

The 2-Pager+: Research Proposal Template for Empirical Work

Throughout the quarter, we will be using this template to track your progress toward writing a research paper. You don't need to have perfect or even complete answers, but our hope is that by the end of the quarter, you will have thought through each bullet carefully and taken a few iterations at answering it.

Due Dates:

- **March 29, 2022:** Complete sections 1-3 (max. 2 pages total)
- **April 24, 2022:** Revise sections 1-3. Complete section 4 (max. one page per section)
- **May 8, 2022:** Revise sections 1-4. Complete section 5 (max. one page per section)

Please attempt to answer all questions. If you're stuck on answering a particular question, try to describe what you're stuck on and which solutions you've considered. We can discuss the barriers you're facing in our meetings.

General Page limits: at most one page per section, plus an appendix for details (if necessary).

1. Research Question

- What are you trying to learn in this project?

Imagine the data, estimation, analysis, etc. all work out perfectly. How will you explain what the bottom line of your paper is to a colleague? Try to keep your answer focused and succinct: there should be one "main" question that you're after, not twenty.

- Why is this question interesting and important?

What are the economic lessons (e.g. tests of theories, estimates of portable parameters, new intuition) from the paper? Which policies might be designed differently based on the answer to this question? How large are the potential consequences?

- Where is the economics in your question?

Why is it important that you, an economist, and not a sociologist, political scientist, computer scientist, etc. answer this question? What perspective does economics add?

- How does answering this question contribute to the existing economic literature?

List and very briefly describe three related papers that you've recently read. How is what these papers explore similar/different to what you have in mind? What is your contribution compared to these papers?

- How does answering this question contribute to public discourse?

Does your question relate to issues being debated by regulators, politicians, activists, etc.? What is the state of that debate?

2. Institutional Setting

- What are some concrete examples of settings (an industry, a population affected by a particular set of policies, etc.) in which you can study your research question?

Be very concrete – treat this question as asking: where will you seek out data to answer your research question? Explain what characteristics of these settings make them good environments to ask your broader economic question.

- List three papers that study a similar setting. What types of questions do they examine?
It's okay to only skim abstracts/introductions for this question. The important thing is to identify people working on your industry/policy/etc. and get a sense of what they study.
- Could a similar question be asked in other settings?
Now zoom out. How general/common are the characteristics you listed above? Is what you can learn in these settings portable to other settings, industries, locations, or policy areas?

3. Data & Feasibility

- Which data exists, or could be collected, to answer your question?
What is your ideal data set? If there are multiple options, describe the one most easily available to you. You could use the appendix to describe back-up options.
- Which steps should you take to collect new data or obtain access to existing data?
If you will be creating your own data, e.g. through an experiment or by scraping a website, how will you collect it? Are there any data sets you need that are not in the public domain? If so, who needs to give you permission to use those data sets? Will you need to purchase the data or will you need to go through a formal approval process? Will your project require a data use agreement? Where will you store and analyze private data once you get such an agreement in place? Will you need IRB approval to collect or work with the data?
- Is it likely that you will have enough power to answer your research question?
Try to take at least a rough pass at a power calculation.
- Will you need to link different datasets? If so, which information will be used to perform the link, and is an imperfect link likely to affect your sample size?
Suppose you can only link 30% of records across data sets. Will you still have a large enough sample for your analysis?
- Which tests would you run on the data to check the validity of your research design?
- If you don't have access yet, are there any shortcuts to running these kinds of tests, e.g. can you ask other researchers who have worked with the data for help, or look at published research? Is there publicly available data you can use in the interim to learn basic facts?

4. Research Design

- What are the parameters you need to identify in order to answer your research question?
Your economic question will involve unknown economic quantities—for instance, treatment effects on productivity; price elasticities; agents' degree of risk aversion. What are the relevant objects that you will need to identify?
- What do you need to be true about the data-generating process in order for these parameters to be identified?
The answer to this question should have two parts: a list of assumptions, and an explanation of how these assumptions help in constructing a mapping from the data to the target parameters

(your identification argument). For the first part, you can describe a rough economic model or a quasi-experiment. Explain what the key ingredients are. For instance: Which variables do you assume to be exogenously determined? Who are the actors making decisions? What decisions do they have to make? What do they know/not know? What are the important factors driving observed outcomes? For the second part, loosely describe how you will map the observed data into the target parameters described in the previous question, using your assumptions. It may be helpful to refer to a paper that does something similar to what you have in mind.

- What is (roughly speaking) the order of most to least reasonable / credible among your assumptions?

Some assumptions are more odious than others. Some assumptions are more necessary than others. It's good to get a headstart in communicating clearly about this.

- What empirical facts will you want to show to describe features of the institutional setting that help support your identification argument?

You will need to explain how your setting works and argue that your model is consistent with the real world through empirical facts. What will you need to see in the data for this to work well? Optional: include a sketch of the key figures you would ideally create in the appendix.

- How will you implement estimation?

What will you need to do in order to estimate your parameters? Will you be able to apply existing methods? How "standard" are the methods? Will you likely need to modify them? Is there an existing software package to implement them?

- How will you calculate standard errors?

How will you incorporate measures of uncertainty around your parameter estimates? It's ok to give a short, high-level answer to this question. E.g., is there anything about the context that might require deeper thinking about the covariance matrix (e.g., units are not iid, spillover effects, etc)?

5. Planning

- What is a conservative timeline for the project? What are potential bottlenecks? If you are planning to use this for your second year paper, will you have enough time to get a draft together by the end of August?

The answer to this question is obviously not binding. Do your best to assess what timeline is feasible and we will discuss it in our meeting.

- What are the potential logistical issues?

Will you need to travel to collect or access data? Do you need special access to secure environments? Would pandemic-related constraints limit your ability to work on this project?

- What are the benchmarks that you'll need to hit for this project to be going well?

What will you need to have done, and by when? What will you do if you miss a benchmark?

- Will you require funding? If so, have you identified potential funding sources?

Be explicit here. E.g., include links to funding sources and be prepared to give examples of people who have recently successfully attained funding from them.

- What are your main concerns about this project? How do you aim to address those?