Guidelines for the resesarch proposal

Nurfatima Jandarova

Content

The proposal should include

- a clear research question
- motivation for research question
- contribution to existing works

The rest of the guideline depends whether you want to make a **theoretical** proposal or an **empirical** one. Both proposal types may include elements of the other. For example, a theoretical proposal may include discussion of how to test the theory with the data. Or an empirical proposal may include discussion of a theoretical framework that could help make sense of the estimates. The type you choose should be closely related to the research question and stated contribution. If the main objective of the proposal is to improve upon theory, then it is theoretical. If the objective is to get new estimates, it is empirical.

Theoretical proposal

- Clear description of model components and parameters
  You should clearly describe
  - who is an agent making a decision
  - what kind of information she observes
  - what decisions does she make
  - what are the constraints she is facing
- Simple predictions, comparative statics or simulations
- Bonus: outline of the plan to estimate model parameters with the data

Empirical proposal

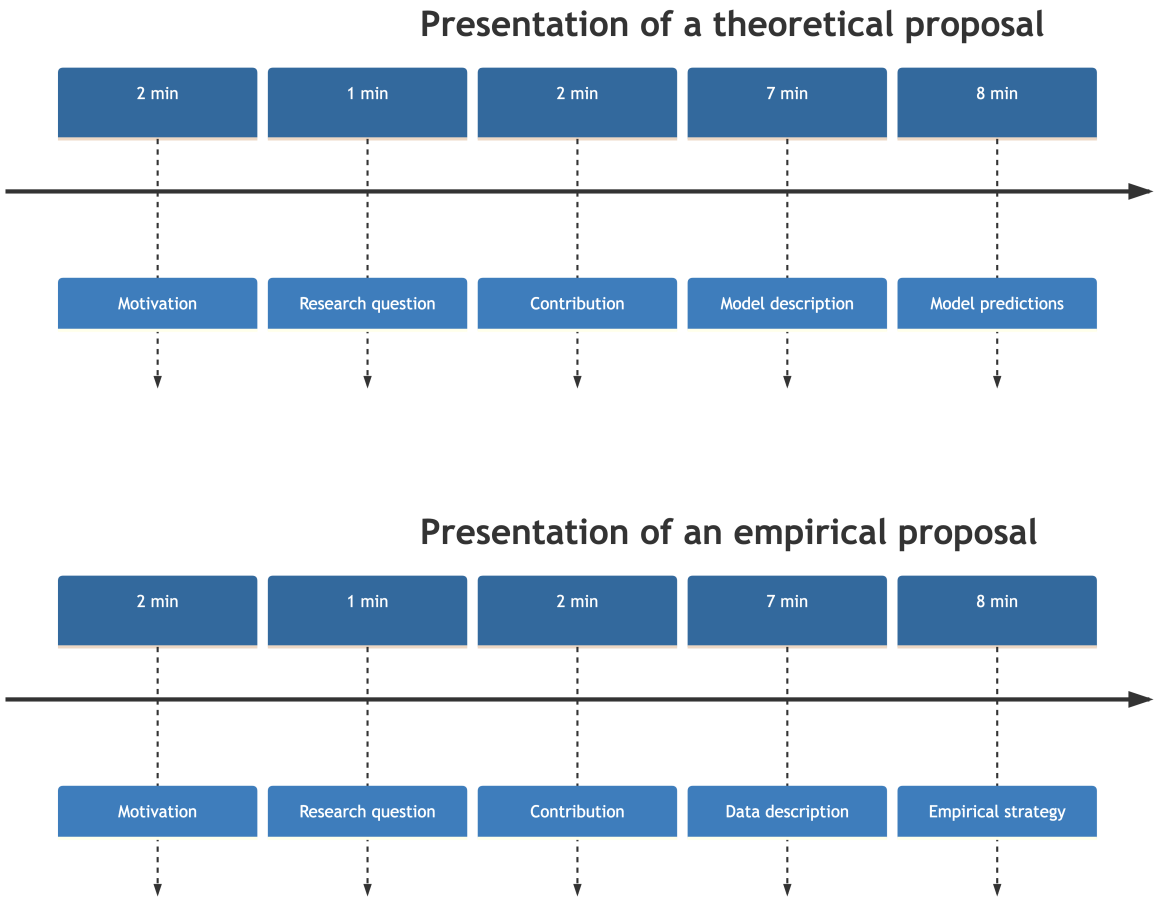
- Description of a potential dataset
  The question you want to answer here is why this dataset is suitable to study the question at hand.
  - *Sample members*: does the dataset include the people whose behaviour you would like to study?
    For example, if you want to study minimum wages in the fast food industry, you need to have sufficient amount of workers in that industry and maybe similar people that work elsewhere or don't work at all. Can the chosen dataset contain information about these people? You may not find exact numbers without actual access to the data, which is ok. But try to give as best an answer to this question as you can.
  - Key variables: does the dataset include the variables that you would like to study?
    For example, does it include information whether people were working or not, which industry they worked in, maybe the occupation they had? This information should be available in the documentation of any dataset, so you should be very precise about it.
    - Bonus point: you can discuss the ideal definition of the variable of interest.
      For example, would you like to study employment patterns of minimum wage workers compared to workers in other industries? Or compared to unemployed workers? Which of these are more appropriate for answering the research question.
- Discussion of empirical strategy
  - Clear statement of empirical strategy.
    For example, instrumental variables, or regression discontinuity, etc.
  - Clear definition of necessary components.
    For example, an instrument variable for IV, or running variable for RDD.
  - Discussion of assumptions required for the chosen method.
    For example, exogeneity and relevance for IV. You should discuss why you think chosen instrument may be relevant instrument and what could endanger exogeneity of the instrument.
- Bonus: outline possible mechanisms

Time structure of the presentation

Aim to spend first 5 minutes on motivation, statement of research question and potential contributions.

Rest 15 minutes should cover the details of the proposal.

Here are example time breakdowns of a 20 minute proposal presentation.



Examples

As I have mentioned at the beginning of the class, best examples are already published papers. Feel free to find your own papers to use as a template.

Below I provide two papers and point relevant parts demonstrating content points above.

Example for a theoretical proposal

Example based on Autor, D. H., Levy, F. and Murnane, R. J. "The Skill Content of Recent Technological Change: An Empirical Exploration*," *The Quarterly Journal of Economics* 118, no. 4 (November 1, 2003): 1279–1333.

Research question

> "what it is that computers do— or what it is that people do with computers—that causes educated workers to be relatively more in demand." (p. 1280)

Motivation

The authors give rather short motivation and contribution because the work was quite novel. The main motivation is by stressing the "gap" in the literature: previous literature documented positive correlation between computerisation and increased employment of educated workers without providing the causal framework.

Read first paragraph of the Introduction on pages 1280-81.

Contribution

The contribution is essentially given by the motivation. Once you identify the gap in the literature, filling that gap is usually the contribution. The authors provide some discussion of possible contributions in footnote 2 at the bottom of page 1280.

Model description

All the elements, underlying assumptions and rationale for those are discussed in detail in subsection I.A. on pages 1286-1289.

Model predictions

Some predictions are given in the last three paragraphs of subsection I.A. on page 1289. Industry predictions are described in subsection I.B. on pages 1289-91.

It is part of your job to identify what predictions from the model are of interest. You do not need to generate predictions in terms of every single input into the model. Choose one or two that are central to answering the research question.

Bonus: estimation in the data

Entire section II on pages 1291-95.

Example for an empirical proposal

Example based on Kleven, H., Landais, C. and Søgaard, J. E. "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics* 11, no. 4 (October 2019): 181–209.

Research question

> "the effects of children on the careers of women relative to men" (p. 181)

Motivation

First paragraph on page 181. The authors relate to a larger debate about gender pay gap and how despite convergence over the years, it is still not eliminated. They propose their hypothesis about the role of childrearing as a possible explanation.

Contribution

The contributions are discussed at length on pages 183-184. Pay attention to sentences that highlight the differences of the current paper relative to those that have already been published such as last sentence of the first paragraph on page 184.

Data description

Subsection I.B. on pages 186-87. Pay attention to sample selection criteria described in the second paragraph of the subsection. Many of the key variables are simply listed in the first paragraph because they are self-explanatory and their definitions are standard. However, pay attention to the discussion of construction of hours worked and wage rate variables. Especially, how the data limitations in their definition may contribute to biased results.

Empirical strategy

Subsection II.A on pages 187-88 provides description of the chosen methodology. Notice how careful the authors are in defining different elements of the regression equation.

Identification paragraphs in subsection II.B. on pages 195-96 discuss at length necessary assumptions for their empirical strategy to work as well as outline of robustness checks to strengthen their case.

Bonus: mechanisms

Section IV on pages 201-04.

References

Autor, David H., Frank Levy, and Richard J. Murnane. "The Skill Content of Recent Technological Change: An Empirical Exploration*." *The Quarterly Journal of Economics* 118, no. 4 (November 1, 2003): 1279–1333.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11, no. 4 (October 2019): 181–209.