

Report checklist (NA310G, HT26)

Go through this before you hand in. It lists the things that most often sent a report back for a revision in earlier terms. Tick it only when you have checked it in your own file.

The question and the data

- ☐ One research question, stated in the first paragraph, with your outcome, `ai_firm`, your industry and your county.
- ☐ One relevant paper summarised (the reading list in the brief is a good start), and a hypothesis you can test, written as a null and an alternative.
- ☐ The sample in words: which firms, which years, how many firm-years and firms, and what you dropped and why.
- ☐ Descriptive table (N, mean, median, SD, min, max) and a correlation matrix, both discussed in the text.
- ☐ Costs are stored as negative numbers in the accounts: you flipped the sign before taking logs.
- ☐ `x > 0` in Stata is true for missing values: every count or dummy on a variable with missings has `& !missing(x)`. `misstable summarize` shows which variables have missings.

The models

- ☐ Pooled OLS, fixed effects and random effects in one table, year effects in all three.
- ☐ Standard errors clustered by firm, and the table note says so.
- ☐ Hausman test run on models WITHOUT clustered errors (`sigmamore`), or the robust version with the firm averages of your regressors AND of the year dummies (as in Lab 2); null hypothesis stated; your choice follows from it.
- ☐ One heteroskedasticity test and one multicollinearity check, each with its number and what you did (heteroskedasticity: you already cluster, so say so; a high VIF on a control: say whether it matters for your coefficient).
- ☐ Optional IV: only with an argument for relevance AND exclusion, the cluster-robust first-stage F reported, or a short paragraph on why no instrument is credible.

The interpretation

- ☐ The main coefficient in words: sign, size, unit, significance. A dummy in a log outcome: about $100 \times b$ per cent (exactly $100 \times (\exp(b) - 1)$).
- ☐ Percent and percentage points kept apart.
- ☐ Association, not effect, unless your design earns the word: fixed effects remove permanent differences between firms, not everything.
- ☐ For fixed effects you report the within R-squared, and you did not choose a model by R-squared.
- ☐ The conclusion answers the question, gives one policy implication and one limitation.

The form

- ☐ At most 2,500 words, tables, references and appendix not counted.
- ☐ Title page, page numbers, Harvard references (Wooldridge, J.M. (2025) *Introductory Econometrics: A Modern Approach*. 8th edn. Cengage.), English decimal points (0.25, not 0,25).

- ☐ Word count (at most 2,500): the text of sections 1-8, the article summary included; not the title page, tables with notes, references or appendix. Copy that text into a new document and count there.
- ☐ Every table and figure has a title, a note (sample, standard errors, stars) and is discussed in the text.

The appendix (one PDF by e-mail, subject “NA310G: Assignment”, named with your code, e.g. 2HA9JR-FINAL.pdf: the report, then the appendix)

- ☐ Your do-files and the final logs (not the ones with every trial and error), each printed to PDF and named so you can tell them apart: 0_settings-do.pdf, 0_settings-log.pdf, and so on.
- ☐ The verification line: “I deleted data/processed and output, ran master.do, and got N = ... and b = ... as in Table ...”. data/raw is never touched by this; copy your whole folder first if you are nervous.
- ☐ The AI declaration, three lines (see the AI page): tools and versions or “none”; what for and how, and which lines carry * AI :; the responsibility sentence.
- ☐ No data and no listings of individual firms’ rows (list, browse) in the files; no names anywhere, not even your own: your code on the title page and in the file name.

Finally

- ☐ You can explain every line in your do-file and every number in your tables without notes. Some students are invited to a short follow-up conversation about exactly that.