
  `http://clinicaltrials.gov`
- Even better if registry requires registering outcomes after study. Currently limited, but NIH is moving on this.

Registrations

- Publicly stating all research you will do, and what hypotheses you will test, prospectively.
- Store this statement in a public registry.
- Near universal adoption in medical RCTs. Top journals won't publish if it's not registered.
  `http://clinicaltrials.gov`
- Even better if registry requires registering outcomes after study. Currently limited, but NIH is moving on this.

Registrations

- Publicly stating all research you will do, and what hypotheses you will test, prospectively.
- Store this statement in a public registry.
- Near universal adoption in medical RCTs. Top journals won't publish if it's not registered.
  `http://clinicaltrials.gov`
- Even better if registry requires registering outcomes after study. Currently limited, but NIH is moving on this.

Registrations

- Publicly stating all research you will do, and what hypotheses you will test, prospectively.
- Store this statement in a public registry.
- Near universal adoption in medical RCTs. Top journals won't publish if it's not registered.
  `http://clinicaltrials.gov`
- Even better if registry requires registering outcomes after study. Currently limited, but NIH is moving on this.

Registrations

- Publicly stating all research you will do, and what hypotheses you will test, prospectively.
- Store this statement in a public registry.
- Near universal adoption in medical RCTs. Top journals won't publish if it's not registered.
  `http://clinicaltrials.gov`
- Even better if registry requires registering outcomes after study. Currently limited, but NIH is moving on this.

Registrations

- Newer to social sciences, but:
  - AEA registry, currently only for RCTs.
    `http://socialscienceregistry.org`
  - EGAP registry
    `http://egap.org/design-registration`
  - 3ie registry, for developing country evaluations.
    `http://ridie.3ieimpact.org`
  - Open Science Framework `http://osf.io`
    - Open format
    - Will soon sync with above

Registrations

- Newer to social sciences, but:
  - AEA registry, currently only for RCTs.
    `http://socialscienceregistry.org`
  - EGAP registry
    `http://egap.org/design-registration`
  - 3ie registry, for developing country evaluations.
    `http://ridie.3ieimpact.org`
  - Open Science Framework `http://osf.io`
    - Open format
    - Will soon sync with above

Registrations

- Newer to social sciences, but:
  - AEA registry, currently only for RCTs.
    `http://socialscienceregistry.org`
  - EGAP registry
    `http://egap.org/design-registration`
  - 3ie registry, for developing country evaluations.
    `http://ridie.3ieimpact.org`
  - Open Science Framework `http://osf.io`
    - Open format
    - Will soon sync with above

Registrations

- Newer to social sciences, but:
  - AEA registry, currently only for RCTs.
    `http://socialscienceregistry.org`
  - EGAP registry
    `http://egap.org/design-registration`
  - 3ie registry, for developing country evaluations.
    `http://ridie.3ieimpact.org`
  - Open Science Framework `http://osf.io`
    - Open format
    - Will soon sync with above

Registrations

- Newer to social sciences, but:
  - AEA registry, currently only for RCTs.
    `http://socialscienceregistry.org`
  - EGAP registry
    `http://egap.org/design-registration`
  - 3ie registry, for developing country evaluations.
    `http://ridie.3ieimpact.org`
  - Open Science Framework `http://osf.io`
    - Open format
    - Will soon sync with above

Registrations

- Newer to social sciences, but:
  - AEA registry, currently only for RCTs. `http://socialscienceregistry.org`
  - EGAP registry `http://egap.org/design-registration`
  - 3ie registry, for developing country evaluations. `http://ridie.3ieimpact.org`
  - Open Science Framework `http://osf.io`
    - Open format
    - Will soon sync with above

Registrations

- Newer to social sciences, but:
  - AEA registry, currently only for RCTs. `http://socialscienceregistry.org`
  - EGAP registry `http://egap.org/design-registration`
  - 3ie registry, for developing country evaluations. `http://ridie.3ieimpact.org`
  - Open Science Framework `http://osf.io`
    - Open format
    - Will soon sync with above

Pre-Analysis Plan

- Often part of a registration
- From 3ie: "A pre-analysis plan is a detailed description of the analysis to be conducted that is written in advance of seeing the data on impacts of the program being evaluated. It may specify hypotheses to be tested, variable construction, equations to be estimated, controls to be used, and other aspects of the analysis. A key function of the pre-analysis plan is to increase transparency in the research. By setting out the details in advance of what will be done and before knowing the results, the plan guards against data mining and specification searching. Researchers are encouraged to develop and upload such a plan with their study registration, but it is not required for registration."

Pre-Analysis Plan

- Often part of a registration
- From 3ie: “A pre-analysis plan is a detailed description of the analysis to be conducted that is written in advance of seeing the data on impacts of the program being evaluated. It may specify hypotheses to be tested, variable construction, equations to be estimated, controls to be used, and other aspects of the analysis. A key function of the pre-analysis plan is to increase transparency in the research. By setting out the details in advance of what will be done and before knowing the results, the plan guards against data mining and specification searching. Researchers are encouraged to develop and upload such a plan with their study registration, but it is not required for registration.”

Pre-Analysis Plan

- Often part of a registration
- From 3ie: “A pre-analysis plan is a detailed description of the analysis to be conducted that is written in advance of seeing the data on impacts of the program being evaluated. It may specify hypotheses to be tested, variable construction, equations to be estimated, controls to be used, and other aspects of the analysis. A key function of the pre-analysis plan is to increase transparency in the research. By setting out the details in advance of what will be done and before knowing the results, the plan guards against data mining and specification searching. Researchers are encouraged to develop and upload such a plan with their study registration, but it is not required for registration.”

Glennerster, Takavarasha Suggestions

Running Randomized Evaluations

1. the main outcome measures,
2. which outcome measures are primary and which are secondary,
3. the precise composition of any families that will be used for mean effects analysis,
4. the subgroups that will be analyzed,
5. the direction of expected impact if we want to use a one-sided test, and
6. the primary specification to be used for the analysis.

Glennerster, Takavarasha Suggestions

Running Randomized Evaluations

1. the main outcome measures,
2. which outcome measures are primary and which are secondary,
3. the precise composition of any families that will be used for mean effects analysis,
4. the subgroups that will be analyzed,
5. the direction of expected impact if we want to use a one-sided test, and
6. the primary specification to be used for the analysis.

Glennerster, Takavarasha Suggestions

Running Randomized Evaluations

1. the main outcome measures,
2. which outcome measures are primary and which are secondary,
3. the precise composition of any families that will be used for mean effects analysis,
4. the subgroups that will be analyzed,
5. the direction of expected impact if we want to use a one-sided test, and
6. the primary specification to be used for the analysis.

Glennerster, Takavarasha Suggestions

Running Randomized Evaluations

1. the main outcome measures,
2. which outcome measures are primary and which are secondary,
3. the precise composition of any families that will be used for mean effects analysis,
4. the subgroups that will be analyzed,
5. the direction of expected impact if we want to use a one-sided test, and
6. the primary specification to be used for the analysis.

Glennerster, Takavarasha Suggestions

Running Randomized Evaluations

1. the main outcome measures,
2. which outcome measures are primary and which are secondary,
3. the precise composition of any families that will be used for mean effects analysis,
4. the subgroups that will be analyzed,
5. the direction of expected impact if we want to use a one-sided test, and
6. the primary specification to be used for the analysis.

Glennerster, Takavarasha Suggestions

Running Randomized Evaluations

1. the main outcome measures,
2. which outcome measures are primary and which are secondary,
3. the precise composition of any families that will be used for mean effects analysis,
4. the subgroups that will be analyzed,
5. the direction of expected impact if we want to use a one-sided test, and
6. the primary specification to be used for the analysis.

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

McKenzie Suggestions

World Bank Development Impact Blog

1. Description of the sample to be used in the study
2. Key data sources
3. Hypotheses to be tested throughout the causal chain
4. Specify how variables will be constructed
5. Specify the treatment effect equation to be estimated
6. What is the plan for how to deal with multiple outcomes and multiple hypothesis testing?
7. Procedures to be used for addressing survey attrition
8. How will the study deal with outcomes with limited variation?
9. If you are going to be testing a model, include the model
10. Remember to archive it

Simmons, Nelson, Simonsohn (2011)

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Simmons, Nelson, Simonsohn (2011)

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Simmons, Nelson, Simonsohn (2011)

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Simmons, Nelson, Simonsohn (2011)

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Simmons, Nelson, Simonsohn (2011)

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Simmons, Nelson, Simonsohn (2011)

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Simmons, Nelson, Simonsohn (2011)

1. Authors must decide the rule for terminating data collection before data collection begins and report this rule in the article.
2. Authors must collect at least 20 observations per cell or else provide a compelling cost-of-data-collection justification.
3. Authors must list all variables collected in a study.
4. Authors must report all experimental conditions, including failed manipulations.
5. If observations are eliminated, authors must also report what the statistical results are if those observations are included.
6. If an analysis includes a covariate, authors must report the statistical results of the analysis without the covariate.

Examples

Wide range of when to write and how detailed to make the plan. At the extreme level of detail you would have your entire code already written before you got any data.

- J-PAL Hypothesis Registry (11), see `http://www.povertyactionlab.org/Hypothesis-Registry`
- AEA Registry has relatively few, plentiful in EGAP.
- Casey, Glennerster, Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan" *QJE* 2012. (Paper,Plan)
  - Government-sponsored program.
  - Broad program (Community Driven Development)
  - Broad outcomes (trust, public goods, public services, community groups, information, participation, crime, welfare, attitudes)

Examples

Wide range of when to write and how detailed to make the plan. At the extreme level of detail you would have your entire code already written before you got any data.

- J-PAL Hypothesis Registry (11), see `http://www.povertyactionlab.org/Hypothesis-Registry`
- AEA Registry has relatively few, plentiful in EGAP.
- Casey, Glennerster, Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan" *QJE* 2012. (Paper,Plan)
  - Government-sponsored program.
  - Broad program (Community Driven Development)
  - Broad outcomes (trust, public goods, public services, community groups, information, participation, crime, welfare, attitudes)

Examples

Wide range of when to write and how detailed to make the plan. At the extreme level of detail you would have your entire code already written before you got any data.

- J-PAL Hypothesis Registry (11), see `http://www.povertyactionlab.org/Hypothesis-Registry`
- AEA Registry has relatively few, plentiful in EGAP.
- Casey, Glennerster, Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan" *QJE* 2012. (Paper,Plan)
  - Government-sponsored program.
  - Broad program (Community Driven Development)
  - Broad outcomes (trust, public goods, public services, community groups, information, participation, crime, welfare, attitudes)

Examples

Wide range of when to write and how detailed to make the plan. At the extreme level of detail you would have your entire code already written before you got any data.

- J-PAL Hypothesis Registry (11), see `http://www.povertyactionlab.org/Hypothesis-Registry`
- AEA Registry has relatively few, plentiful in EGAP.
- Casey, Glennerster, Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan" *QJE* 2012. (Paper,Plan)
  - Government-sponsored program.
  - Broad program (Community Driven Development)
  - Broad outcomes (trust, public goods, public services, community groups, information, participation, crime, welfare, attitudes)

Examples

Wide range of when to write and how detailed to make the plan. At the extreme level of detail you would have your entire code already written before you got any data.

- J-PAL Hypothesis Registry (11), see `http://www.povertyactionlab.org/Hypothesis-Registry`
- AEA Registry has relatively few, plentiful in EGAP.
- Casey, Glennerster, Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan" *QJE* 2012. (Paper,Plan)
  - Government-sponsored program.
  - Broad program (Community Driven Development)
  - Broad outcomes (trust, public goods, public services, community groups, information, participation, crime, welfare, attitudes)

Examples

Wide range of when to write and how detailed to make the plan. At the extreme level of detail you would have your entire code already written before you got any data.

- J-PAL Hypothesis Registry (11), see `http://www.povertyactionlab.org/Hypothesis-Registry`
- AEA Registry has relatively few, plentiful in EGAP.
- Casey, Glennerster, Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan" *QJE* 2012. (Paper,Plan)
  - Government-sponsored program.
  - Broad program (Community Driven Development)
  - Broad outcomes (trust, public goods, public services, community groups, information, participation, crime, welfare, attitudes)

Outcome variable	(1) Mean for controls	(2) Treatment effect
Panel A: GoBifo "weakened" institutions		
Attended meeting to decide what to do with the tarp	0.81	–0.04^{+}
Everybody had equal say in deciding how to use the tarp	0.51	–0.11^{+}
Community used the tarp (verified by physical assessment)	0.90	–0.08^{+}
Community can show research team the tarp	0.84	–0.12*
Respondent would like to be a member of the VDC	0.36	–0.04*
Respondent voted in the local government election (2008)	0.85	–0.04*
Panel B: GoBifo "strengthened" institutions		
Community teachers have been trained	0.47	0.12^{+}
Respondent is a member of a women's group	0.24	0.06**
Someone took minutes at the most recent community meeting	0.30	0.14*
Building materials stored in a public place when not in use	0.13	0.25*
Chiefdom official did not have the most influence over tarp use	0.54	0.06*
Respondent agrees with "Responsible young people can be good leaders" and not "Only older people are mature enough to be leaders"	0.76	0.04*
Correctly able to name the year of the next general elections	0.19	0.04*

Conclusion

- Neither registration nor pre-analysis plans are limited to randomized trials.
- Spectrum from confirmatory to exploratory research, all has value.
- I'd just like know which research is which.

Conclusion

- Neither registration nor pre-analysis plans are limited to randomized trials.
- Spectrum from confirmatory to exploratory research, all has value.
- I'd just like know which research is which.

Conclusion

- Neither registration nor pre-analysis plans are limited to randomized trials.
- Spectrum from confirmatory to exploratory research, all has value.
- I'd just like know which research is which.

Conclusion

- Neither registration nor pre-analysis plans are limited to randomized trials.
- Spectrum from confirmatory to exploratory research, all has value.
- I'd just like know which research is which.