Projects:
Writing Tips and More

Jesper Riis–Vestergaard Sørensen

University of Copenhagen, Department of Economics

February 17, 2025

Assessment and Feedback

Assignment open-ended—no "correct" result

Assessment mainly based on your:

- Presentation of econometric theory.
- Ability to apply theory in empirical analysis.
- Ability to describe underlying thought process.
- Explanation and arguments for choices made.

Assessment criteria on Absalon. Read carefully.

Peer feedback based on assessment criteria.

- Focus on how to improve each report (incl. your own).

Writing and Structure: Assessment Criteria

Assignments must be written in a formal and precise language with an academic structure (...)

In particular (...) it should be clear which question is addressed, how it is done, and what the conclusion is.

Writing Tips

Crucial:

- Your report must be readable on its own.
- Presume no prior knowledge of assignment, but general knowledge about econometrics.
- Write it for student who completed course in past.
- Explain what you have done and why you have done it
  - Such that your empirical analysis is reproducible.

Writing Tips

Additionally:

- Write in a formal academic language.
- Think about the structure of your assignment.
- Be precise and write in a compact way!
- Be consistent with notation and avoid typos.
- Present theory and results in logical order.

Structure (Suggestive)

1. Introduction
2. (Data)
3. Econometrics
4. Empirical analysis
5. Discussion
6. (Conclusion)

Introduction

What to include:

1. Main question(s): What is it? Just one or two main questions, not ten!
2. Motivation: Why of interest?
3. Econometric model/method: Which econometric model(s) do you use to answer the question?
4. Conclusion: What do you conclude? Results in line with expectations/economic theory?

Do:

- Write as if no prior knowledge of assignment.
- Short and to the point!

Econometric Theory: Assessment Criteria

Sections involving econometric theory should include:

1. A precise definition and interpretation of the model considered and its properties.
2. A precise description of the estimator used. In particular, a precise account of the assumptions used to derive the estimator must be given.
3. A precise presentation of the necessary assumptions for consistent estimation and valid inference. This includes a precise presentation of the null and alternative hypotheses, test statistics, decision rules, and (asymptotic) distributions used to test relevant hypotheses.
4. The theory must be presented in a logical order with a correct and consistent use of notation.

Econometric Theory

What to include:

Model: Present econometric model(s) you use. Start with equation, then describe assumptions, properties and interpretation.

Estimator: Which estimator(s) do you use? Describe assumptions used to derive estimator. State estimator (either equation or optimization problem that the estimator solves).

Properties: What are properties of estimator? Under which assumptions? How can you do inference? How can you test restrictions imposed?

Econometric Theory

Do:

- Write general model for your problem.
- Present model, then discuss. Define variables/parameters.
- Explain if economic theory imposes model restrictions, e.g. parameter restrictions?

Don't:

- Include irrelevant theory.
- Explain estimator you use before stating model.
- Copy-paste "generic" expressions from slides/notebooks. E.g., expression for estimator $\widehat{\boldsymbol{\beta}}$ based on (y_i, x_i) when you have $\widehat{\boldsymbol{\theta}}$ and $(\texttt{wage}_i, \texttt{educ}_i)$.

Discussing Assumptions: Example

Model $y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + c_i + u_{it}$ with y_{it} wage, and $\mathbf{x}_{it}$ contains education. Want to discuss possibility: $E(c_i|\mathbf{x}_i) = 0$.

Do:

- *"IQ is not observed and could result in omitted variable bias in POLS: If IQ is time-constant and affects worker productivity, it might affect wages and thus belongs in c_i. Furthermore, if high-ability individuals take longer education, then there is a correlation between c_i and* educ_{it}, *invalidating $E(c_i|\mathbf{x}_i) = 0$."*

Don't

- *"There could be omitted variables."* (Which?)
- *"Maybe c_i is correlated with $\mathbf{x}_{it}$."* (Why?)
- *"High-ability workers have longer education."* (Where?)

Brief and on Point

Do: Enough to illustrate a mechanism

- *"Time-constant worker ability, such as IQ, may affect productivity and thereby wages. Thus, it belongs in c_i."*

Don't: Give too much detail.

- *"Imagine one worker that is very smart, and another worker that is not. The firm may be able to produce more output with the smart worker, for example if the smart worker can improve management skills by introducing a LEAN process. Conversely, the less smart worker is unable to do so and would, comparatively, be less productive. A more productive worker might command a higher wage if labor markets are competitive, since firms will outbid each other in competition for the best workers, raising wages until they equal marginal product. In this way, there is a positive correlation between ability and wages."*

Empirical Analysis: Assessment Criteria

The empirical analysis should include:

1. A description of the model selection process.
2. A presentation and discussion of the relevant empirical results, such that the reader is able to understand and reproduce the steps carried out in the process as well as the conclusions made. The results must be presented in suitably formatted tables and/or figures. Irrelevant information should be left out of the tables/figures.

⋮

Empirical Analysis: Assessment Criteria (Continued)

$\vdots$

3. A discussion of whether the assumptions underlying consistent estimation and valid inference are likely to be satisfied/violated in the present empirical setting. This may include various forms of (mis)specification testing.
4. A clear conclusion to the main question and a discussion of the limitations of the approach used to reach the stated conclusion. Real-world examples of behavior/phenomena that may fit or invalidate specific model assumptions should be related to the econometric model.

Empirical Analysis

What to include:

- Empirical results: Results in formatted tables.
- Specification testing:
  Where do you start? Why?
  Which tests? Conclusion?
  How do you proceed? Why?
  Where do you end? What is your preferred model? Why?
  Model assumptions fulfilled? If not, consequences?
- Interpretation: Implications? Expected?

Empirical Analysis

Do:

- WHY? Justify your choices!
- One of the most important parts of the assignments!
- Explain what you do and present the relevant results, so that your empirical analysis can be reproduced.
- Use the tools from class.
- Refer to equations in earlier sections.
- Discuss misspecification tests before statistical testing.
- Clearly state conclusion to main question.

Empirical Analysis

Don't:

- Repeat equations already stated.
- No canned software. No screen dumps. No 8^{th}(!) decimal.
- Exclude relevant results (e.g., only report estimates for some model parameters).
- Only report results for "final" model.
- Forget to discuss test results.

Discussion and Conclusion

What to include:

- Limitations of approach: Weaknesses? Alternatives? Sensitivity? (E.g. adding/dropping time dummies.)
- Results comparison with existing literature, if available.
- (Concluding remarks): Briefly conclude wrt. main question. Link back to introduction.

Presentation of Results

Do

- Include and feature relevant results (only!):

  Coefficient estimates, s.e.'s , and other relevant (test statistics, p-values, log-likelihood, variance estimates).

  Gives reader understanding of process and choices.

- Format tables/figures: Enumerate, label, add caption, variable/model names, etc.

Don't

- Copy and paste Python output!

- Include irrelevant information in the table.

A Presentable Table

Table 1: Coefficient Estimates: Demand for Apples

Variable\Method	OLS	FE	Sub	All
Constant	-0.13^{***}	–	-0.20^{**}	0.01
	(0.012)		(0.108)	(0.149)
Price	-0.13^{***}	-0.43^{***}	-0.19^{**}	0.01
	(0.014)	(0.015)	(0.112)	(0.135)
Household Dummies	No	Yes	No	Yes
N	10,313	10,313	10,313	10,313
T	10	10	1	10

Notes: Outcome variable is ln(Apples). Heteroskedasticity robust standard errors [based on Eq. (4.20)] are in parentheses. Stars indicate statistical significance at the 1(***), 5(**) and 10(*) percent levels, respectively. "Sub" refers to the $t = 1969$ subsample (...)

Python

- You must submit your code along with your report.
  - One organized Jupyter notebook alongside `.py` files.
- It will only count in your favor.
  - If errors occur, may check if correctly applied.
  - We will not search for errors if report appears correct.
- Code from exercises: does not have to be cited.
- Do not use canned routines from outside course.