

Panel Regression

Test name: Panel Regression

Dataset: D:\DATA ANALYSIS\H Regression Tests and Models\Panel Regression\dataset.csv

Rows in original dataset: 649

Columns in original dataset: 33

Repeated grade columns used: G1, G2, G3

Panel rows created: 1947

Number of panel units/students: 649

Number of time periods: 3

Numeric predictors in pooled model: studytime, failures, absences, age, Medu, Fedu

Categorical predictors in pooled model: school, sex, address

Pooled formula: $\text{panel_grade} \sim \text{time_index} + \text{studytime} + \text{failures} + \text{absences} + \text{age} + \text{Medu} + \text{Fedu} + \text{C(school)} + \text{C(sex)} + \text{C(address)}$

Time fixed-effects formula: $\text{panel_grade} \sim \text{C(panel_time)} + \text{studytime} + \text{failures} + \text{absences} + \text{age} + \text{Medu} + \text{Fedu} + \text{C(school)} + \text{C(sex)} + \text{C(address)}$

Entity fixed-effects formula: $\text{within_panel_grade} \sim \text{within_time_index} - 1$ using student-demeaned variables.

First-difference formula: $\text{grade_change} \sim \text{transition_time}$.

Best adjusted R-squared model among reported models: Pooled OLS

Pooled adjusted R-squared: 0.2724

Time fixed-effects adjusted R-squared: 0.2722

Entity fixed-effects adjusted R-squared: 0.0507

First-difference adjusted R-squared: 0.0028

Interpretation: panel regression uses repeated observations to distinguish within-unit time movement from between-unit differences.

Reporting guide: report panel construction, unit count, time count, model specification, fixed-effects choice, coefficients, fit statistics, and residual diagnostics.

Charts generated: 8

Model fit statistics

model	n	df_model	df_resid	r_squared	adjusted_r_sq	aic	bic	f_statistic	f_p_value	rmse	mae
Pooled OLS	1947.0000	10.0000	1936.0000	0.2761	0.2724	9163.8303	9225.1448	73.8523	0.0000	2.5313	1.9020
Time fixed ef	1947.0000	11.0000	1935.0000	0.2763	0.2722	9165.3714	9232.2600	67.1611	0.0000	2.5310	1.9015
Entity fixed e	1947.0000	1.0000	1946.0000	0.0512	0.0507	5076.5798	5082.1539	105.0609	0.0000	0.8907	0.5860
First difference	1298.0000	1.0000	1296.0000	0.0035	0.0028	4526.2520	4536.5891	4.6154	0.0319	1.3814	0.9243

Key coefficients

model	term	coefficient	standard_error	t_value	p_value	ci95_lower	ci95_upper	decision_alpha_0.05
Pooled OLS	Intercept	9.4769	0.8844	10.7156	0.0000	7.7424	11.2113	Statistically significant
Pooled OLS	C(school)[T.MS]	-1.2579	0.1361	-9.2432	0.0000	-1.5248	-0.9910	Statistically significant
Pooled OLS	C(sex)[T.M]	-0.5800	0.1218	-4.7640	0.0000	-0.8188	-0.3413	Statistically significant
Pooled OLS	C(address)[T.U]	0.2806	0.1347	2.0828	0.0374	0.0164	0.5448	Statistically significant
Pooled OLS	time_index	0.2535	0.0705	3.5974	0.0003	0.1153	0.3917	Statistically significant
Pooled OLS	studytime	0.4640	0.0733	6.3307	0.0000	0.3203	0.6077	Statistically significant
Pooled OLS	failures	-1.5047	0.1056	-14.2507	0.0000	-1.7118	-1.2977	Statistically significant
Pooled OLS	absences	-0.0671	0.0130	-5.1794	0.0000	-0.0925	-0.0417	Statistically significant
Pooled OLS	age	0.0451	0.0507	0.8895	0.3738	-0.0543	0.1445	Not statistically significant
Pooled OLS	Medu	0.2732	0.0684	3.9953	0.0001	0.1391	0.4073	Statistically significant
Pooled OLS	Fedu	0.1576	0.0692	2.2792	0.0228	0.0220	0.2933	Statistically significant
Time fixed effects	Intercept	9.7578	0.8770	11.1263	0.0000	8.0378	11.4778	Statistically significant

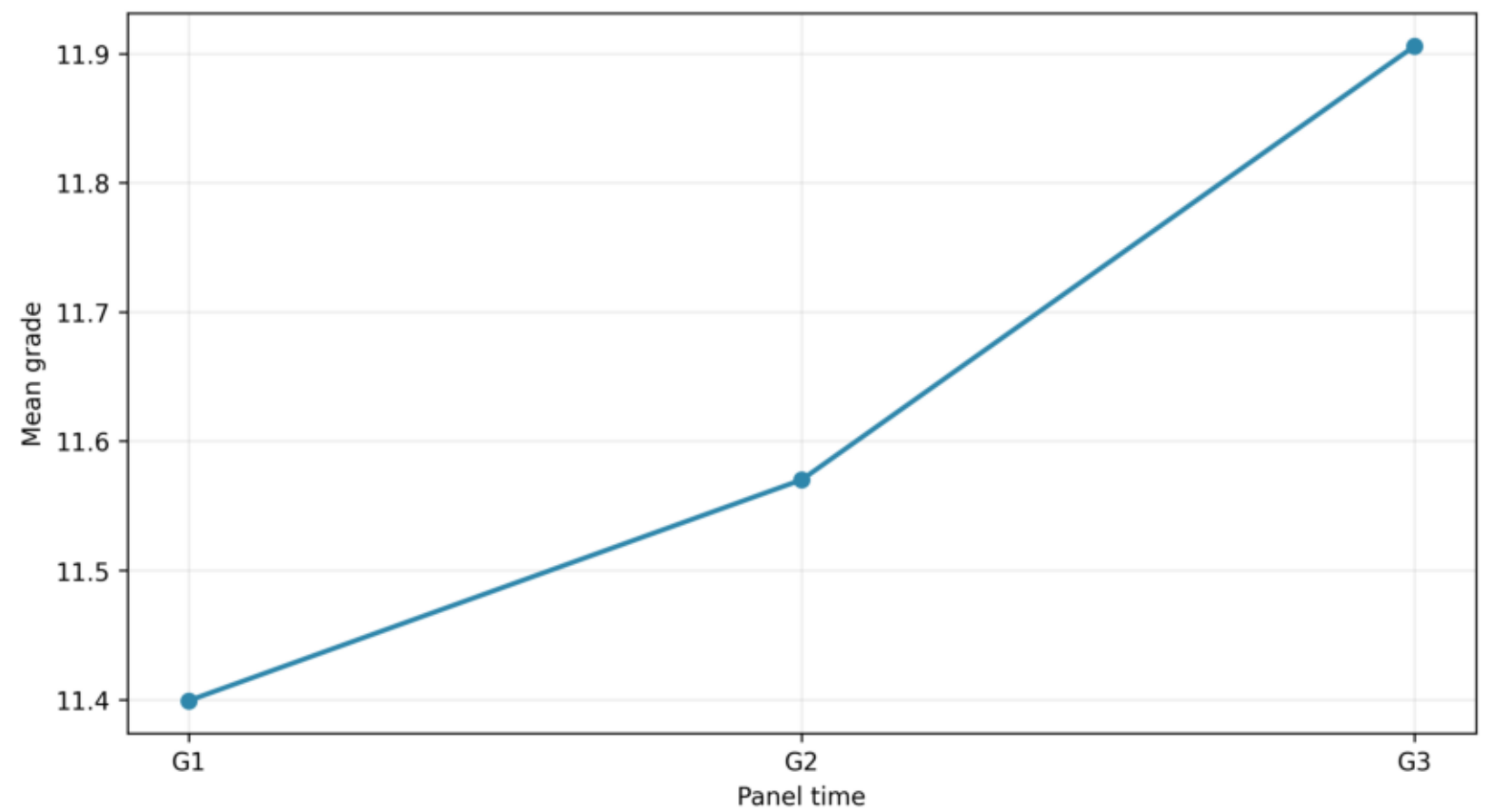
Student panel summary

student_id	student_mean_grade	within_student_sd	min_grade	max_grade	waves_observed
1.0000	7.3333	6.3509	0.0000	11.0000	3.0000
2.0000	10.3333	1.1547	9.0000	11.0000	3.0000
3.0000	12.3333	0.5774	12.0000	13.0000	3.0000
4.0000	14.0000	0.0000	14.0000	14.0000	3.0000
5.0000	12.3333	1.1547	11.0000	13.0000	3.0000
6.0000	12.3333	0.5774	12.0000	13.0000	3.0000
7.0000	12.6667	0.5774	12.0000	13.0000	3.0000
8.0000	12.0000	1.7321	10.0000	13.0000	3.0000
9.0000	16.0000	1.0000	15.0000	17.0000	3.0000
10.0000	12.3333	0.5774	12.0000	13.0000	3.0000
11.0000	14.0000	0.0000	14.0000	14.0000	3.0000
12.0000	11.6667	1.5275	10.0000	13.0000	3.0000

01_panel_mean_grade_by_time.png

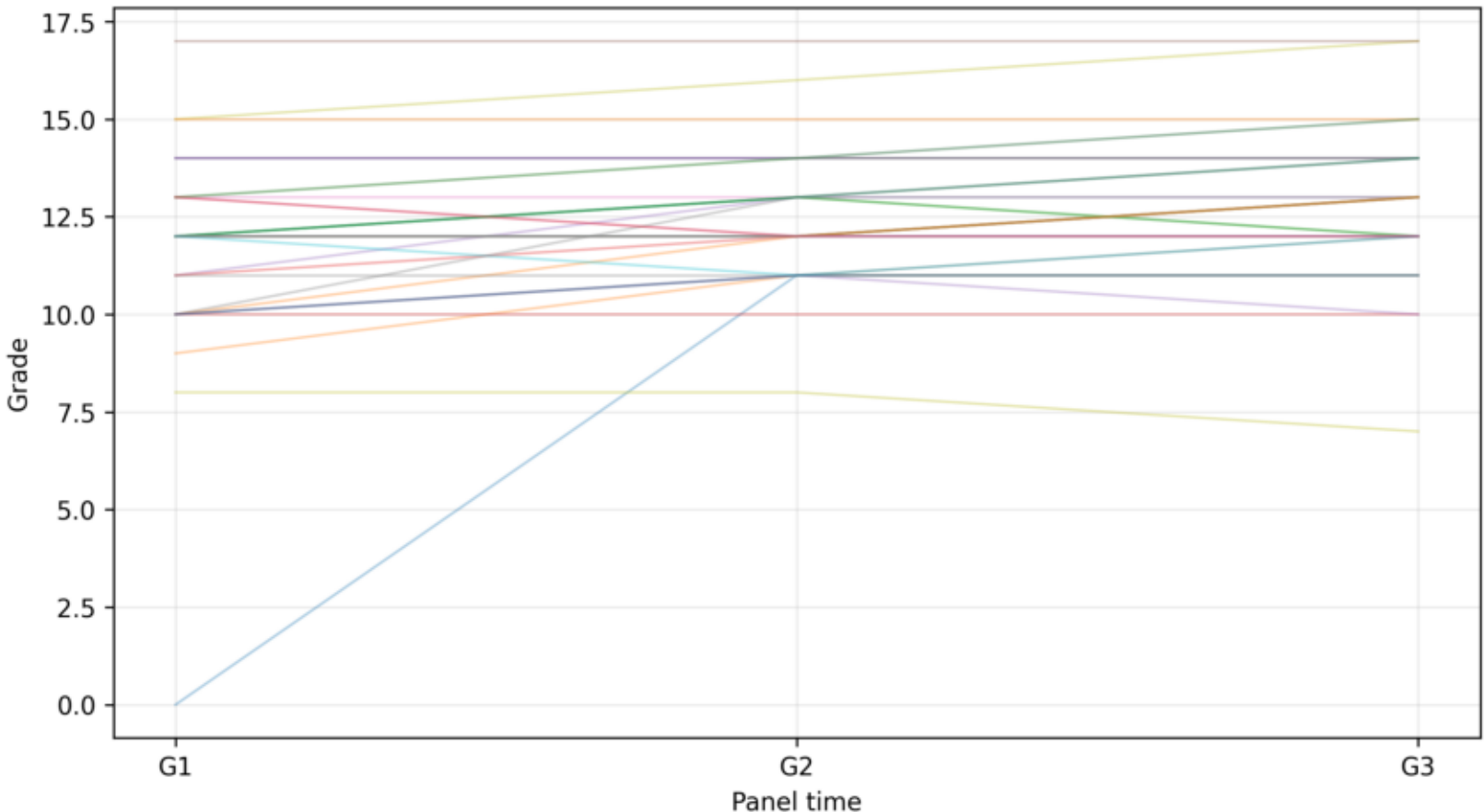
Mean grade trajectory across panel waves

G1, G2, and G3 are treated as repeated observations in a student panel.



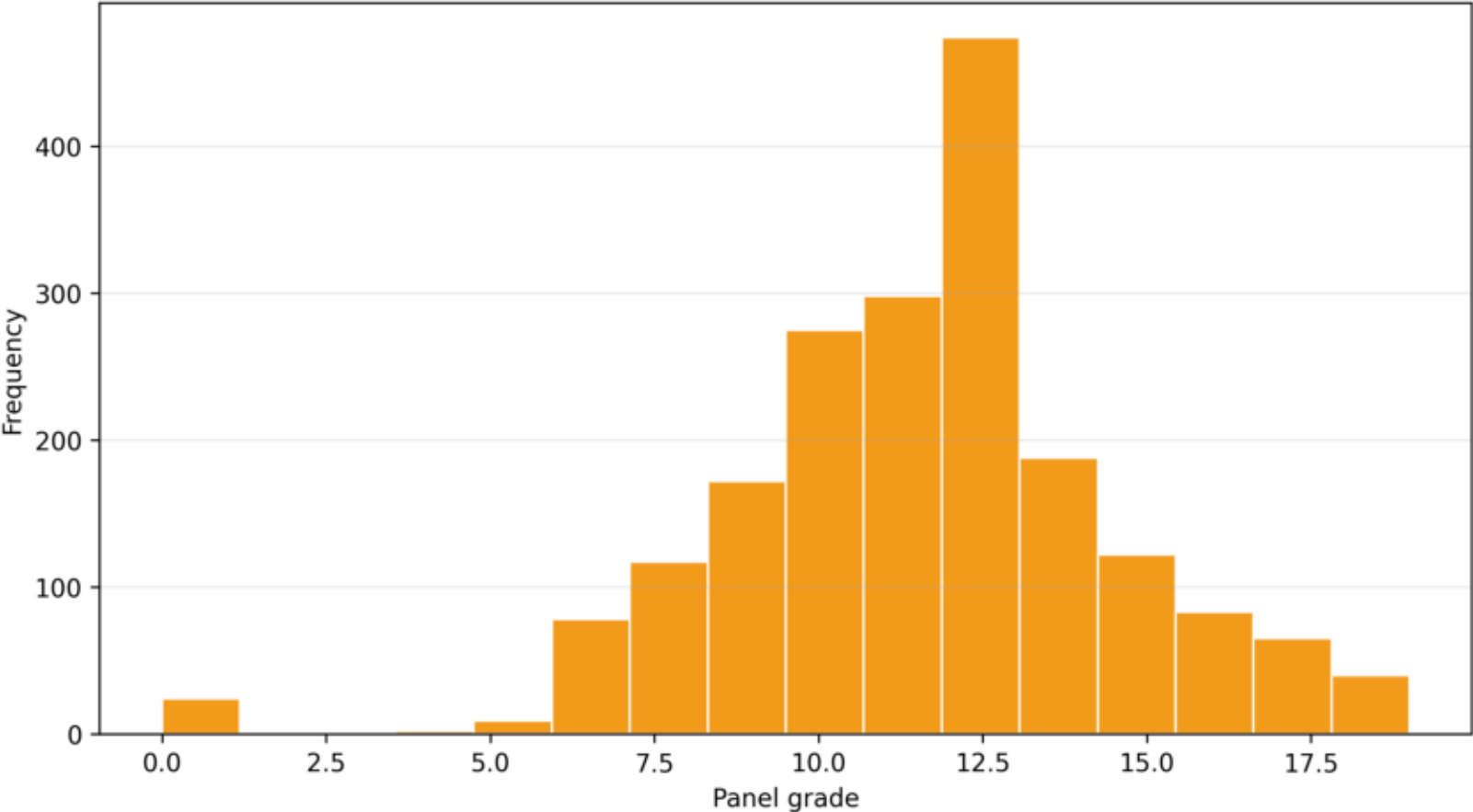
Sample student panel trajectories

Each line represents one student's repeated grade path across available waves.



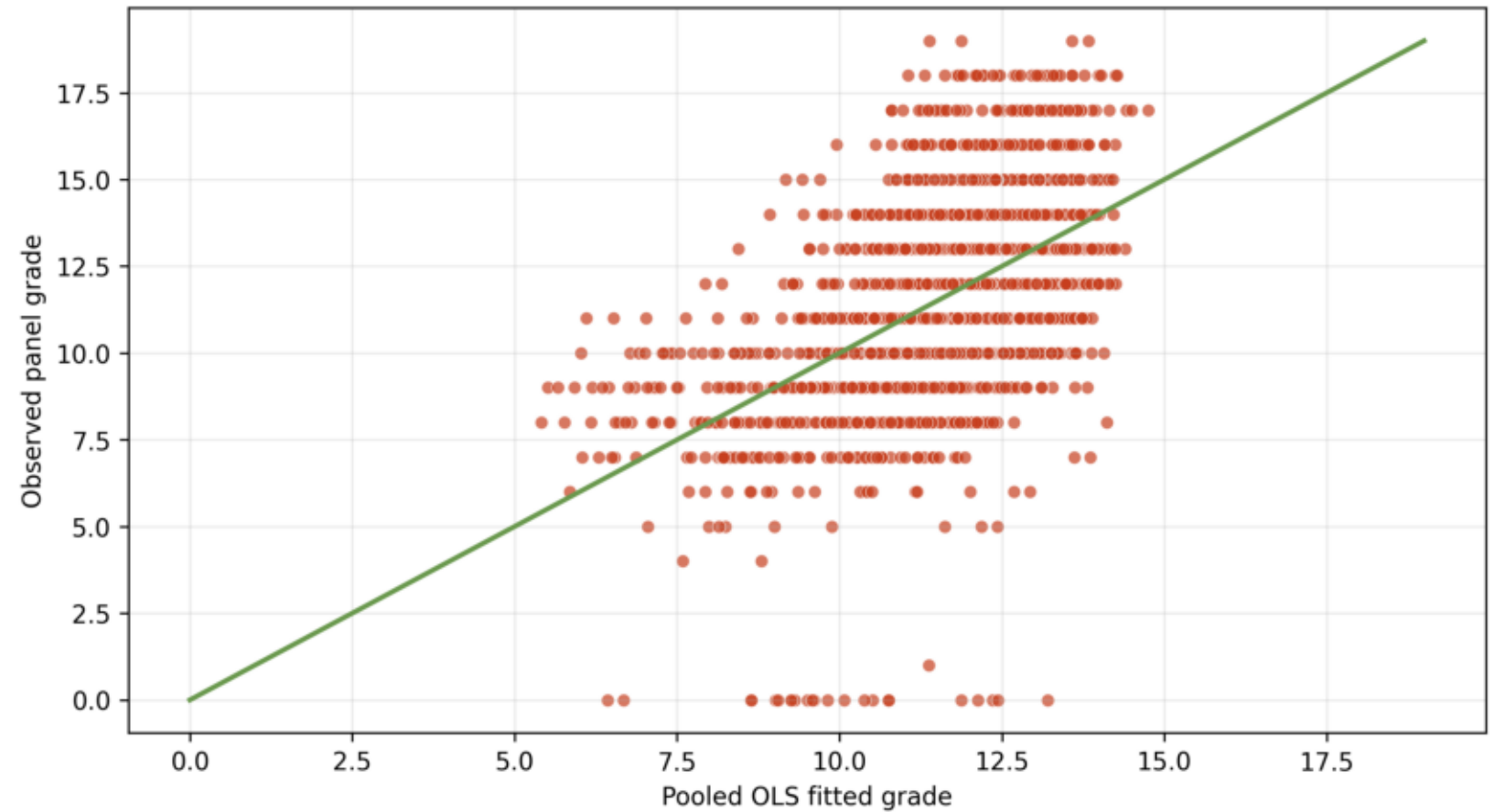
Panel outcome distribution

Distribution of the stacked repeated grade outcome.



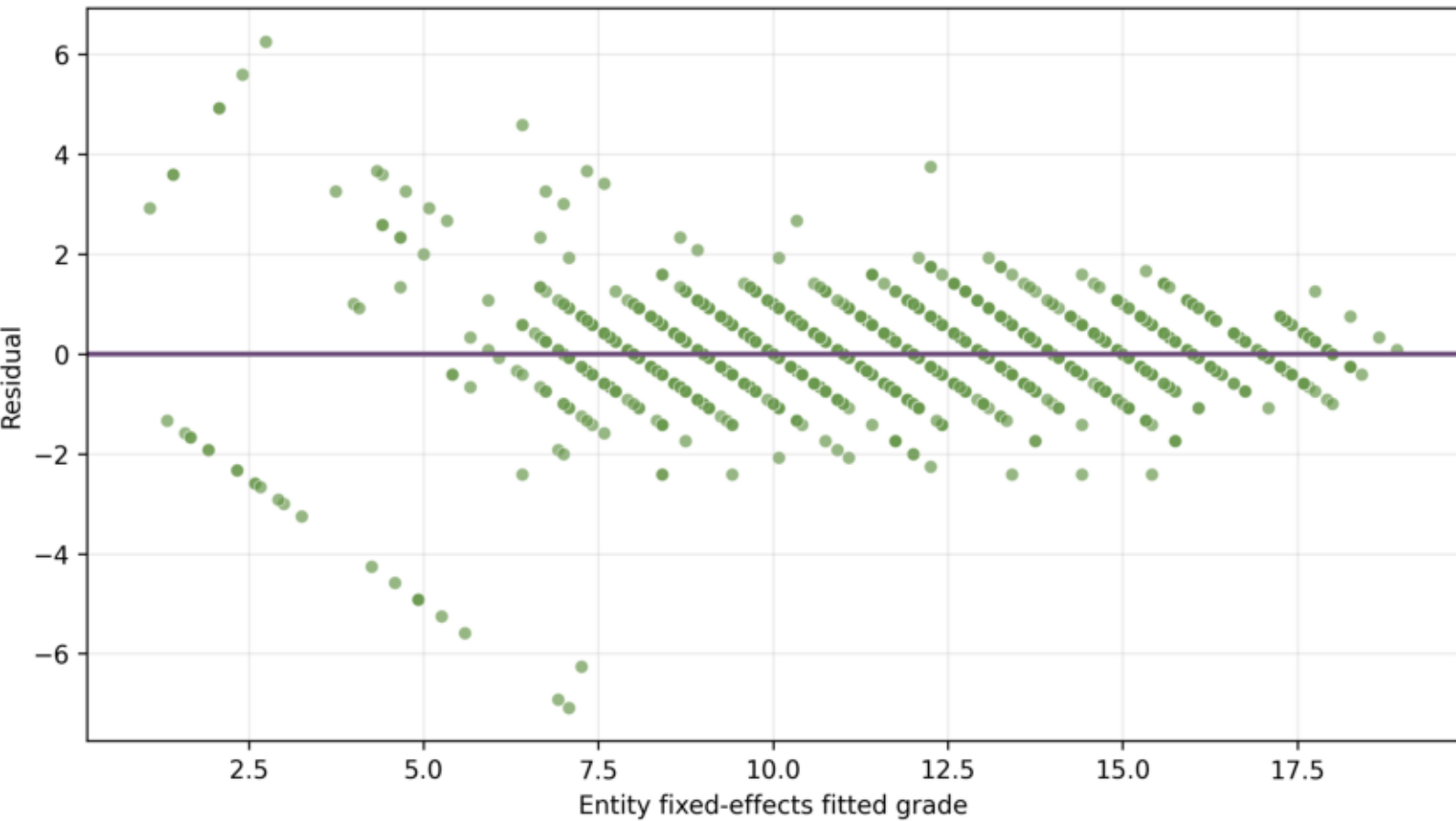
Observed vs fitted values from pooled panel model

Closer points to the diagonal indicate better prediction from the pooled panel specification.



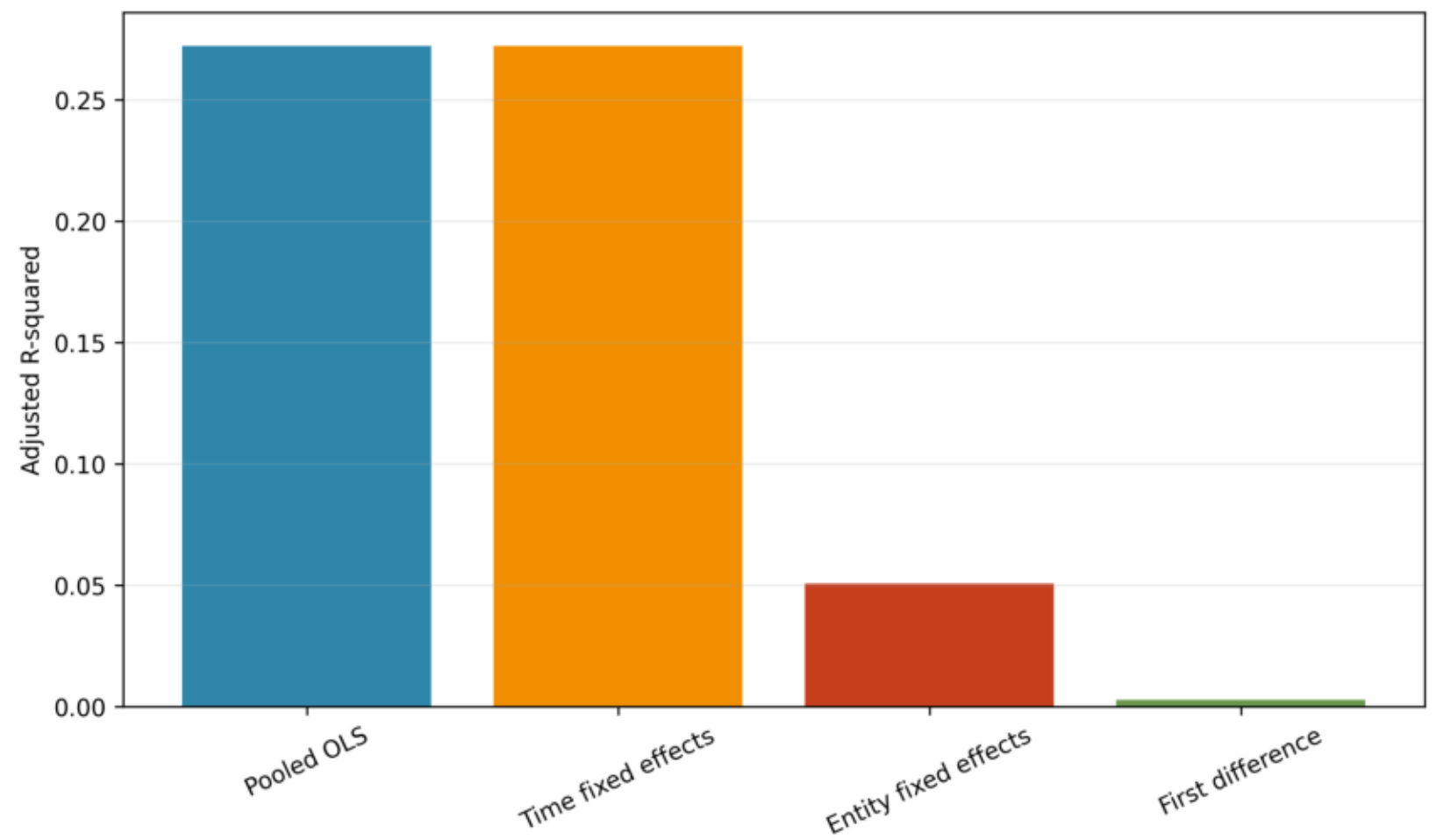
Entity fixed-effects residuals vs fitted values

Residual patterns help assess panel model fit after controlling for student fixed effects.



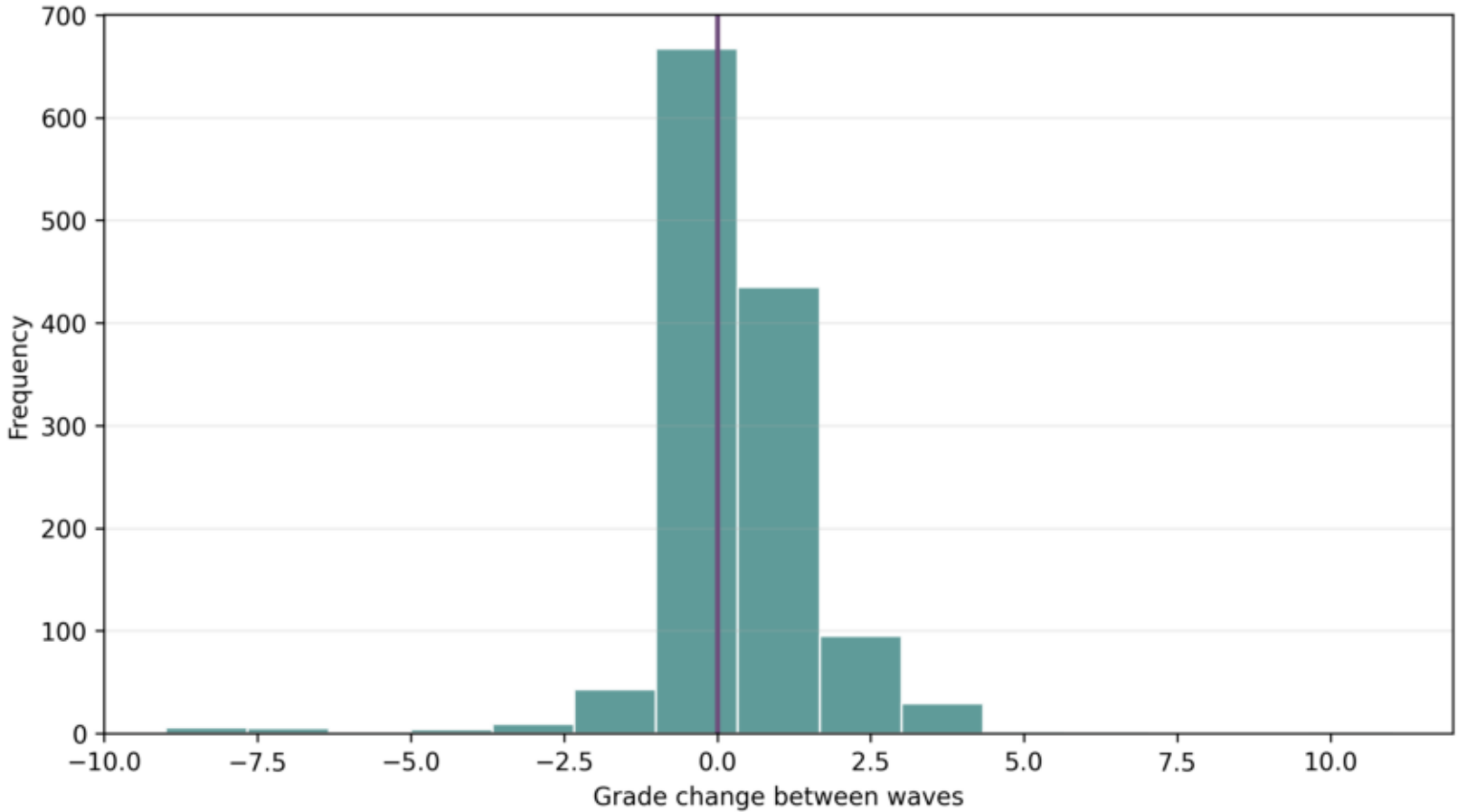
Panel model comparison

Adjusted R-squared comparison across pooled, time, entity, and first-difference specifications.



First-difference grade changes

The first-difference model focuses on within-student changes between adjacent waves.



Between-student level vs within-student variation

Panel regression separates cross-student differences from within-student movement over time.

