Assignment: Replication and Extension of Empirical Paper

Ph.D. Environment and Development Economics
UMN Department of Applied Economics
Instructor: Raahil Madhok

Due Date: October 21st, 2024

1 Overview

The goal of this assignment is to conduct a replication and extension of a published empirical paper in the field of environmental and development economics. This exercise is designed to help you engage deeply with research in your field of interest by reproducing key findings and contributing your own insights through an extension of the analysis. The goal is to enhance your understanding of empirical methods, familiarize yourself with the challenges of coding, and encourage critical thinking about existing research.

You will select an empirical paper of your choice, replicate and extend it, and submit a short write-up explaining what you did. You will also submit your code.

2 Assignment Components

2.1 Select a paper

- Most papers on the syllabus have replication files.
- Choose a paper that aligns with your research interests
- You may choose a different peer-reviewed paper in environmental & development economics that has publicly available data
- Note: you will benefit most from selecting a paper with methods you intend to use in your own research.

2.2 Replication Code

- Replicate **2-3 of the main tables or figures**. These should represent the core empirical findings
- You can use Stata or R
- Write your own code for the analysis. Try to do this yourself before consulting the authors' code. Doing so will help you identify potential sources of divergence between your results and the published findings.
- Extensively comment your code to explain exactly what each part does
- Output all tables to LaTeX. If you do not know how, use ChatGPT to help you.

2.3 Extension

- After replicating the main tables, you will **extend the analysis**
- The extension is up to you, but should be carefully thought out and yield a new insight
  - I reserve the right to determine what is "carefully thought out''
  - E.g. if you drop the five largest outcome values and re-run the analysis, this **does not count**
  - But, if you plot a histogram, find outliers, and truncate at the 95th percentile, **that counts**
- Some ideas for an extension are:
  - **Enhance the research question**: this could mean studying heterogeneity or non-linearities.
  - **Robustness checks**: Verify the original results under different assumptions or specifications.

- **Error identification and correction**: If you identify any mistakes in the original analysis, provide a corrected version with clear explanations.

3 Write-up (6 pages double spaced)

After conducting the replication and extension, you will explain what you did in a short write-up. The report should contain the following:

1. **Summary:** A one paragraph summary of the paper's research question, setting, methodology, and main findings
2. **Replication Process**: Describe which tables/figures you opted to replicate, the steps you took to replicate the results, and any issues you encountered. Include a detailed explanation of your data processing and analysis.
3. **Results**: Present your replicated tables/figures and compare them with the original paper's results. Discuss any differences.
4. **Extension**: Explain the extension you undertook, your motivation for doing so, and the results you obtained. Provide any new tables or figures related to your extension.

3.1 Submission Format

- Submit your write up as a PDF in Canvas
- Submit your code: either link to repository in your write-up or upload through Canvas
- Format:
    - **Length:** Max 6 pages double spaced
    - **Format:** 12 point font, 1 inch margins
    - **Code:** Clear, concise, and commented

3.2 Grading Criteria

You will be evaluated based on:

- **Replication Quality**: Completeness, accuracy, and clarity of your replication.
- **Extension**: Insightfulness and rigor of the extension.
- **Writing**: Clear, well-structured, and concise writing that follows academic conventions.
- **Code**: Organization, clarity, and documentation of the code, including reproducibility of results.

Category	Criteria	Points	Description
Written Report (60 Points)	**Replication Process (15 Points)**		
	Description of Methods and Process	10	Clear explanation of replication steps, including data cleaning and methodology.
	Data and Documentation Review	5	Thoughtful review of data quality and documentation.
Replication Results (15 Points)	Presentation of Results	10	Clear presentation and formatting of replicated tables/figures.
	Comparison with Original	5	Detailed comparison of replicated results with original findings.
Extension (20 Points)	Description and Justification	10	Clear explanation of extension and rationale for its connection to the original paper.
	Results of the Extension	10	Presentation of findings from the extension, including new tables/figures.
Writing Quality (10 Points)	Clarity and Structure	5	Report is logically structured and clearly written.
	Grammar and Style	5	Proper grammar, spelling, and adherence to scientific writing conventions.
Code (40 Points)	**Code Quality and Organization (15 Points)**		
	Clarity and Commenting	10	Code is clearly written with sufficient comments.
	Organization	5	Code is well-organized with logical structure and consistent formatting.
Correctness and Functionality (15 Points)	Accuracy of Replication	10	Code accurately replicates the main results.
	Extension Implementation	5	Correct implementation of the extension in the code.
Reproducibility and Efficiency (10 Points)	Reproducibility	5	Code can be easily reproduced with clear instructions.
	Efficiency	5	Code is efficient, avoiding unnecessary complexity or overhead.
		Total	**100 Points**