Pre-registration for Economists: Exhaustive templates for primary and secondary data

Leibniz Open Science Day 2025 - Berlin

October 27th, 2025

Thibaut Arpinon
Chair Professor Junior (CPJ)

Savoie Mont-Blanc University (USMB)
Research Institute in Management in Economics (IREGE)

Key paper points

Starting point

- Evidence suggests questionable practices are widespread in economics.
  - → *Questionable Research Practices (QRPs)*
  - → *Questionable Publishing Practices (QPPs)*
- This hinders the credibility of evidence.

Identify current limitations in open science practices

- Registered Reports are most efficient, but implementation is slow…
- Pre-registration is most popular, but practices remain highly heterogeneous.
  - → *Absence of common and peer-reviewed guidelines.*
  - → *This causes increased ex-post degree of freedom and pre-registration as noisy quality signal.*

Paper's contribution

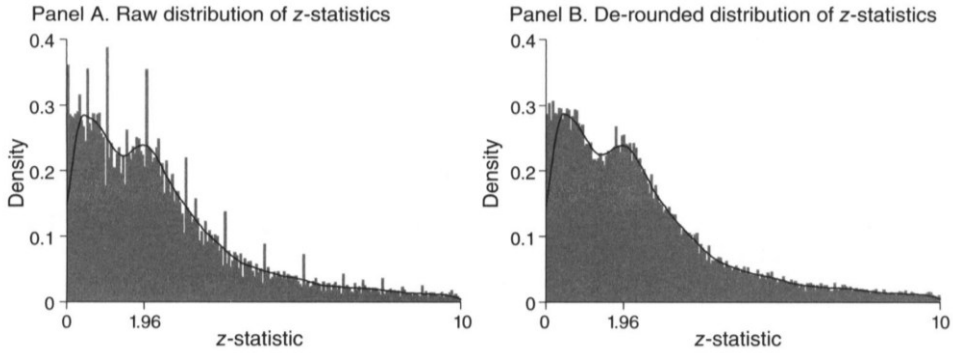
- Improve pre-registration with thorough primary and secondary data templates for economists.
  - → *Latest open science practices with resources and justification to allow full identification of analysis space.*
  - → *Deviation template to report ex-post deviations.*
- Decision tree to identify optimal research path depending on time, budget, and priors.
- Discussion of barriers to remove and how the templates can contribute.

A replication crisis in economics and management?

Some initial evidence suggests Questionable Research Practices (QRPs)

Biased distribution of tests statistics in economics - Selective reporting

- 50 000 tests published in AER, JPE and QJE between 2005 and 2011 [Brodeur et al. (2016)].



p-value	z-statistic
0.10	1.645
0.05	1.960
0.01	2.576
0.001	3.291

- Two humped camel-hump shape, with missing values between 0.25 et 0.10, which are retrieved right after the 0.05 threshold (in p-values)…

Francesca Gino's fraudulent data - Data falsification

- Accused of falsifying data by hand-modifying data on Excel [http://datacolada.org/109].

- Francesca Gino sued for defamation and sought for $25 million, but the case was dismissed

A replication crisis in economics and management?

Questionable Publishing Practices (QPPs), notably publication bias

- By nature, publication bias is difficult to measure. Why?

- It occurs at two levels:

  1. At the journal-level, at the stage of desk rejects and first-round rejections.

  → *To measure its effect, one would need to access this data, which is not easy…*

  2. At the researcher-level, when null result research projects are abandoned.

  « File Drawer Issue » [Rosenthal (1979)].

  → *To measure its effect, one would need to access all abandoned / unsubmitted projects…*

Evidence suggests the presence of QPPs in economics

- Economists' perceptions of null results with a vignette experiment [Chopra et al. (2023)].

→ Holding all other studies' features constant, null results are perceived as:

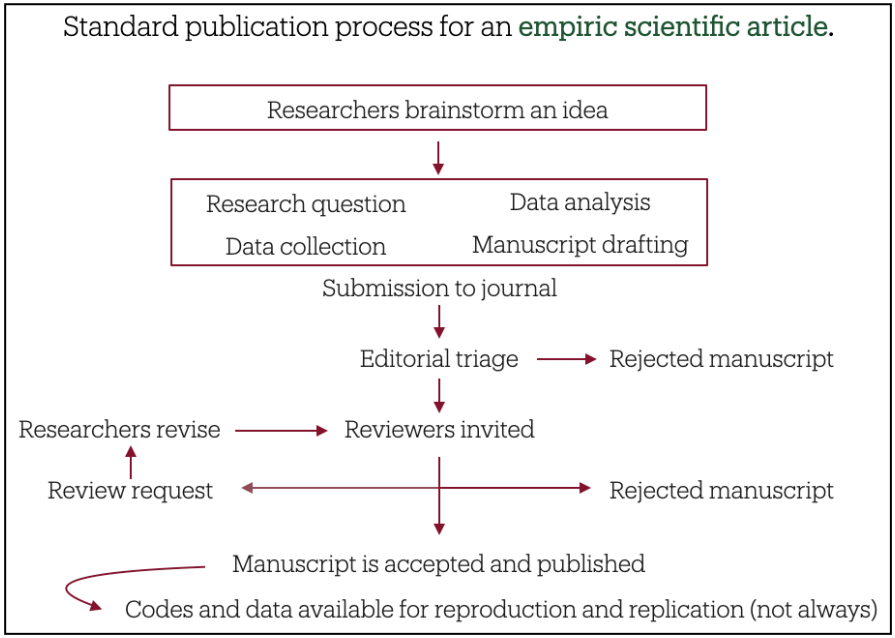
- *Less publishable*
- *Less important*
- *Of lower quality*
- *Less precisely estimated*
- *Penalty is larger when expressed with p-values (vs. standard errors)*

The current publication system

The current publication system is misaligned with scientific objectives.

- Focus on results rather than method.
- Journals have full power to accept or reject, based on the results.
- Admitting that publication bias is real, then incentives to find a positive result are high.

  → *Encourages selective reporting and p-hacking.*
- Researchers bear the cost, not journals.
- Admitting that publication bias is real, only positive results are available in the literature.



Undermines the credibility of research findings.

How can we solve this issue?

Registered Reports

The first best option to prevent QRPs and QPPS is the Registered Report format

- Registered Reports prevent both QRPs and QPPs through a two-stage publishing process.
- But, implementation is slow in economics compared to psychology or other disciplines.
- Only 13 economic journals accept Registered Reports…

List of economic journals accepting Registered Report submissions

Search...

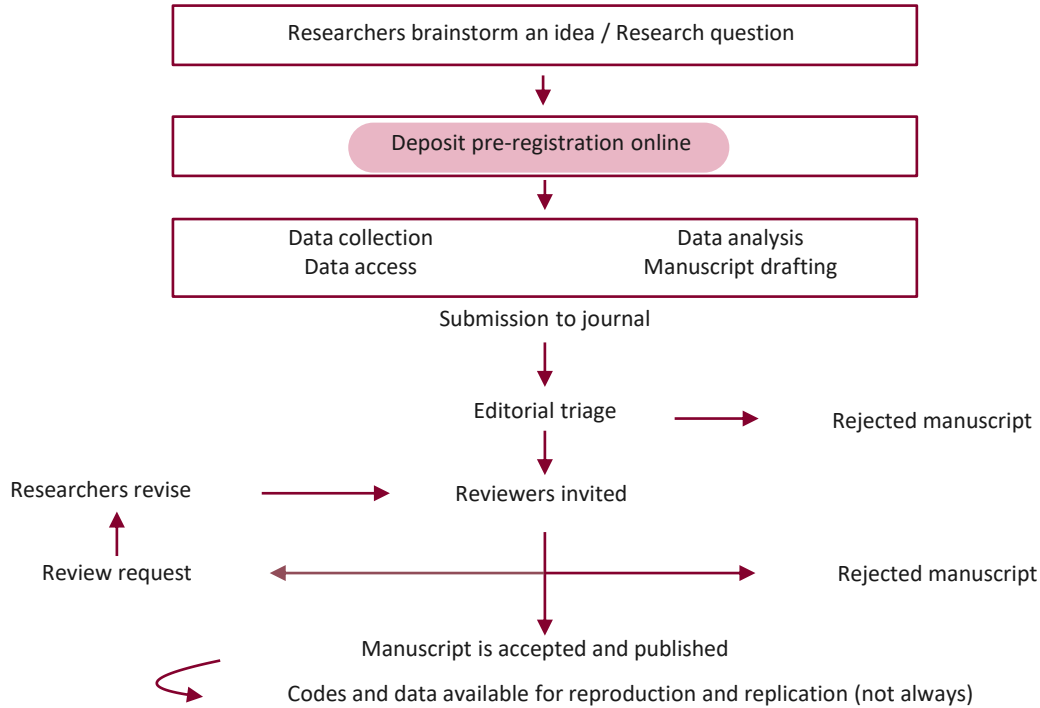
Journal	Position regarding Registered Report submission	Link
Journal of Development Economics	Accept submissions	http://jde-preresultsreview.org/
Q-Open	Accept submissions	https://academic.oup.com/qopen/pages/general_instructions
Journal of Behavioral and Experimental Economics	Accept submissions	https://www.journals.elsevier.com/journal-of-behavioral-and-experimental-economics/call-for-papers/transparency-reproducibility-and-generalizability-of-behavioral-economics-experiments

The complete list is available here: https://sites.google.com/view/thibautarpinon/registered-reports

Pre-registration

A more flexible option

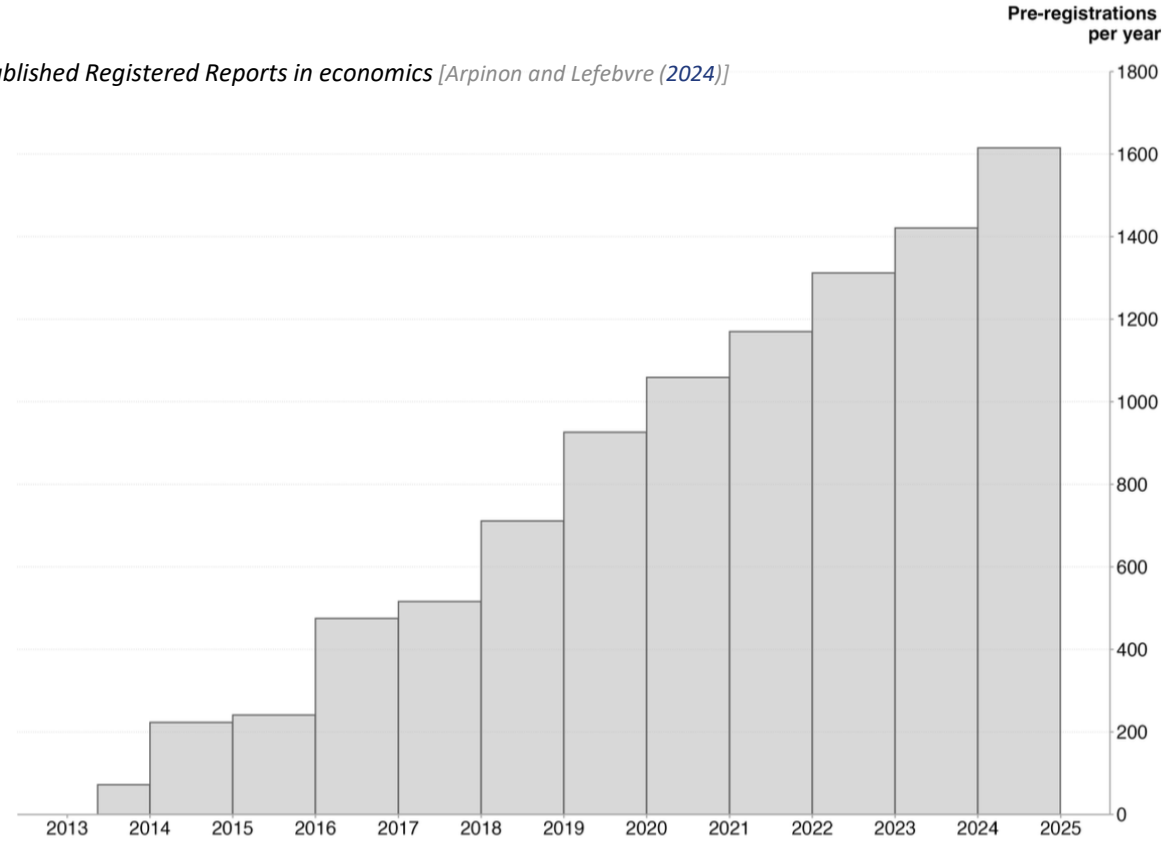
- Pre-registration is depositing a document online before collecting / accessing data.
  - → *What you will do and how you will do it, before doing it.*
  - → *Examples: Open Science Framework, AsPredicted, or AEA Registry.*
- Reviewers can access it at the revision stage to verify analysis conformity.



Pre-registration in economics

Pre-registration is popular in economics, much more so than Registered Reports

→ *Only 8 published Registered Reports in economics* [Arpinon and Lefebvre (2024)]

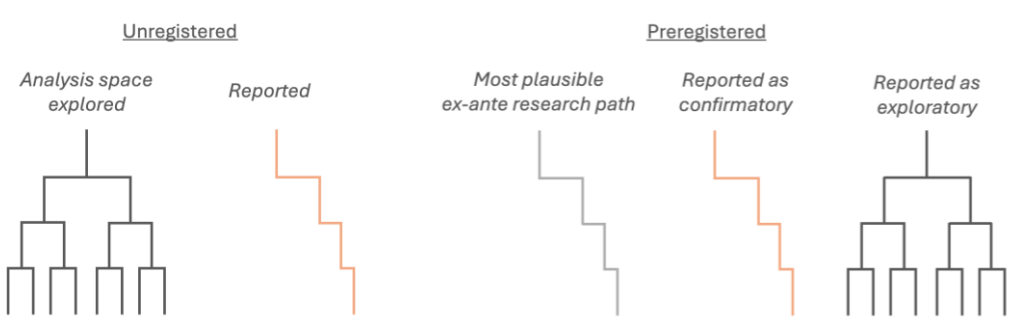


Notes: The data for this graph was taken from the `https://www.socialscienceregistry.org/`.

Pre-registration is limited…

Pre-registration can be efficient against QRPs, but not QPPs

- Pre-registration restricts researchers' degree of freedom by choosing most plausible ex-ante path.



But in its current form, pre-registration is (almost) useless…

- Heterogeneity in practices due to absence of common guidelines and practices.
  - → *Current pre-registration practices do not sufficiently reduce degrees of freedom.*
  - → *Available templates are incomplete and do not allow to efficiently identify analysis space.*
- As a result, pre-registration becomes an imperfect quality signal.
  - → *If reviewers never check the document, a bad pre-registration can be passed on as a quality signal.*
- Inconsistent standards allow for ex-post flexibility.
  - → *Voluntarily leaving out elements increases ex-post analysis space.*
- Researchers bear the weight, not journals.
  - → *A perfectly executed pre-registration ties the researchers' hands in case of null result.*
  - → *It is totally inefficient against QPPs.*

Pre-registration for Economists

Two exhaustive templates for primary and secondary data

Motivation: Since pre-registration is the most popular open science practice, let's maximise it's efficiency.

Methodology: Propose template / guideline with the latest open science practices relevant to economists.

Objectives:

- Allow optimal identification of the total analysis space to reduce ex-post degrees of freedom.
  - → Increasing severity of testing (i.e., test's ability to rigorously challenge a prediction).
- Clearly distinguish confirmatory from exploratory results.
  - → Disallow postdictions to be passed on as predictions.
- Standardize practices in economics, with a peer-reviewed resource, to decrease heterogeneity.
  - → The format allows for flexibility while maximizing transparency.

Who are the templates for:

- Researchers looking for a thorough resource to pre-register.
- Editors and reviewers looking to assess the quality of a pre-registration.

The overarching objective is to enhance research credibility in economics.

Pre-registration for Economists

The pre-registration templates

1. Administrative information
   - Type of project / Type of data
   - Authors' contributions CRediT
   - Research guarantor
   - Link to experiment / dataset
   - Ethics and/or IRB approval
   - Funding Source and amount
   - Project description
   - Project timeline
   - Other specific elements

Administrative information

Before collecting the data, researchers should anticipate administrative information surrounding a research project. This information is important as it allows editors, reviewers, and readers to quickly identify authors' pre-defined roles and data structure. Failure to report some of the elements can increase analysis space and should be justified.	
Do the researchers describe the type of research project/data collection? *Researchers should mention whether the research project is a field, lab, or online data collection.*	**Type of research project/data collection:** O Survey O Field experiment O Lab experiment O Online experiment O Other:
Does the data already exist?	O Yes *If yes, the researchers should refer to the secondary data pre-registration template.* O No
Do the researchers list the plan contributions for each co-author? *In a transparency effort, researchers can provide a list of the contributions made by each contributor. This list can be added in the pre-registration or in the final manuscript. However, listing the researchers' contributions at the pre-registration stage clarifies who drafted the pre-registration. For example, Nature journals recommend the CRediT designation available here: https://credit.niso.org/.*	O Yes If yes, please enter each author's contribution:

Pre-registration for Economists

The pre-registration templates

2. Sampling plan

- Data access/collection procedure (when, where, from whom, how)
- Termination rule / Incentivizing
- Data description
- Sample size rationale (lots of detail on the procedure)
- Design (primary and secondary outcomes)
- Randomization procedure and blinding
- Other specific elements

Sampling plan

Description of the data collection procedure *Researchers should have a clear idea of the data to be collected. Providing detailed elements on the data to be collected decreases the likelihood of forgetting elements and limits the ex-post analysis space. However, unpredicted events can alter the final data collection process. In this case, researchers have two options. If the alteration occurs before the pre-registration is submitted, researchers should append the document before the data is collected rather than attempt to rationalize the changes ex-post. If the alteration occurs once the pre-registration has been submitted, researchers should report the changes transparently in their final manuscript and discuss how these changes affect the severity of the testing. This should be reported in the "Deviations from pre-registration" template section.*	
Do the researchers indicate where the data will be collected? *Researchers should provide detail, whenever possible, on the location of the participants (e.g., countries, regions, villages). For projects collecting data in different locations, each location should be described, and the researchers should indicate how the data will be unified (e.g., data coding in the final database).*	O Yes *If yes, please describe where the data will be collected. Researchers should provide a concise description.* O No *If no, researchers do not describe where the data will be collected.* Justification:
Do the researchers indicate when the data will be collected? *Researchers should provide a timeframe for when the data collection will occur. This can be a fixed date or a timespan. If the data collection data changes due to unpredicted events, researchers should report it in their final manuscripts.* *Note of caution: The data collection should be dated after the pre-registration submission date. If the data collection is dated before, the researchers' degree of freedom and the risk of Pre-registering After the Results are Known (PARKing) greatly increase.*	O Yes *If yes, please describe when the data will be collected. Researchers should provide a concise description.* O No *If no, researchers do not describe when the data will be collected.* Justification:

Pre-registration for Economists

The pre-registration templates

3. Analysis plan

- Inclusion and exclusion rules
- Outliers and missing values correction
- Hypotheses and questions of interest
- Statistical testing

- Multiple Hypothesis Testing (MHT)
- Multiverse and specification curves
- Reliability and outcome neutral tests
- Exploratory analyses
- Other specific elements

Analysis plan

Inclusion and exclusion rules *At the pre-registration stage, researchers should anticipate potential observations in the data that will not be included in the final analysis and provide exclusion rules/inclusion rules. These rules need to be pre-specified to limit the analysis space. As mentioned earlier, researchers might run into unanticipated issues during the data collection process that force deviations from the pre-registered exclusion rules. In this case, the changes should be reported as deviations and a discussion on how this impacts the severity of testing should be provided. For more elements on deviations, please refer to the template "Deviations from the pre-registration".*		
Do the researchers provide exclusion rules? *Researchers should provide a set of pre-specified exclusion rules for observations that will not be included in the final analysis.*	O Yes *If yes is selected, see elements in right column.* O No *If no is selected, researchers do not provide exclusion rules.* Justification:	**If yes, do the researchers describe which observations will be excluded from the final analysis?** *Researchers should describe as accurately as possible the combination of variables and observations that will not be included in the final analysis. This can include, but is not limited to, the following:* • *Missing observations* • *Erroneous observations* • *Completion time above/below a predefined threshold* • *Inconsistent or overly consistent values* • *Observations with values above/below a predefined threshold* • *Failure of predefined attention or comprehension checks* • *Suspicious responses (e.g., bot generated)* • *Individual-specific characteristics (e.g., demographics)* • *Methodology based rules (e.g., inconsistent preferences)* • *Equipment errors* • *Any other data-specific occurrences* *Note of caution: Researchers should attempt to anticipate as many exclusion rules as possible. Any exclusion rule that is not pre-registered but added to the final analysis increases the analysis space and decreases the severity of testing. In this case, the deviations should be reported and a discussion of whether they affect the severity of testing should be provided.* O Yes If yes, please describe the exclusion rules:

Pre-registration for Economists

Sometimes, research projects don't go as planned and deviations from pre-registration must occur

- An additional template to report ex-post deviations.

What is the deviation?	*Examples: Added exclusion due to the presence of outliers; change in statistical model due to data structure; more data was collected than originally planned.*
Why did the deviation occur?	Examples: Mistake in the pre-registration; population has a high proportion of participants outside of a pre-registered threshold
How does the deviation impact the pre-registered path?	*Examples:* *Small deviation: A supplementary covariate variable was collected/accessed and added to the statistical modeling;* *Large deviation: The primary outcome data structure is different than anticipated, changing a pre-registered exclusion rule and statistical modeling.*
How does the deviation impact the severity of testing?	*Negative: An added unregistered exclusion rule increases the analysis space and decreases the severity of testing;* *Positive: An added multiple hypothesis correction procedure changes the alpha level and increases the severity of testing*

Objective: allow flexibility but with optimal analysis space identification.

But in the end, it is up to the editors and reviewers to decide whether the deviation increases or decreases the severity of testing.

Pre-registration for Economists

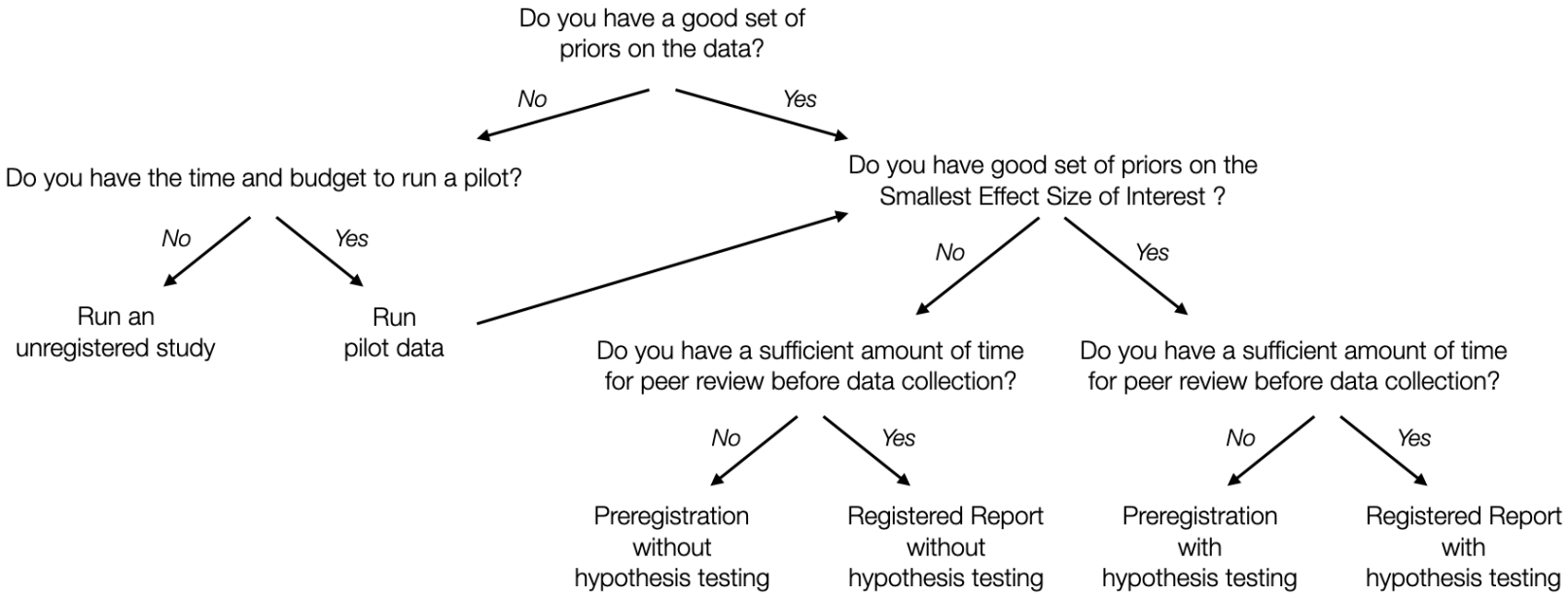
The templates surpass current existing alternatives

Feature	AEA RCT	AsPredicted	OSF	Preregistration for Economists
Primary and secondary data	✗	✗	✓	✓
Methodological references	✗	✗	✗	✓
Exhaustive guideline for:				
Administrative info.	*Incomplete*	✗	*Incomplete*	✓
Sampling plan	*Incomplete*	*Incomplete*	*Incomplete*	✓
Analysis plan	✗	*Incomplete*	*Incomplete*	✓
Justify omitted elements	✗	✗	✗	✓
Maximize analysis space identification	✗	✗	✗	✓
Document ex-post deviations	✗	✗	✗	✓

Notes: The mention *incomplete* refers to the absence of fundamental elements to restrict degrees of freedom efficiently and maximize transparency.

Should you always pre-register?

The decision of pre-registering depends on a combination of time, priors, and priors



Resources

How can you access the templates?

A simple pdf/word version to fill manually

Pre-registration for Economists: exhaustive templates for primary and secondary data

Thibaut Arpinon[1]

[1]Savoie Mont-Blanc University and IREGE*

Hopefully a version available soon on Open Science Framework (*submitted and currently under review*)

Additional resource for Registered Reports in economics

A practical guide to Registered Reports for economists

Published online by Cambridge University Press: **17 January 2025**

Thibaut Arpinon and Romain Espinosa

Show author details

Article | Supplementary materials | Metrics

Get access | Share | Cite | Rights & Permissions

Appendices