Project
Econometrics II (Course No. 2534)
February 13, 2026

In this course, we will ask you to apply your Econometrics knowledge in a **project**. The goal is for you to go through a proper "research process" – starting with idea generation, thinking of an identification strategy, and then writing a concise project report. This should serve as an opportunity for you to

- find sub-fields of Economics that interest you,
- deepen your knowledge of Econometrics and get a feeling for how it is applied, and
- practice doing research before you have to do it anyways for your Bachelor's Thesis.

We recommend you to follow these **steps** while working on your project:

1. Find a sub-field and broad topic of interest.
2. Come up with a research question that includes a causal claim.
3. Think about your motivation to research this topic and why you think it is important.
4. Go through the literature to discover what is already out there, and if needed, adapt your question.
5. Come up with a model and search for appropriate data sources.
6. Identify issues of endogeneity and try to tackle them.
7. Iteratively refine model, data, and identification over the course of the semester.
8. Provide relevant summary statistics for your research question and data.
9. Finish your project report and hand it in.

The **first submission**, on November 16, should be a one-page (excluding references) document that briefly summarizes Steps 1–4 of the above.

The **final submission**, on February 13, should be a project report. The project report should have around **10 pages** (excluding references) and be **structured** like this:

1. Introduction (approx. 1.5 pages)
    - What is your motivation?
    - Give an account of the literature.
2. Method (approx. 4 pages)
    - The empirical specification
    - Discuss endogeneity and threats to identification.
    - Detail your identification strategy.
3. Data (approx. 3 pages)
    - Discuss data requirements of your model.
    - Discuss available data sources.
    - Show relevant summary statistics of the data you found.
4. First Results and Research Outlook (approx. 1.5 pages)
    - If you already have results, include them here.
    - Mention where you could go on from this (what would be the next steps, are there more interesting data sources, what would you need to add for this to become a proper thesis/paper?)

Please observe the following **rules** for the Project:

- You can work either **alone** or in a **group of two people** on the Project.
- The idea is for you to start thinking about the project **now**. During the semester, we will be available to support you. We recommend you to take at least one meeting with us. At the halfway point of the semester, we will ask you to submit a short account of your idea, and at the end of the semester, you will submit your final project document.
- The project counts **30 points** towards your grade,
  - **3 points** of which will be awarded for the short midterm document, and
  - **27 points** of which will be awarded to the final project.
- You may use **Large Language Models (LLMs) such as ChatGPT or Claude, but only under the following conditions**:
  - You may use LLMs only for help regarding the **code** in the empirical component (e.g., help adjusting a figure, questions about implementing certain models), or for **narrow, specific theoretical questions** (e.g. "Is a mediator in a DAG a type of confounder?", but NOT "Come up with an IV design for this research question."). You may also use LLMs to ask about **data sources** (e.g. "Recommend a dataset on housing prices in Angola"), but of course you have to follow any links you get and are not allowed to take data directly out of your LLM.
  - You are **explicitly NOT** allowed to use LLMs for **idea generation**, **interpretation** or for generating your submission **text** (you are, however, allowed to use it for correcting typos etc.).
  - You must save and submit documentation (e.g., PDF printout) of all conversations with LLMs (see below). This record must include both your **questions** and the **answers** of the LLM.
- At the end, **three files** must be submitted by the group:
  - Your main submission as a **PDF document**. We recommend to use LaTeX (e.g. via Overleaf) or **RMarkdown/Quarto**. The document should not contain any code. Other file formats than PDF (e.g. Word) will not be accepted and will be graded with 0 points.
  - An **R script** or a comparable file if you use Python, Stata, etc. This file must contain the **entire code used** and be executable.
  - A **record of all conversations with Large Language Models (LLMs)**, e.g. ChatGPT or Claude; for example, by exporting the conversation as a PDF. If there are multiple PDFs, you may also submit them in a ZIP file. Also observe the rules above for handling LLMs.
- If you are a group of two, each group member must submit an **assessment of both of your contributions** as a text submission. If a group consists of Persons A and B, Person A must submit an assessment in the following format:

```
Person A:        contributed 40 %
Person B:        contributed 60 %
```

  This serves to prevent "free-riding" and gives you a way to anonymously clarify such situations. **The specification of "contributions" is not a weighting of the points. Each group member receives the same score as long as the differences in contributions remain within reason.** Only if clear and blatant "free-riding" is evident, we will adjust the number of points given to that person.