

ECON 340

Economics Research Methods

Div Bhagia

Final Submission Guidelines

Final Paper: Overview

- The paper should reflect:
 - your ability to perform and interpret statistical analyses
 - your writing ability (use Grammarly)
- Carefully read all the instructions on course website
- Incorporate the feedback you have gotten so far
- Rubric provides a breakdown of the score for different components

Final Paper: Submission

- The paper should be in essay format divided into sections (around 6-8 pages)
- Comply with the formatting requirements
- Submit your write-up in PDF format on Canvas

Research Paper: Sections

- Structure and content:
 - Introduction
 - Literature Review
 - Descriptive Analysis
 - Empirical Strategy
 - Results and Analysis
 - Conclusion
 - References
 - Appendix (optional)

Instructions include the *suggested* length for each section.

I. Introduction

- What is your research question? Why is your research question interesting? Why is your topic important?
 - This was Task 1 for the interim submission
- Preview your results

People who do not have time to go through the entire paper should get the general idea and conclusion of the paper through the introduction.

II. Literature Review

- What other work has been done on these issues? What has been found?
- Look for papers on Google Scholar (ChatGPT makes up papers that don't exist!)
- The literature review must include at least two academic papers.
- Cite sources using the Chicago Manual of Style Author-Date system

III. Descriptive Analysis

Tasks 3-5 from Submission 2:

1. Describe the data and your variables of interest. Mention if you constructed new variables or deleted observations.
2. Present summary statistics for all your variables and comment.
3. Present a graph showing the relationship between your dependent and primary independent variable
 - Scatter plot if both variables are continuous
 - Bar plot if one variable is binary

Summarize what you learn.

IV. Empirical Strategy

Explain how you plan to answer your question(s).

- Write down the main regression model and explain why it makes sense.
 - Interpret the main coefficient and what you hope to learn from it.
 - Why did you pick the controls that you picked? Talk about how excluding these controls may impact your main coefficient of interest.

Empirical Strategy: Example

The objective of this paper is to investigate whether class size has an impact on test performance. To address this question, we estimate the following regression model:

$$testscr = \beta_0 + \beta_1 str + \beta_2 high_comp_stu + \beta_3 meal_pct + u$$

The main coefficient of interest is β_1 , which corresponds to the student-teacher ratio variable. This coefficient captures the change in test scores associated with an additional student per teacher, holding other variables in the model constant. We expect β_1 to be negative if smaller class sizes are associated with higher test scores.

Empirical Strategy: Example (cont.)

However, it is possible that β_1 could be capturing the impact of other factors that are correlated with class size and test scores, such as better infrastructure or funding in the school district. To control for these potential confounders, we include two additional variables in our regression model. These variables are *high_comp_stu* and *meal_pct*.

high_comp_stu is an indicator variable that takes the value of 1 if the school has more computers per student than the median school and 0 otherwise. We expect that schools with more computers per student are associated with higher test scores, as computers can provide students with access to a wider range of educational resources and facilitate learning. Additionally, we expect schools with more computers also to have smaller class sizes. Therefore omitting this control variable could overestimate the effect of class size on test scores.

We also control for *meal_pct*...

V. Results and Analysis

- In one or more tables, present the results of your regression analysis. Adding controls one by one.
- For this example, the code would be:

```
1 model1 <- lm(testscr ~ str, data)
2 model2 <- lm(testscr ~ str + high_comp_stu, data)
3 model3 <- lm(testscr ~ str + high_comp_stu + meal_pct, data)
4 stargazer(model1, model2, model3, type="text", keep.stat =
  c('n', 'adj.rsq'))
```

V. Results and Analysis

- Interpret the sign, value, and significance of the main coefficient across all specifications.
- Comment on how and why the coefficient changes or does not change as you add each control.
- Comment on the coefficients on your control variables.
- Interpret the adjusted R^2 of your models
- Discuss what you can conclude from your regression results. Correlation vs causation? Are there other omitted variables you are worried about?

	testscr		
	(1)	(2)	(3)
str	-2.28*** (0.48)	-1.74*** (0.49)	-1.00*** (0.25)
high_comp_stu		7.99*** (1.84)	1.92** (0.94)
meal_pct			-0.59*** (0.02)
Constant	698.93*** (9.47)	684.32*** (9.86)	699.36*** (5.00)
Observations	420	420	420
Adjusted R ²	0.05	0.09	0.77

VI. Conclusion

- Summarize your results. Given this summary, answer your research question.
- Expand on how your research findings connect to broader implications. Future research? Policy? Wider context?