Chapter 15 Discussion Problems

Jake Anderson

ECON 104 – UCLA

Outline

Problem 15.21

STAR Experiment (Teacher Effects)

This exercise uses data from the STAR experiment on class sizes.

Model:

$$\begin{aligned} READSCORE_{it} = {} & \beta_1 + \beta_2 SMALL + \beta_3 AIDE + \beta_4 TCHEXPER \\ & + \beta_5 TCHMASTERS + \beta_6 BOY + \beta_7 WHITE\_ASIAN \\ & + \beta_8 FREELUNCH + u_i + e_{it} \end{aligned}$$

This exercise uses data from the STAR experiment on class sizes.

Model:

$$
\begin{aligned}
READSCORE_{it} = {} & \beta_1 + \beta_2 SMALL + \beta_3 AIDE + \beta_4 TCHEXPER \\
& + \beta_5 TCHMASTERS + \beta_6 BOY + \beta_7 WHITE\_ASIAN \\
& + \beta_8 FREELUNCH + u_i + e_{it}
\end{aligned}
$$

Variables:

- $SMALL = 1$ if small class (13–17 students)
- $AIDE = 1$ if teacher aide present
- $TCHEXPER =$ teacher experience (years)
- BOY, $WHITE\_ASIAN$, $FREELUNCH =$ student characteristics

Problem 15.21
(a): Pooled OLS

(a) Estimate the regression with no fixed or random effects. Do students perform better in small classes?

Problem 15.21

(a): Pooled OLS

- (a) Estimate the regression with no fixed or random effects. Do students perform better in small classes?

Solution:

Variable	Estimate	Std. Error	t-value	
(Intercept)	437.956	1.350	324.43	***
SMALL	5.749	0.990	5.81	***
AIDE	0.809	0.953	0.85	
TCHEXPER	0.522	0.071	7.31	***
TCHMASTERS	−1.589	0.861	−1.84	.
BOY	−6.153	0.796	−7.73	***
WHITE_ASIAN	4.083	0.958	4.26	***
FREELUNCH	−14.767	0.890	−16.59	***

Problem 15.21

(a): Pooled OLS

(a) Estimate the regression with no fixed or random effects. Do students perform better in small classes?

Solution:

Variable	Estimate	Std. Error	t-value	
(Intercept)	437.956	1.350	324.43	***
SMALL	5.749	0.990	5.81	***
AIDE	0.809	0.953	0.85	
TCHEXPER	0.522	0.071	7.31	***
TCHMASTERS	−1.589	0.861	−1.84	.
BOY	−6.153	0.796	−7.73	***
WHITE_ASIAN	4.083	0.958	4.26	***
FREELUNCH	−14.767	0.890	−16.59	***

Key finding: Small classes improve reading scores by about 5.75 points on average.

Problem 15.21

(b): Cluster-Robust Standard Errors

(b) Repeat using cluster-robust standard errors (clustered by teacher). Compare 95% CI for *SMALL*.

- (b) Repeat using cluster-robust standard errors (clustered by teacher). Compare 95% CI for *SMALL*.

Solution:

	OLS SE		Cluster SE	
Variable	Coef	95% CI	Coef	95% CI
SMALL	5.749 ***	[3.81 , 7.69]	5.749 **	[1.31 , 10.19]
AIDE	0.809	[−1.06 , 2.68]	0.809	[−3.60 , 5.21]
TCHEXPER	0.522 ***	[0.38 , 0.66]	0.522 ***	[0.16 , 0.89]
FREELUNCH	−14.767 ***	[−16.51 , −13.02]	−14.767 ***	[−16.89 , −12.64]

(b) Repeat using cluster-robust standard errors (clustered by teacher). Compare 95% CI for *SMALL*.

Solution:

	OLS SE		Cluster SE	
Variable	Coef	95% CI	Coef	95% CI
SMALL	5.749 ***	[3.81 , 7.69]	5.749 **	[1.31 , 10.19]
AIDE	0.809	[−1.06 , 2.68]	0.809	[−3.60 , 5.21]
TCHEXPER	0.522 ***	[0.38 , 0.66]	0.522 ***	[0.16 , 0.89]
FREELUNCH	−14.767 ***	[−16.51 , −13.02]	−14.767 ***	[−16.89 , −12.64]

Why are clustered SEs larger?

- *SMALL* varies at the **class level** — every student in the same class has the same value
- A good teacher in a small class lifts *all* her students, but OLS counts each as independent evidence that small classes work
- Clustering corrects this: effective sample size ≈ number of **classes**, not students

Problem 15.21

(c): Random Effects

- (c) Reestimate with teacher random effects using both conventional and cluster-robust SEs.

- Reestimate with teacher random effects using both conventional and cluster-robust SEs.

Solution:

	Pooled (clust.)		RE		RE (clust.)	
Variable	Est.	SE	Est.	SE	Est.	SE
SMALL	5.749	2.267	5.625	2.215	5.625	2.263
AIDE	0.809	2.247	0.833	2.322	0.833	2.234
TCHEXPER	0.522	0.185	0.442	0.162	0.442	0.179
TCHMASTERS	−1.589	1.953	−1.705	1.974	−1.705	1.911
BOY	−6.153	0.831	−5.133	0.704	−5.133	0.762
WHITE_ASIAN	4.083	1.898	6.138	1.275	6.138	1.417
FREELUNCH	−14.767	1.083	−14.653	0.843	−14.653	0.858

Problem 15.21

(c): Random Effects

- Reestimate with teacher random effects using both conventional and cluster-robust SEs.

Solution:

	Pooled (clust.)		RE		RE (clust.)	
Variable	Est.	SE	Est.	SE	Est.	SE
SMALL	5.749	2.267	5.625	2.215	5.625	2.263
AIDE	0.809	2.247	0.833	2.322	0.833	2.234
TCHEXPER	0.522	0.185	0.442	0.162	0.442	0.179
TCHMASTERS	−1.589	1.953	−1.705	1.974	−1.705	1.911
BOY	−6.153	0.831	−5.133	0.704	−5.133	0.762
WHITE_ASIAN	4.083	1.898	6.138	1.275	6.138	1.417
FREELUNCH	−14.767	1.083	−14.653	0.843	−14.653	0.858

Key point: RE estimates are very similar to pooled OLS with clustered SEs. RE assumes $\text{Cov}(u_i, X) = 0$ — teacher effects are uncorrelated with regressors.

Problem 15.21

(d): Mundlak Test

(d) Are any variables correlated with teacher effects? Conduct the Mundlak test using teacher-level averages.

Problem 15.21

(d): Mundlak Test

(d) Are any variables correlated with teacher effects? Conduct the Mundlak test using teacher-level averages.

Mundlak test: Tests whether RE assumption holds by checking if group means predict residuals. Add teacher-level means $\overline{BOY}_i$, $\overline{WHITE\_ASIAN}_i$, $\overline{FREELUNCH}_i$ to the RE model.

Problem 15.21

(d): Mundlak Test

(d) Are any variables correlated with teacher effects? Conduct the Mundlak test using teacher-level averages.

Mundlak test: Tests whether RE assumption holds by checking if group means predict residuals. Add teacher-level means $\overline{BOY}_i$, $\overline{WHITE\_ASIAN}_i$, $\overline{FREELUNCH}_i$ to the RE model.

Test: Are the coefficients on these means jointly significant?

- Test statistic: 7.209
- Critical value: $\chi^2_{0.95,3} = 7.815$

Problem 15.21

(d): Mundlak Test

(d) Are any variables correlated with teacher effects? Conduct the Mundlak test using teacher-level averages.

Mundlak test: Tests whether RE assumption holds by checking if group means predict residuals. Add teacher-level means $\overline{BOY}_i$, $\overline{WHITE\_ASIAN}_i$, $\overline{FREELUNCH}_i$ to the RE model.

Test: Are the coefficients on these means jointly significant?

- Test statistic: 7.209
- Critical value: $\chi^2_{0.95,3} = 7.815$

$$\boxed{\text{Fail to reject } H_0\text{: RE is appropriate}}$$

Problem 15.21
(d): Mundlak Test

(d) Are any variables correlated with teacher effects? Conduct the Mundlak test using teacher-level averages.

Mundlak test: Tests whether RE assumption holds by checking if group means predict residuals. Add teacher-level means $\overline{BOY}_i$, $\overline{WHITE\_ASIAN}_i$, $\overline{FREELUNCH}_i$ to the RE model.

Test: Are the coefficients on these means jointly significant?

- Test statistic: 7.209
- Critical value: $\chi^2_{0.95,3} = 7.815$

$\boxed{\text{Fail to reject } H_0\text{: RE is appropriate}}$

Why? Teachers were randomly assigned *within* schools, so teacher effects are likely uncorrelated with student characteristics.

Problem 15.21

(e): Hausman-Taylor Estimator

- Treat *FREELUNCH* as endogenous. Use Hausman-Taylor estimation.

Problem 15.21

(e): Hausman-Taylor Estimator

(e) Treat *FREELUNCH* as endogenous. Use Hausman-Taylor estimation.

Why might *FREELUNCH* be endogenous? Free lunch status reflects family income, which likely correlates with the teacher effect u_i — e.g., schools in poorer neighborhoods may systematically attract different teachers.

Problem 15.21

(e): Hausman-Taylor Estimator

(e) Treat *FREELUNCH* as endogenous. Use Hausman-Taylor estimation.

Why might *FREELUNCH* be endogenous? Free lunch status reflects family income, which likely correlates with the teacher effect u_i — e.g., schools in poorer neighborhoods may systematically attract different teachers.

Problem: FE would remove u_i, but it also removes any time-invariant variables. If *FREELUNCH* doesn't change over time for a student, FE can't estimate its effect.

Problem 15.21

(e): Hausman-Taylor Estimator

(e) Treat *FREELUNCH* as endogenous. Use Hausman-Taylor estimation.

Why might *FREELUNCH* be endogenous? Free lunch status reflects family income, which likely correlates with the teacher effect u_i — e.g., schools in poorer neighborhoods may systematically attract different teachers.

Problem: FE would remove u_i, but it also removes any time-invariant variables. If *FREELUNCH* doesn't change over time for a student, FE can't estimate its effect.

Hausman-Taylor: Uses the time-varying exogenous variables as instruments for the endogenous ones, allowing us to estimate *FREELUNCH*'s coefficient while accounting for its correlation with u_i.

Problem 15.21

(e): Hausman-Taylor Estimator

- (e) Treat *FREELUNCH* as endogenous. Use Hausman-Taylor estimation.

Why might *FREELUNCH* be endogenous? Free lunch status reflects family income, which likely correlates with the teacher effect u_i — e.g., schools in poorer neighborhoods may systematically attract different teachers.

Problem: FE would remove u_i, but it also removes any time-invariant variables. If *FREELUNCH* doesn't change over time for a student, FE can't estimate its effect.

Hausman-Taylor: Uses the time-varying exogenous variables as instruments for the endogenous ones, allowing us to estimate *FREELUNCH*'s coefficient while accounting for its correlation with u_i.

	Pooled (clust.)		**RE**		**Hausman-Taylor**	
Variable	Est.	SE	Est.	SE	Est.	SE
SMALL	5.749	2.267	5.625	2.215	5.730	2.281
AIDE	0.809	2.247	0.833	2.322	0.851	2.195
TCHEXPER	0.522	0.185	0.442	0.162	0.512	0.164
BOY	−6.153	0.831	−5.133	0.704	−5.159	0.704
WHITE_ASIAN	4.083	1.898	6.138	1.275	6.045	1.274
FREELUNCH	−14.767	1.083	−14.653	0.843	−14.603	0.843

Problem 15.21

(e): Hausman-Taylor Estimator

- (e) Treat *FREELUNCH* as endogenous. Use Hausman-Taylor estimation.

Why might *FREELUNCH* be endogenous? Free lunch status reflects family income, which likely correlates with the teacher effect u_i — e.g., schools in poorer neighborhoods may systematically attract different teachers.

Problem: FE would remove u_i, but it also removes any time-invariant variables. If *FREELUNCH* doesn't change over time for a student, FE can't estimate its effect.

Hausman-Taylor: Uses the time-varying exogenous variables as instruments for the endogenous ones, allowing us to estimate *FREELUNCH*'s coefficient while accounting for its correlation with u_i.

	Pooled (clust.)		**RE**		**Hausman-Taylor**	
Variable	Est.	SE	Est.	SE	Est.	SE
SMALL	5.749	2.267	5.625	2.215	5.730	2.281
AIDE	0.809	2.247	0.833	2.322	0.851	2.195
TCHEXPER	0.522	0.185	0.442	0.162	0.512	0.164
BOY	−6.153	0.831	−5.133	0.704	−5.159	0.704
WHITE_ASIAN	4.083	1.898	6.138	1.275	6.045	1.274
FREELUNCH	−14.767	1.083	−14.653	0.843	−14.603	0.843

Outline

Production relationship with materials input:

$$\ln(SALES_{it}) = \beta_1 + \beta_2 \ln(CAPITAL_{it}) + \beta_3 \ln(LABOR_{it}) + \beta_4 \ln(MATERIALS_{it}) + u_i + e_{it}$$

Problem 15.25

Chinese Firms Production Function

Production relationship with materials input:

$$\ln(SALES_{it}) = \beta_1 + \beta_2 \ln(CAPITAL_{it}) + \beta_3 \ln(LABOR_{it}) + \beta_4 \ln(MATERIALS_{it}) + u_i + e_{it}$$

Question: What is the elasticity of sales with respect to each input?

Problem 15.25

(a): OLS with Different SEs

(a) Estimate by OLS. Compute conventional, HC-robust, and cluster-robust SEs. Compare 95% CIs for materials elasticity.

Problem 15.25

(a): OLS with Different SEs

(a) Estimate by OLS. Compute conventional, HC-robust, and cluster-robust SEs. Compare 95% CIs for materials elasticity.

Solution:

95% CI for $\hat{\beta}_4$ (materials elasticity):

SE Type	Lower	Upper	Width
OLS	0.7295	0.7545	0.025
HC-robust	0.7217	0.7621	0.040
Cluster-robust	0.7139	0.7699	0.056

Problem 15.25

(a): OLS with Different SEs

(a) Estimate by OLS. Compute conventional, HC-robust, and cluster-robust SEs. Compare 95% CIs for materials elasticity.

Solution:

95% CI for $\hat{\beta}_4$ (materials elasticity):

SE Type	Lower	Upper	Width
OLS	0.7295	0.7545	0.025
HC-robust	0.7217	0.7621	0.040
Cluster-robust	0.7139	0.7699	0.056

Pattern: CIs become progressively wider as we account for heteroskedasticity and within-firm correlation.

Problem 15.25

(b): Testing Returns to Scale

- (b) Test $H_0 : \beta_2 + \beta_3 + \beta_4 = 1$ using each type of SE.

Problem 15.25

(b): Testing Returns to Scale

- (b) Test $H_0 : \beta_2 + \beta_3 + \beta_4 = 1$ using each type of SE.

Solution:

SE Type	χ^2 statistic	p-value
OLS	65.29	< 0.001
HC-robust	58.03	< 0.001
Cluster-robust	29.00	< 0.001

Problem 15.25

(b): Testing Returns to Scale

(b) Test $H_0 : \beta_2 + \beta_3 + \beta_4 = 1$ using each type of SE.

Solution:

SE Type	χ^2 statistic	p-value
OLS	65.29	< 0.001
HC-robust	58.03	< 0.001
Cluster-robust	29.00	< 0.001

Conclusion: Reject H_0 at 1% level with all three SE types. Evidence of non-constant returns to scale.

Problem 15.25

(c): Testing for Serial Correlation

- (c) Test for AR(1) serial correlation in OLS residuals using 2005–2006 data.

Problem 15.25

(c): Testing for Serial Correlation

- (c) Test for AR(1) serial correlation in OLS residuals using 2005–2006 data.

Solution:
Breusch-Godfrey/Wooldridge test: $\chi^2 = 249.8$, $p < 0.001$. **Reject** H_0 of no serial correlation.

- (c) Test for AR(1) serial correlation in OLS residuals using 2005–2006 data.

Solution:
Breusch-Godfrey/Wooldridge test: $\chi^2 = 249.8$, $p < 0.001$. **Reject** H_0 of no serial correlation.

Why serial correlation?

- Unobserved firm heterogeneity (u_i) persists across years
- Pooled OLS ignores this, causing correlated errors

Problem 15.25

(c): Testing for Serial Correlation

(c) Test for AR(1) serial correlation in OLS residuals using 2005–2006 data.

Solution:

Breusch-Godfrey/Wooldridge test: $\chi^2 = 249.8$, $p < 0.001$. **Reject** H_0 of no serial correlation.

Why serial correlation?

- Unobserved firm heterogeneity (u_i) persists across years
- Pooled OLS ignores this, causing correlated errors

Implication: Need panel methods (RE or FE) to address this.

Problem 15.25

(d): Random Effects

(d) Estimate with random effects. Test for constant returns to scale. Test for presence of random effects.

(d) Estimate with random effects. Test for constant returns to scale. Test for presence of random effects.

Solution:

	OLS		**RE**	
Variable	Est.	SE	Est.	SE
ln(*CAPITAL*)	0.104	0.007	0.102	0.008
ln(*LABOR*)	0.105	0.010	0.130	0.012
ln(*MATERIALS*)	0.742	0.006	0.700	0.008
Constant	1.641	0.049	1.948	0.064

Problem 15.25

(d): Random Effects

- (d) Estimate with random effects. Test for constant returns to scale. Test for presence of random effects.

Solution:

	OLS		RE	
Variable	Est.	SE	Est.	SE
ln($CAPITAL$)	0.104	0.007	0.102	0.008
ln($LABOR$)	0.105	0.010	0.130	0.012
ln($MATERIALS$)	0.742	0.006	0.700	0.008
Constant	1.641	0.049	1.948	0.064

Returns to scale: Still reject H_0 ($\chi^2 = 68.59$, $p < 0.001$)

(d) Estimate with random effects. Test for constant returns to scale. Test for presence of random effects.

Solution:

	OLS		**RE**	
Variable	Est.	SE	Est.	SE
ln(*CAPITAL*)	0.104	0.007	0.102	0.008
ln(*LABOR*)	0.105	0.010	0.130	0.012
ln(*MATERIALS*)	0.742	0.006	0.700	0.008
Constant	1.641	0.049	1.948	0.064

Returns to scale: Still reject H_0 ($\chi^2 = 68.59$, $p < 0.001$)

LM test for random effects: Tests whether unobserved firm heterogeneity exists at all ($H_0 : \sigma_u^2 = 0$).

- $\chi^2 = 802$, $p < 0.001$. **Reject —** significant firm heterogeneity exists.

Problem 15.25

(e): Fixed Effects and Hausman Test

(e) Estimate with fixed effects. Use Hausman test to compare FE and RE.

- (e) Estimate with fixed effects. Use Hausman test to compare FE and RE.

Solution:

	OLS		**RE**		**FE**	
Variable	Est.	SE	Est.	SE	Est.	SE
$\ln(CAPITAL)$	0.104	0.007	0.102	0.008	0.052	0.013
$\ln(LABOR)$	0.105	0.010	0.130	0.012	0.106	0.021
$\ln(MATERIALS)$	0.742	0.006	0.700	0.008	0.597	0.012

Problem 15.25

(e): Fixed Effects and Hausman Test

- (e) Estimate with fixed effects. Use Hausman test to compare FE and RE.

Solution:

	OLS		**RE**		**FE**	
Variable	Est.	SE	Est.	SE	Est.	SE
ln(*CAPITAL*)	0.104	0.007	0.102	0.008	0.052	0.013
ln(*LABOR*)	0.105	0.010	0.130	0.012	0.106	0.021
ln(*MATERIALS*)	0.742	0.006	0.700	0.008	0.597	0.012

Hausman test:

- H_0: Both FE and RE consistent; RE efficient
- Test statistic: 158.39, critical value: $\chi^2_{0.95,3} = 7.815$

- Estimate with fixed effects. Use Hausman test to compare FE and RE.

Solution:

	OLS		RE		FE	
Variable	Est.	SE	Est.	SE	Est.	SE
$\ln(CAPITAL)$	0.104	0.007	0.102	0.008	0.052	0.013
$\ln(LABOR)$	0.105	0.010	0.130	0.012	0.106	0.021
$\ln(MATERIALS)$	0.742	0.006	0.700	0.008	0.597	0.012

Hausman test:

- H_0: Both FE and RE consistent; RE efficient
- Test statistic: 158.39, critical value: $\chi^2_{0.95,3} = 7.815$

$$\boxed{\text{Reject } H_0\text{: Use fixed effects}}$$

Firm effects are correlated with inputs — more productive firms use more of everything.

Problem 15.25

(f): Serial Correlation in FE Residuals

- (f) Test FE residuals for serial correlation using $\tilde{e}_{it} = \rho \tilde{e}_{i,t-1}$.

Problem 15.25

(f): Serial Correlation in FE Residuals

(f) Test FE residuals for serial correlation using $\tilde{e}_{it} = \rho \tilde{e}_{i,t-1}$.

Key fact: Under no serial correlation in e_{it}, the FE (demeaned) residuals have $\rho = -1/(T-1)$. This dataset has $T = 3$ periods, so the null value is $\rho = -1/(3-1) = -0.5$.

Problem 15.25

(f): Serial Correlation in FE Residuals

(f) Test FE residuals for serial correlation using $\tilde{e}_{it} = \rho\tilde{e}_{i,t-1}$.

Key fact: Under no serial correlation in e_{it}, the FE (demeaned) residuals have $\rho = -1/(T-1)$. This dataset has $T = 3$ periods, so the null value is $\rho = -1/(3-1) = -0.5$.

Test: $\chi^2 = 34.72$, $p < 0.001$.

Problem 15.25

(f): Serial Correlation in FE Residuals

(f) Test FE residuals for serial correlation using $\tilde{e}_{it} = \rho\tilde{e}_{i,t-1}$.

Key fact: Under no serial correlation in e_{it}, the FE (demeaned) residuals have $\rho = -1/(T-1)$. This dataset has $T = 3$ periods, so the null value is $\rho = -1/(3-1) = -0.5$.

Test: $\chi^2 = 34.72$, $p < 0.001$.

$$\boxed{\text{Reject: serial correlation present}}$$

(g): Cluster-Robust SEs for FE

(g) Re-estimate FE with cluster-robust standard errors.

- (g) Re-estimate FE with cluster-robust standard errors.

Solution:

	FE (standard)		**FE (cluster-robust)**	
Variable	Est.	SE	Est.	SE
$\ln(CAPITAL)$	0.052	0.013	0.052	0.016
$\ln(LABOR)$	0.106	0.021	0.106	0.026
$\ln(MATERIALS)$	0.597	0.012	0.597	0.029

Problem 15.25

(g): Cluster-Robust SEs for FE

(g) Re-estimate FE with cluster-robust standard errors.

Solution:

	FE (standard)		**FE (cluster-robust)**	
Variable	Est.	SE	Est.	SE
$\ln(CAPITAL)$	0.052	0.013	0.052	0.016
$\ln(LABOR)$	0.106	0.021	0.106	0.026
$\ln(MATERIALS)$	0.597	0.012	0.597	0.029

Takeaway: SEs for capital and labor grew modestly, but the SE for materials **more than doubled** $(0.012 \rightarrow 0.029)$. Cluster-robust SEs are essential when there's within-firm correlation.

Problem 15.25

Model Comparison Summary

	OLS	RE	FE	FE (clust.)
$\ln(CAPITAL)$	0.104	0.102	0.052	0.052
$\ln(LABOR)$	0.105	0.130	0.106	0.106
$\ln(MATERIALS)$	0.742	0.700	0.597	0.597

Problem 15.25

Model Comparison Summary

	OLS	RE	FE	FE (clust.)
ln(*CAPITAL*)	0.104	0.102	0.052	0.052
ln(*LABOR*)	0.105	0.130	0.106	0.106
ln(*MATERIALS*)	0.742	0.700	0.597	0.597

What changed and why?

- **OLS → RE:** Accounts for firm heterogeneity. Materials elasticity drops.
- **RE → FE:** Hausman test rejects RE. Capital elasticity *halves*, materials drops further.
- **FE → FE (clust.):** Coefficients unchanged, but SEs grow — especially for materials.

Problem 15.25

Model Comparison Summary

	OLS	RE	FE	FE (clust.)
$\ln(CAPITAL)$	0.104	0.102	0.052	0.052
$\ln(LABOR)$	0.105	0.130	0.106	0.106
$\ln(MATERIALS)$	0.742	0.700	0.597	0.597

What changed and why?

- **OLS $\to$ RE:** Accounts for firm heterogeneity. Materials elasticity drops.
- **RE $\to$ FE:** Hausman test rejects RE. Capital elasticity *halves*, materials drops further.
- **FE $\to$ FE (clust.):** Coefficients unchanged, but SEs grow — especially for materials.

Key insight: Each step corrects a different problem. FE handles correlated effects; clustering handles serial correlation in errors.

Outline

Problem 15.28

College Cost Analysis

Model for cost per student at public universities:

$$\ln(TC_{it}) = \beta_1 + \beta_2 FTESTU + \beta_3 FTGRAD + \beta_4 TT + \beta_5 GA + \beta_6 CF + \sum \gamma_t D_t + u_i + e_{it}$$

Problem 15.28

College Cost Analysis

Model for cost per student at public universities:

$$\ln(TC_{it}) = \beta_1 + \beta_2 FTESTU + \beta_3 FTGRAD + \beta_4 TT + \beta_5 GA + \beta_6 CF + \sum \gamma_t D_t + u_i + e_{it}$$

Variables:

- TC = total cost per student
- $FTESTU$ = full-time equivalent students
- TT = tenure track faculty per 100 students
- CF = contract faculty (year-to-year hires)

Problem 15.28
College Cost Analysis

Model for cost per student at public universities:

$$\ln(TC_{it}) = \beta_1 + \beta_2 FTESTU + \beta_3 FTGRAD + \beta_4 TT + \beta_5 GA + \beta_6 CF + \sum \gamma_t D_t + u_i + e_{it}$$

Variables:

- TC = total cost per student
- $FTESTU$ = full-time equivalent students
- TT = tenure track faculty per 100 students
- CF = contract faculty (year-to-year hires)

First differences: Differencing eliminates u_i: $\Delta y_{it} = y_{it} - y_{i,t-1}$, so the university fixed effect cancels.

Problem 15.28

(a): First Differences (2011)

1. Estimate first-difference model using 2011 data (no intercept).

Problem 15.28

(a): First Differences (2011)

(a) Estimate first-difference model using 2011 data (no intercept).

Solution:

$$\Delta \ln(TC_{it}) = \beta_2 \Delta FTESTU + \beta_3 \Delta FTGRAD + \ldots + \Delta e_{it}$$

(a) Estimate first-difference model using 2011 data (no intercept).

Solution:

$$\Delta \ln(TC_{it}) = \beta_2 \Delta FTESTU + \beta_3 \Delta FTGRAD + \ldots + \Delta e_{it}$$

Variable	Estimate	Std. Error	
$\Delta FTESTU$	-0.0235	0.0043	***
$\Delta FTGRAD$	0.0124	0.0204	
ΔTT	0.0081	0.0071	
ΔGA	-0.0005	0.0014	
ΔCF	0.0124	0.0052	*

Problem 15.28

(a): First Differences (2011)

(a) Estimate first-difference model using 2011 data (no intercept).

Solution:

$$\Delta \ln(TC_{it}) = \beta_2 \Delta FTESTU + \beta_3 \Delta FTGRAD + \ldots + \Delta e_{it}$$

Variable	Estimate	Std. Error	
$\Delta FTESTU$	-0.0235	0.0043	***
$\Delta FTGRAD$	0.0124	0.0204	
ΔTT	0.0081	0.0071	
ΔGA	-0.0005	0.0014	
ΔCF	0.0124	0.0052	*

Key results: $\Delta FTESTU$: -2.4% (significant) — more students = lower ATC. ΔCF: $+1.2\%$ (significant) — more contract faculty = higher ATC.

Problem 15.28

(b): Adding an Intercept

(b) Add an intercept to the first-difference model. Interpret it.

(b) Add an intercept to the first-difference model. Interpret it.

Solution:

Variable	Estimate	Std. Error	
(Intercept)	0.0214	0.0045	***
$\Delta FTESTU$	-0.0246	0.0040	***
$\Delta FTGRAD$	0.0062	0.0190	
ΔTT	0.0417	0.0097	***
ΔGA	-0.0001	0.0013	
ΔCF	0.0111	0.0048	*

Problem 15.28

(b): Adding an Intercept

(b) Add an intercept to the first-difference model. Interpret it.

Solution:

Variable	Estimate	Std. Error	
(Intercept)	0.0214	0.0045	***
$\Delta FTESTU$	-0.0246	0.0040	***
$\Delta FTGRAD$	0.0062	0.0190	
ΔTT	0.0417	0.0097	***
ΔGA	-0.0001	0.0013	
ΔCF	0.0111	0.0048	*

Intercept: $+2.1\%$ — average college increased ATC by 2.1% from 2010 to 2011, holding inputs fixed.

Problem 15.28

(b): Adding an Intercept

(b) Add an intercept to the first-difference model. Interpret it.

Solution:

Variable	Estimate	Std. Error	
(Intercept)	0.0214	0.0045	***
$\Delta FTESTU$	-0.0246	0.0040	***
$\Delta FTGRAD$	0.0062	0.0190	
ΔTT	0.0417	0.0097	***
ΔGA	-0.0001	0.0013	
ΔCF	0.0111	0.0048	*

Intercept: $+2.1\%$ — average college increased ATC by 2.1% from 2010 to 2011, holding inputs fixed.

Key change: Coefficient on tenure track faculty now significant at $+4.2\%$.

Problem 15.28

(c): Strict Exogeneity — Level Variables

- (c) Test strict exogeneity by adding 2011 levels to the FD model.

Problem 15.28

(c): Strict Exogeneity — Level Variables

(c) Test strict exogeneity by adding 2011 levels to the FD model.

What is strict exogeneity? It requires $E(e_{it} \mid X_{i1}, \ldots, X_{iT}) = 0$ for all t — errors are unrelated to regressors in *every* period, not just the current one.

(c) Test strict exogeneity by adding 2011 levels to the FD model.

What is strict exogeneity? It requires $E(e_{it} \mid X_{i1}, \ldots, X_{iT}) = 0$ for all t — errors are unrelated to regressors in *every* period, not just the current one.

F-test for joint significance of level variables:

- $F = 1.03 < F_{0.95,5,130} = 2.28$

- (c) Test strict exogeneity by adding 2011 levels to the FD model.

What is strict exogeneity? It requires $E(e_{it} \mid X_{i1}, \ldots, X_{iT}) = 0$ for all t — errors are unrelated to regressors in *every* period, not just the current one.

F-test for joint significance of level variables:

- $F = 1.03 < F_{0.95,5,130} = 2.28$

$\boxed{\text{Fail to reject: strict exogeneity holds}}$

Problem 15.28

(d): Strict Exogeneity — Lead Variables

(d) Test using forward (lead) values in FE model.

Problem 15.28

(d): Strict Exogeneity — Lead Variables

(d) Test using forward (lead) values in FE model.

Solution:
Lead values test whether future regressors predict current errors (feedback effects).

Problem 15.28

(d): Strict Exogeneity — Lead Variables

(d) Test using forward (lead) values in FE model.

Solution:

Lead values test whether future regressors predict current errors (feedback effects).

$F = 0.661$, $p = 0.653$. Lead values are **not jointly significant**.

(d): Strict Exogeneity — Lead Variables

(d) Test using forward (lead) values in FE model.

Solution:
Lead values test whether future regressors predict current errors (feedback effects).

$F = 0.661$, $p = 0.653$. Lead values are **not jointly significant**.

Conclusion: Both tests confirm strict exogeneity holds — FE and FD estimators are consistent.

Thank you!

jakeanderson@g.ucla.edu