Econometrics (Econ 308)
Project Guidelines

Fall 2021
Instructor: Andra Hiriscau

1 Introduction

Due December 17^{th} by 11:59 PM. The paper must be uploaded to Blackboard in a PDF file. The Turnitin plagiarism checker is enabled, meaning that your paper is compared to a massive database of student work, websites, books and articles.

This is a group project. You must form a team of two or three individuals by September 26^{th}, and register your team in excel spreadsheet (available here). You will not be allowed to switch teams after September 26^{th}. The final paper must include all names. If your partner drops the class during the semester, you will need to complete the project on your own.

2 Research Question

The purpose of this project is to apply the analytical and quantitative skills acquired in this course. It would be best if you found an interesting research question. Examples for typical research questions: Environmental Economics: Can city air quality be improved by implementing driving restrictions? Crime Economics: How does increased police presence affect crime rates? Education Economics: Do smaller class sizes promote better student performance? Family Economics: Do children contribute to increasing gender wage gap?

Additional Resources

Econometric Analysis Undergraduate Research Papers website includes examples of Econometric projects authored by undergraduate students. All materials are protected under U.S. Copyright Law and all rights are reserved, but it should provide you with an idea of how the final project should look like. This Sample Paper in Econometrics shows how to communicate econometric work in written form.

Your research project should **add something new**:

- Add a new variable whose influence has not been studied before. You can download the data for a specific paper, use the data provided by the book or find your own data.
- Expand economic questions to include factors from other sciences.
- Study an existing question with more recent data.
- Use a new data set or study a question for a different country.
- Find a completely new question (hard but possible).

3 Literature Review and Data

Search for published papers on the chosen topic using tools such as EconLit, Google Scholar, the Journal of Economic Literature (JEL). A literature review is essential to place your paper into context. It can be part of the introduction or a separate section. It would help if you summarized the findings for at **least two**

papers. You need to decide on the appropriate data set. Many questions can, in principle, be studied using cross-section data.

Data Sources:

- IPUMS. It provides census and survey data from around the world (9 datasets).
- U.S. Census
- Panel of Study of Income Dynamics
- You are welcome to find other data

4 Writing an Empirical Project

A successful empirical paper combines a careful, convincing data analysis with good explanations and a clear exposition.

1. **Formatting (5 pts)**
   - **3 pts.** Choose a title that is exciting and reflects the paper's topic. Papers should be typed, double spaced, 12pt and Times New Roman font. The paper should be no more than 10 pages, excluding title, graphs, tables and references.
   - **2 pts.** Number equations, graphs and tables. Refer to papers by author and date (eg. Angrist (2020)). Use APA citation style (APA Guide). You can use the first person point of view when discussing your research steps ("I studied ...") and when referring to yourself and your co-authors ("We examined the literature ...").
2. **Introduction (20 pts)**
   - **5 pts.** State basic objectives and explain why the topic is important.
   - **5 pts.** Include literature review: What has been done? Discuss the results of **at least two** papers.
   - **10 pts.** *What do you add to the literature? (see section 2)*
     You can grab the reader's attention by presenting simple statistics, paradoxical evidence, topical examples, or challenges to common wisdom. One may give a short summary of the results in the introduction.
3. **Data (20 pts)**
   - **5 pts.** Name the sources of your data and how they can be obtained.
   - **5 pts.** Discuss the units of measurement of the variables of interest.
   - **10 pts.** Present summary statistics for the variables used in the analysis in a table (mean, standard deviation, minimum, maximum, number of observations) and include a summary table with them. All categorical variables, variables that are not continuous (e.g., race, gender), need to be presented by category.
4. **Empirical Methodology (20 pts)**
   - **5 pts.** Specify the population model you have in mind and write the equation you estimate. Example: Effects of alcohol consumption on college GPA.

$$colGPA = \beta_0 + \beta_1 alcohol + \beta_2 hsGPA + \beta_3 SAT + \epsilon$$

   - **5 pts.** A convincing discussion of what variables to control for is essential. Are your results causal or not? Discuss potential concerns with sample selection, omitted variables etc along with the consequences of those problems.

- **10 pts.** Then discuss estimation methods and how you measure the variables in your model:
  - When using OLS: Discuss why exogeneity assumptions hold and how you deal with heteroskedasticity, serial correlation, and the like.
  - When using panel methods: Explain what the unobserved individual-specific effects stand for and how they are removed/accounted for.
  - When using IV/2SLS: Explain why your instrumental variables fulfill the assumptions: 1) exclusion, 2) exogeneity, 3) partial correlation. (time permitting)

5. **Results (30 pts)**
   - **5 pts.** Present your results in a table that includes coefficients, standard error, R-squared and number of observations.
   - **10 pts.** Write about the magnitude and interpretations of your coefficients. Are they statistically significant (includes significance levels in tables)? Do they have or not the expected sign (this may indicate a specification problem, for example, omitted variables)?
   - **10 pts.** Hypothesis testing: discuss R^2 and conduct an F-test on your regression results.
   - **5 pts.** Predict outcomes for two hypothetical people or situations.
6. **Conclusion (5 pts)**
   - **4 pts.** Summarize main results and conclusion from them.
   - **1 pts.** You can suggest directions for further research.

5 General Guidelines

- Your paper should be a coherent essay.
- Carefully choose functional form specifications (logs, squares etc.).
- Do not include variables that are listed as numerical values but have no quantitative meaning (e.g., 3-digit occupations). Transform such variables to categorical variables.
- For ordinal dependent variables, there are ordered logit/probit models.
- Understand how missing values are coded (e.g., in some data missing values are coded as "999" or "-1"), understand units of measurement of your variables, check your data for implausible values.
- Tables and Figures:
  - Use *estout command* to create the tables; do not cut and paste the STATA output.
  - Each table or figure should have a title.
  - Rename the variable names in tables and figure (e.g., if you have *momeduc* in your data, you should have *Mother's Education* in the table).
  - There should be a footnote which explains what is shown (e.g., Each entry is an OLS coefficient with standard errors in parentheses. Significance levels: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$)
  - Make sure you label X-axis and Y-axis

6 Deadlines

Description	Details	Deadline
Topic	Paper title and 250 word summary	7-Oct
Data and Literature Review	Discussion of the data sources, summary statistics and discuss at least two related studies.	9-Nov
Rough Draft and In-Class Peer Reviewed Discussions	Paper draft	2-Dec
Final paper	Final paper	12-Dec