Lab 3: an AI agent's report, checked

What a four-minute analysis gets right, what it misses, and how your report avoids both

Magnus Lodefalk
Örebro University

Economics Research and Communication (NA310G)

28 October 2026



Today’s goals

By the end of today you can:

1. Check one AI-written results section: does the text match the numbers and the code?
2. Stack files with `append`, and write the covariates once with a local
3. Know the five things that most often send a report back for revision, and check your own against them
4. Use AI as this course allows: you write, it explains and hints, and you answer for every number

From last time

Answer on your own first, then compare with a neighbour

1. Hausman rejects. Which model, and why?
2. Why did our instrument lose its first stage with fixed effects?

We gave an AI agent the assignment

- A fresh project, the course data, the codebook, and the prompt a student might write:

  "Do firms that advertise for AI skills have higher salaries per employee? I need descriptive statistics, pooled OLS, fixed effects and random effects with a Hausman test, the usual diagnostic tests, and if possible an IV … write my results section."

- Under four minutes later: six do-files that run without an error, two tables, 1,045 words (it was asked for 600 to 800)
- Then we planted a few errors of the kind we see in reports

Files: na310g-lab3-agent-run.zip on the course site

Its headline result

	Pooled, raw	Pooled	Fixed effects	Random effects
AI firm (per cent)	17.5	17.1	3.1	4.3

- “The within-firm estimate is a modest 3 per cent (0.031, SE 0.007)”; its log gives a 95% interval of 1.7 to 4.6 per cent
- Hausman rejects: fixed effects. 327 firms switch AI status at least once

How many serious problems will we find in it?

Your bet first: hands up. The answer is on the next slide.

A. None: it is ready to hand in
B. One or two
C. Three or more

Your turn: referee it (40 minutes, in pairs)

- In the zip: `referee_sheet.pdf` (A text, B regressions, C descriptives, D sample code); answers in `referee_form.docx`

1–3. **Numbers, tests, code:** is every number and test read right? Does B do what A says?

4–5. **Words:** two claims beyond what the data can show, or a number never explained

6–7. **Data:** two things in the descriptives (C) the text should have noticed and handled

One finding per pair on the board, or a tick. No chatbot: the follow-up conversation tests this skill.

Why you still could not hand it in

- The course rule: AI may explain, debug, hint and improve your language; it may not run the analysis or write the report. This is both
- It answers no question of yours: no outcome, industry and county you chose
- The follow-up conversation would ask about your own numbers and your own code
- It even marked every do-file “code in this file was written with Claude”: honest, and not allowed

Researchers let AI draft because they can already verify. Here you learn to verify: you write, AI tutors.

append: stack files with the same variables

```
use "$root/data/processed/sample.dta", clear
keep if year <= 2020
save "$root/data/processed/early.dta", replace
use "$root/data/processed/sample.dta", clear
keep if year > 2020
append using "$root/data/processed/early.dta"
isid firmID year
```

- `merge` adds *variables* (on a key); `append` adds *rows*
- Same variable, two names (`emp`, `employees`)? You get both, each half empty, and no error: check the names first

---

Count before and after (5,020 + 3,802 = 8,822), then `isid`.

A local macro: write the covariates once

```
local controls ln_emp ln_cap
regress ln_fin ai_firm `controls' i.year, vce(cluster firmID)
testparm `controls'
xtreg ln_fin ai_firm `controls' i.year, fe vce(cluster firmID)
xtreg ln_fin ai_firm `controls' i.year, re vce(cluster firmID)
```

- `local controls` … names a piece of text; `` `controls' `` (backtick, apostrophe) puts it back. 15 controls? Change one line, not five
- The agent's do-file did the same: `local X ai_firm ln_emp ln_kpe ny_solid`

Run the `local` line *with* the models. Run `regress` alone and `` `controls' `` is empty: no controls, and no error.

Five things that most often mean a revision

1. A test read backwards: write the null hypothesis before the verdict
2. The main coefficient without sign, size, unit and significance in words
3. “Effect” where the design only shows an association
4. Tables without notes (which errors, which sample, which stars)
5. The appendix: do-files, final log, AI declaration, verification line

---

`report_checklist.pdf` **in your project: go through it before you hand in.**

Two short lines in your appendix (examples)

The verification line (copy your folder first if nervous; `data/raw` is never touched):

```
I deleted data/processed and output, ran master.do, and got
N = 4,212 and b = 0.031 as in Table 2.
```

The AI declaration (three lines, as journals ask):

```
1. ChatGPT (GPT-5). 2. Explaining the error r(198), checking my
reading of the Hausman test; two lines in 4_regressions.do carry
* AI:. 3. I reviewed everything these tools produced, and I take full
responsibility for the content of this report, every line and number.
```

Now: your report, with help in the room

- **First 30 minutes:** read your own draft with the three lenses: *numbers* (does the text match your log?), *words* (does a claim go beyond your data?), *code* (does it do what the text says?). Then `report_checklist.pdf`
- **AI now:** as a tutor (the tutor prompt on the AI page); paste code and error messages, never data; tag lines from AI with `* AI:`
- **Do-files:** add your code, never your name (`2_sample_vars_K7Q2MX.do`), and update `master.do`
- Raise a hand; today's questions become FAQ entries on the site (no names)
- After 15:00: help desk in N4008 until 16:00, then the help desks until 6 Nov

---

Deadline Sun 8 Nov, 16:00.

Summary

1. An agent writes a running analysis in minutes; its numbers matched, its understanding did not
2. Check what the words claim against what the data can show, and the code against the text
3. Your report: the checklist, the verification line, the AI declaration, and every number yours

After the deadline: the thesis template do-file. For your thesis: Serrano, the real data behind our twin, free for ORU students.

Test yourself before next time

Without AI

1. Why can salaries per employee rise in a firm that pays nobody more?
2. Write your own main coefficient in one sentence: sign, size, unit, significance.
3. What can your outcome not show? One sentence.
4. Write the null hypothesis of the Hausman test in one sentence.

Appendix: three more things about locals

```
local controls ln_emp ln_cap
display "`controls'"
local controls_long ln_emp ln_cap ///
                    ny_solid
display "`controls_long'"
```

```
ln_emp ln_cap
ln_emp ln_cap                       ny_solid
```

1. `display` shows what a local holds: check it before a long run
2. `///` at the end of a line continues the command, also inside a local
3. Never put `///` inside a `*` comment: Stata joins the next line to the comment, and that line silently never runs