
Homework 2: R / Stata Practice

Prof. Tzu-Ting Yang

May 13, 2026

INSTRUCTION

- You should submit Homework 2 before 5/31 (including 5/31).
- There is no late submission (zero points after the deadline).
- Homework 2 accounts for 15 points in your final grade.
- This homework will help you make progress on your term paper.
- You may use **either Stata or R**. Submit **two files**:
 1. **Empirical analysis record (HTML)**: A single HTML file documenting your current empirical workflow, generated from your code using markdown.
 - **Stata**: Write code and narrative in a `.stmd` file and compile with `markstat` using `"filename", bundle`.
 - **R**: Write code and narrative in an `.Rmd` file and knit to HTML via “Knit to HTML” in RStudio or `rmarkdown::render()`.
 2. **Answer sheet (PDF)**: A PDF answering all questions (1–6), including tables and figures.
- Recommended R packages: `haven`, `dplyr`, `ggplot2`, `fixest`, `rmarkdown`.
- Upload files here: <https://www.dropbox.com/request/1qkUrEkv02ygFaIJcrrX>
- File name format: `StudentID_YourName_HW2.html` and `StudentID_YourName_HW2.pdf`.

QUESTION 1 EMPIRICAL ANALYSIS RECORD

Submit an HTML file that documents your **current R or Stata code** and results. Generate the HTML from your code using markdown. The HTML should include:

- All code you have written so far (data loading, cleaning, visualization, regression).
- Output from each code block (tables, figures, regression results).
- Brief narrative text explaining what each step does.

Hint (Stata): Create a `.stmd` file with markdown text and Stata code blocks, then compile:

```
markstat using "your_file", bundle
```

The `bundle` option embeds all figures directly into the HTML.

Hint (R): Create an `.Rmd` file with YAML header output: `html_document` and embed R code in ```{r}` chunks. Render with:

```
rmarkdown::render("your_file.Rmd")
```

QUESTION 2 INTRODUCTION

Write a paragraph to introduce your research topic. Include:

- A brief description of your research question.
 - Why you chose to study this topic.
-

QUESTION 3 SAMPLE CONSTRUCTION

Write a paragraph describing your sample construction process. Include:

- Data cleaning steps.
 - Sample selection criteria.
 - Final sample size and time period.
-

QUESTION 4 DESCRIPTIVE STATISTICS

Create a table to display descriptive statistics of your sample and provide a brief explanation. Specify this as Table 1 in your answer sheet. Include:

- Means and standard deviations for your outcome and control variables.
 - Comparison between treatment and control groups if applicable.
-

QUESTION 5 EMPIRICAL METHODOLOGY

Describe your empirical methodology for estimating causal relationships. Include:

- Description of your identification strategy.
 - Mathematical expressions of your empirical model.
 - Key assumptions for this method.
-

QUESTION 6 PRELIMINARY RESULTS

Create a table or figure to present your findings, and write paragraphs explaining them. Specify this as Figure 1 (or Table 2) in your answer sheet. Include:

- A description of what you see in the table or figure (e.g., the sign, magnitude, and statistical significance of your main estimates).
 - An interpretation of what these results mean in the context of your research question.
-

QUESTION 7 AI ASSISTANCE

If you used generative AI (e.g., ChatGPT, Claude) to assist your research, describe how you used it. Include:

- What kind of tasks generative AI helped you complete; provide a specific example.
- The prompts you gave to the AI.

If you did **not** use generative AI, describe either (a) a concern you have about using AI in research, or (b) a specific type of task where you find AI unreliable or untrustworthy and explain why.