Session Report: Panel Data I

Applied Quantitative Methods for the Social Sciences II

2026-02-19

Contents

Summary

Artifact	Status	File Path
Slides (presentation)	PASS	`slides/05_panel1/panel1.pdf`
Slides (notes)	PASS	`slides/05_panel1/panel1_notes.pdf`
Assignment	PASS	`assignments/assign5_panel1.pdf`
Solution Part 1	PASS	`assignments/solution/assign5_panel1_part1.pdf`
Solution Part 2	PASS	`assignments/solution/assign5_panel1_part2.pdf`

All artifacts compiled or knitted successfully. Two critical bugs were found and fixed in the assignment.

---

1. Slides

Section Overview

The slide deck (`panel1_body.tex`) contains **28 content frames** organized into **7 sections** plus a wrap-up, with 2 discussion frames added after pedagogy review:

Section	Key Content
1. What is panel data?	Definition, notation, panel structure table
2. Why panel data?	Motivating example (cross-sectional scatter), spurious correlation, OVB intuition
3. The within estimator	Demeaning derivation, within/between variation, FE = LSDV equivalence
4. Interpretation and limits	What FE identifies, what it cannot control for, time-invariant variables dropping out
5. Two-way fixed effects	Adding time FEs, absorbing common shocks, `feols()` syntax
6. Random effects	GLS vs FE, RE assumption, Hausman test
Wrap-up	Summary table, next session preview

Compilation Status

- **Presentation PDF (`panel1.pdf`)**: 132 pages – PASS (0 errors, 2 hbox warnings < 2.15pt)
- **Notes PDF (`panel1_notes.pdf`)**: 18 pages – PASS (0 errors)
- **Overfull hbox > 10pt**: 0 (threshold: 3 or fewer allowed)

Pedagogy Review

One critical issue identified and fixed:

- **Zero Socratic questions** on any slide
- **Fix**: Added two discussion frames:
  1. After cross-sectional scatter: *"Does unemployment cause lower approval? What else might explain this pattern?"*
  2. Before TWFE section: *"FE controls for everything time-invariant. But what if something happens in 2008 that affects all states simultaneously?"*

Visual Audit Fixes

Two vbox overflow issues fixed:

Frame	Overflow Before Fix	Fix Applied
"Panel data: the basic structure"	+7.83pt	Reduced tikz row `minimum height` from 0.7cm to 0.6cm
"The within (demeaning) estimator"	+33.47pt (-> +11.47pt after first attempt)	Restructured to single displayed equation; removed intermediate mean derivation step

One hbox overflow also fixed:

Frame	Overflow	Fix Applied
"Motivating example" tikzpicture	+10.67pt	Repositioned axis labels within frame boundary (x-axis shortened to 9.5, labels moved to x=7.5/5.8)

Images Needed

No external images required. All diagrams are built with TikZ (scatter plot, equations, panel structure table).

---

2. Assignment

File: `assign5_panel1.tex` -> 7-page PDF – PASS

Structure

Part	Section	Description
Part 1 (in-class)	1.1	Setup and data exploration
	1.2	Pooled OLS
	1.3	Entity fixed effects (`fixest`)
	1.4	Two-way fixed effects
Part 2 (take-home)	2.1	Data exploration
	2.2	Pooled OLS baseline
	2.3	Fixed effects models
	2.4	Random effects and Hausman test (`plm`)

Datasets

- **Presidential approval**: `presidential_approval.dta` – 50 U.S. states x 73 years (unbalanced); variables: `State`, `StCode`, `Year`, `PresApprov`, `UnemPct`, `South`
- **Teaching evaluations**: `teaching_evals.dta` – 48 instructors, 254 courses; variables: `Eval`, `Apct`, `Enrollment`, `Required`, `InstrID`, `CourseID`, `Year`
- Both available at `github.com/franvillamil/AQM2`

Checker Results and Fixes

Two critical bugs found and fixed:

1. **State name error**: `"New York"` -> `"NewYork"` – dataset has no spaces in state names; code would have silently returned an empty filtered dataset
2. **`pdata.frame` index**: `c("InstrID", "Year")` -> `c("InstrID", "CourseID")` – instructors teach multiple courses per year, so `InstrID + Year` is not a valid unique panel index

Items Requiring Attention

- **Deadline**: Set to `March 19, before class` – confirm this is correct

---

3. Solutions

Files Created

File	Knit Status	Size
`assign5_panel1_part1.Rmd`	PASS	242 KB PDF
`assign5_panel1_part2.Rmd`	PASS	150 KB PDF
`assign5_panel1_part1.R`	(script, not knitted)	n/a
`assign5_panel1_part2.R`	(script, not knitted)	n/a

Key Design Choices

- Data loaded from GitHub raw URLs (same as existing solution files)
- `output = "markdown"` in all `modelsummary()` calls for proper PDF table rendering
- `fig.height = 3.5` for all plots
- Part 2 explicitly explains why `InstrID + CourseID` is the correct panel index (a common student error)
- OVB direction in teaching evals is framed conditionally (explanation covers what to infer from whichever direction the coefficient shifts)

---

4. Issues Requiring Attention

Priority	Item
Low	Confirm assignment deadline: `March 19, before class`

No other issues. All critical bugs fixed; all artifacts compile/knit cleanly.

---

5. Quality Scores

Artifact	Score	Status
Slides	92/100	PASS – minor hbox warnings (-1 each x2); Socratic questions added
Assignment	100/100	PASS – two critical bugs found and fixed
Solution Part 1	100/100	PASS – knits cleanly, all questions answered
Solution Part 2	100/100	PASS – knits cleanly, all questions answered