

Random Effects Regression

Test name: Random Effects Regression

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

Rows in original dataset: 649

Long-format rows used: 1947

Student clusters: 649

Repeated outcome columns: G1, G2, G3

Numeric predictors: studytime, failures, absences, age, Medu

Categorical predictors: school, sex, address

Full formula: grade_score ~ C(grade_period) + studytime + failures + absences + age + Medu + C(school) + C(sex) + C(address)

Random structure: random intercept for student_id.

Null-model optimizer: lbfgs

Full-model optimizer: lbfgs

Model converged: True

Random intercept variance: 5.235425

Residual variance: 1.187727

Intraclass correlation (ICC): 0.815087

Likelihood-ratio p-value: 1.67632e-58

Interpretation: fixed coefficients are average effects holding other variables constant.

Random intercepts estimate persistent student-specific deviations from the overall intercept.

ICC is the proportion of total variance attributable to between-student differences.

Decision guide: $p < 0.05$ indicates stronger evidence for a fixed-effect predictor.

Charts generated: 10

Model fit statistics

model	n	groups	log_likelihood	aic	bic	converged	r_squared	adjusted_r_squared	rmse	mae
random-intercept	1947.00000	649.00000	-3941.59881	7889.19762	7905.91975	1.00000			0.92626	0.63498
random-effects n	1947.00000	649.00000	-3791.67619	7609.35239	7681.81497	1.00000			0.90535	0.60559
oded OLS compar	1947.00000	649.00000	-4573.29492	9168.58983	9229.90433	1.00000	0.27436	0.27061	2.53439	1.90294

Likelihood-ratio test

comparison	likelihood_ratio_chi_square	df	p_value	decision_alpha_0_05
om-effects model vs null random-inter	299.84523	10.00000	0.00000	Full model improves fit

Fixed effects

model	term	coefficient	standard_error	z_value	p_value	ci95_lower	ci95_upper	decision_alpha_0_05
random-effects model	Intercept	9.98545	1.40485	7.10782	0.00000	7.23199	12.73892	Statistically significant
random-effects model	C(grade_period)[T.G2]	0.17103	0.06050	2.82701	0.00470	0.05246	0.28961	Statistically significant
random-effects model	C(grade_period)[T.G3]	0.50693	0.06050	8.37916	0.00000	0.38836	0.62551	Statistically significant
random-effects model	C(school)[T.MS]	-1.27086	0.22017	-5.77227	0.00000	-1.70237	-0.83934	Statistically significant
random-effects model	C(sex)[T.M]	-0.57878	0.19714	-2.93595	0.00333	-0.96516	-0.19240	Statistically significant
random-effects model	C(address)[T.U]	0.28196	0.21813	1.29262	0.19614	-0.14557	0.70948	Not statistically significant
random-effects model	studytime	0.46031	0.11865	3.87971	0.00010	0.22777	0.69286	Statistically significant
random-effects model	failures	-1.51888	0.17067	-8.89928	0.00000	-1.85339	-1.18436	Statistically significant
random-effects model	absences	-0.06567	0.02096	-3.13314	0.00173	-0.10676	-0.02459	Statistically significant
random-effects model	age	0.03929	0.08195	0.47943	0.63163	-0.12132	0.19990	Not statistically significant
random-effects model	Medu	0.36896	0.08735	4.22395	0.00002	0.19776	0.54016	Statistically significant

Variance components

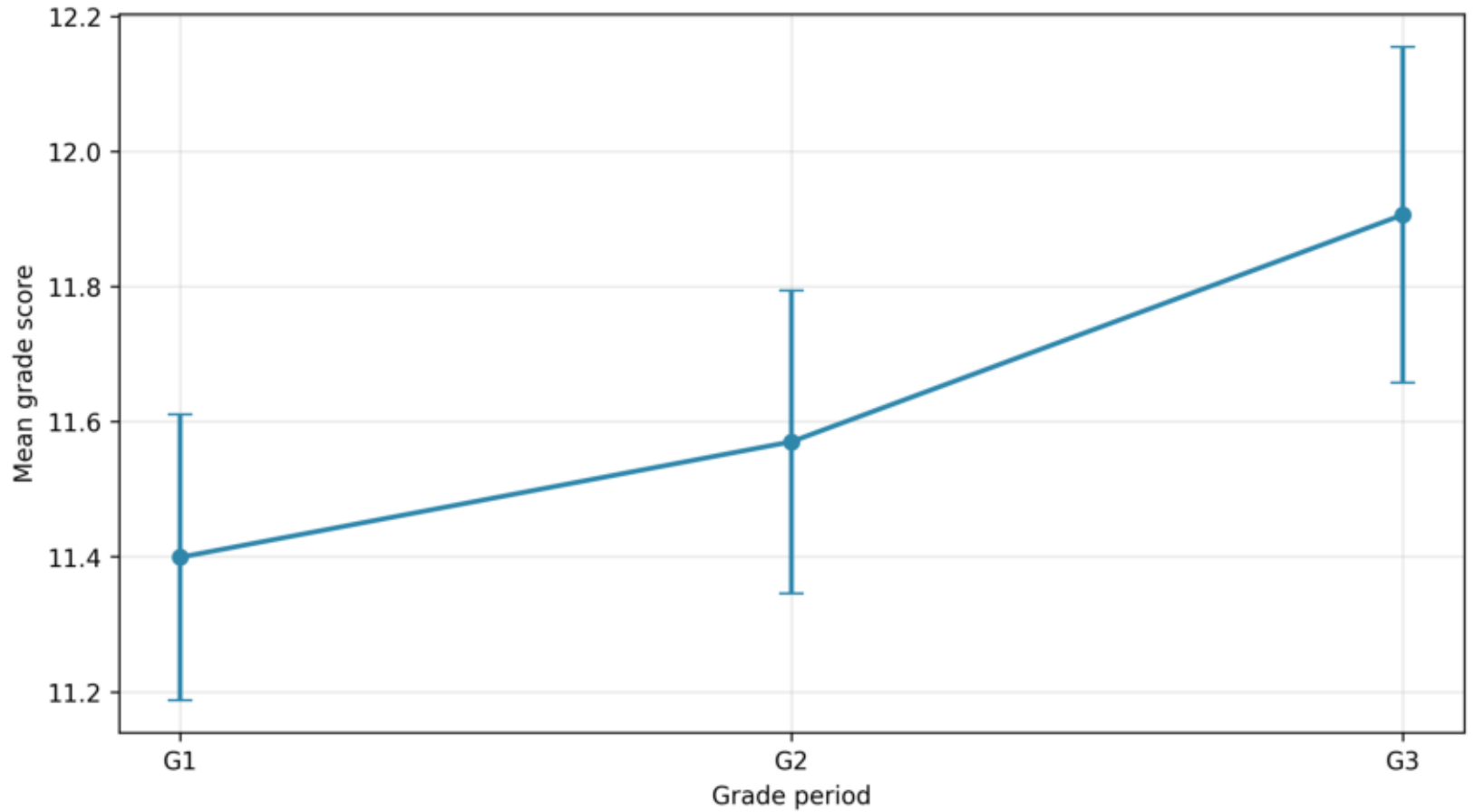
component	value
Random intercept variance	5.23542
Residual variance	1.18773
Total variance	6.42315
Intraclass correlation (ICC)	0.81509

Random intercepts

student_id	random_intercept
1.00000	-5.57913
2.00000	-1.84656
3.00000	0.33010
4.00000	0.05625
5.00000	-0.75880
6.00000	-0.19739
7.00000	0.43221
8.00000	-1.32613
9.00000	3.22470
10.00000	-0.18418
11.00000	0.60631
12.00000	-1.42700
13.00000	-0.09925
14.00000	-0.52720
15.00000	1.59028
16.00000	4.03104
17.00000	0.01048
18.00000	0.60291

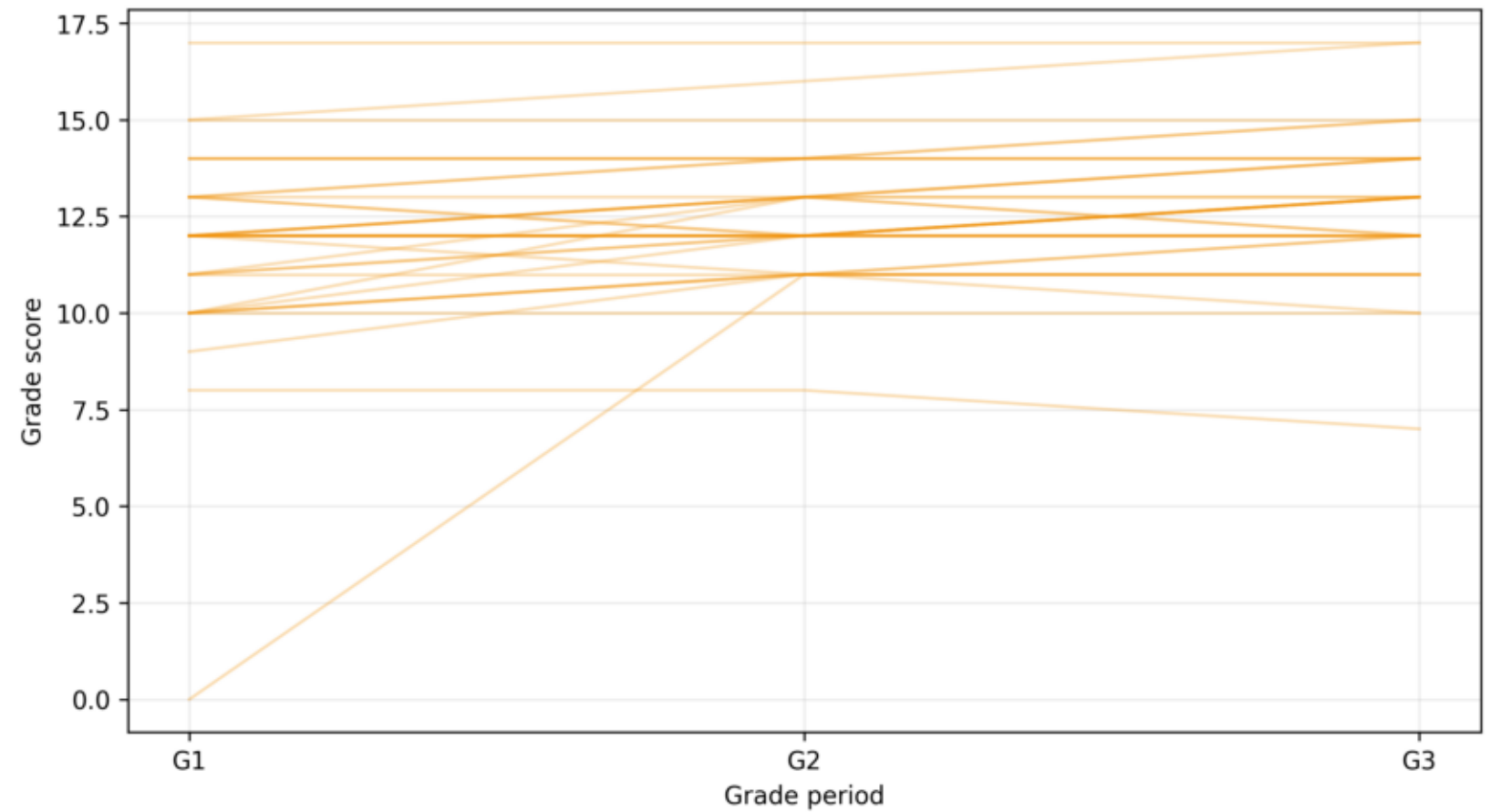
Mean grade trajectory

Error bars show approximate 95% confidence intervals



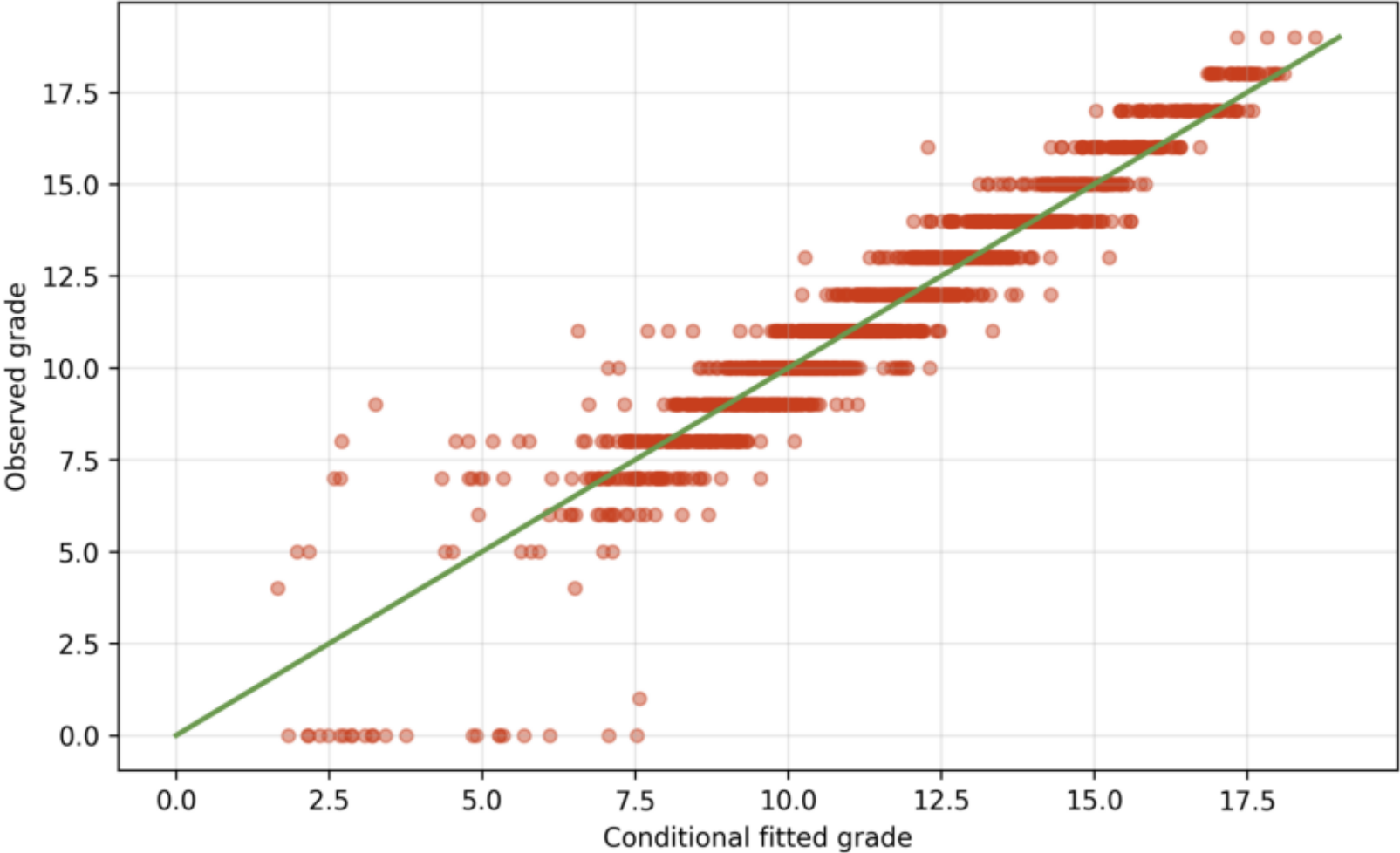
Sample student trajectories

Repeated outcomes demonstrate within-student dependence



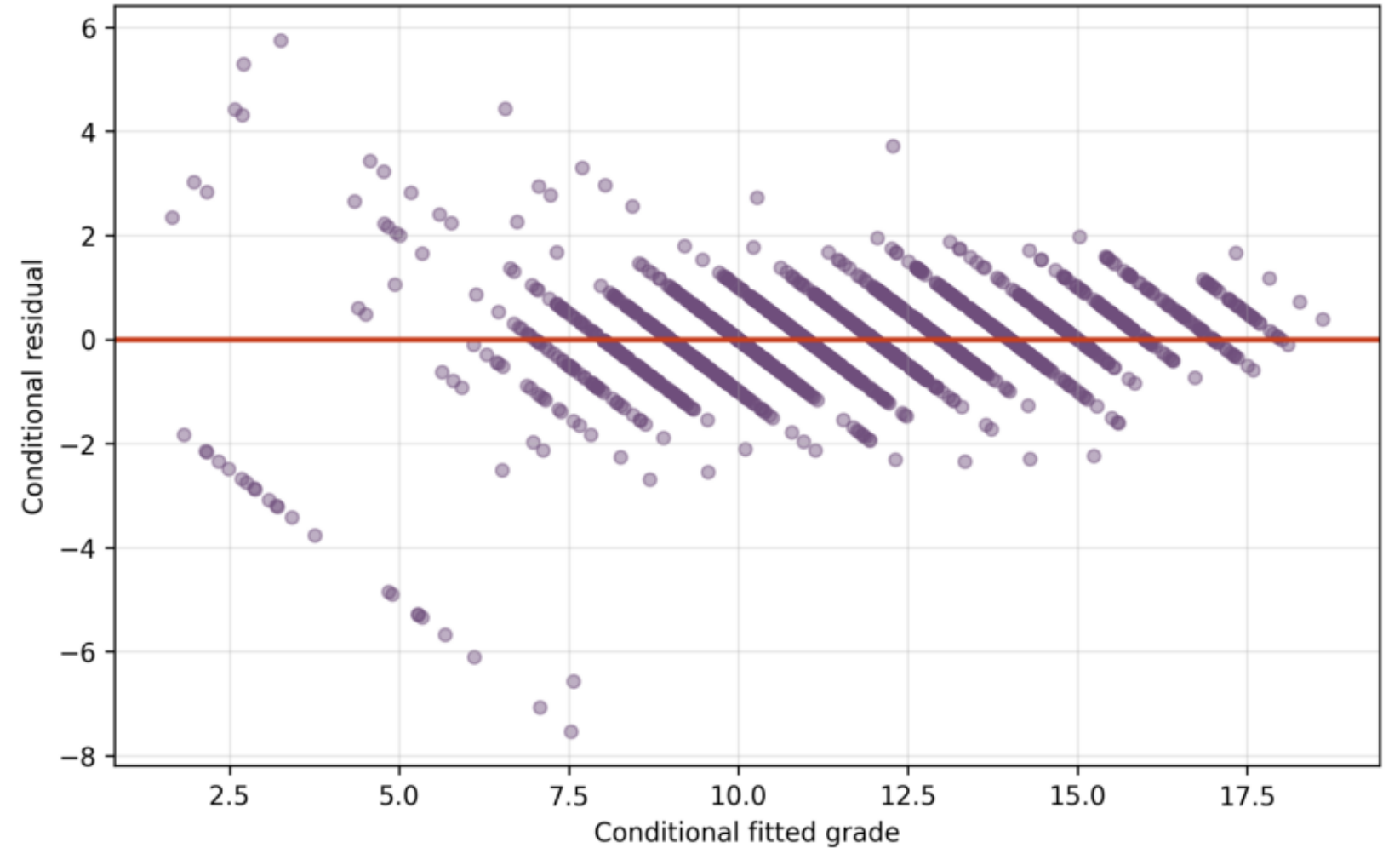
Observed versus fitted values

Points near the diagonal indicate stronger conditional predictions



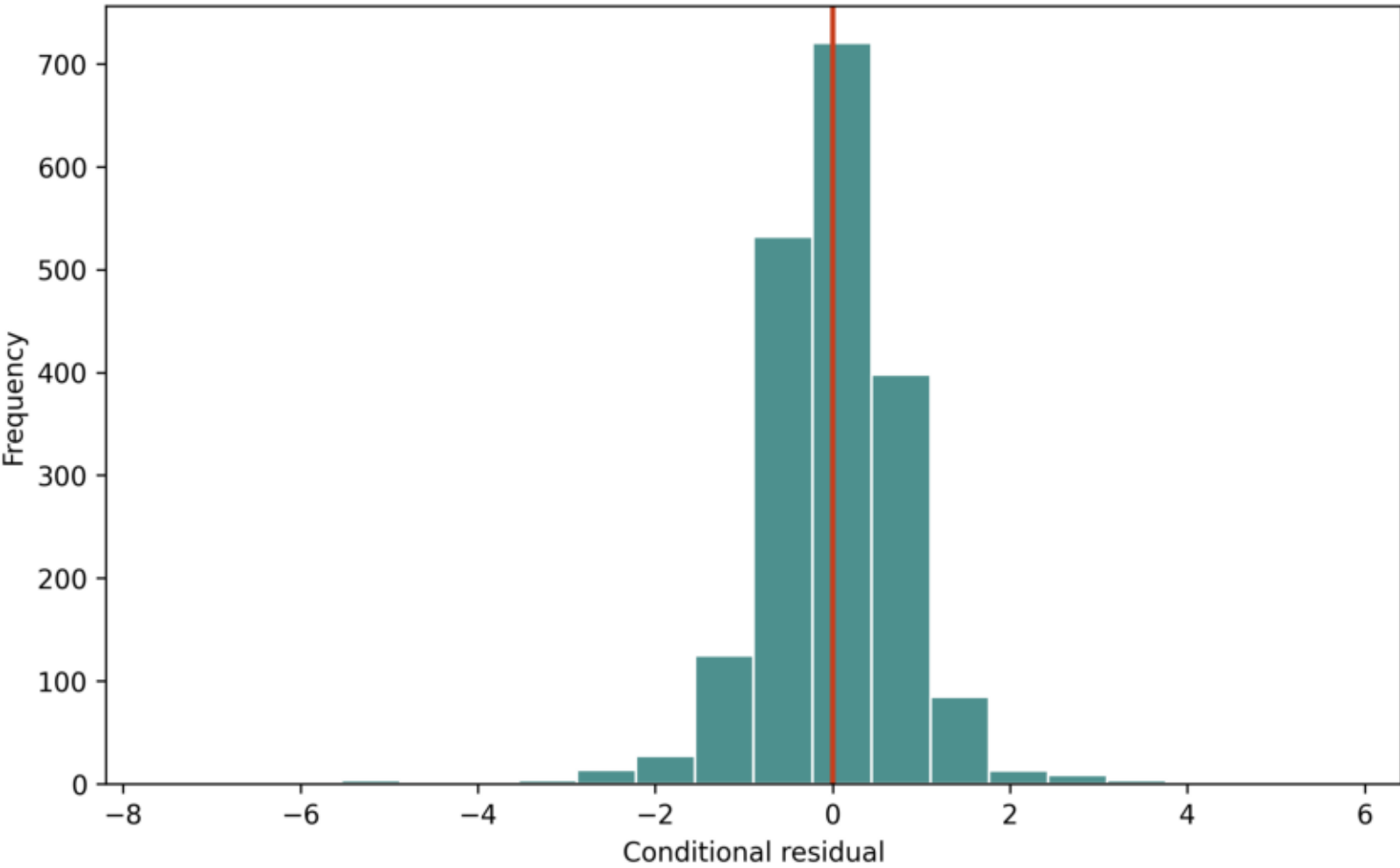
Residuals versus fitted values

Random scatter around zero supports the linear mean structure



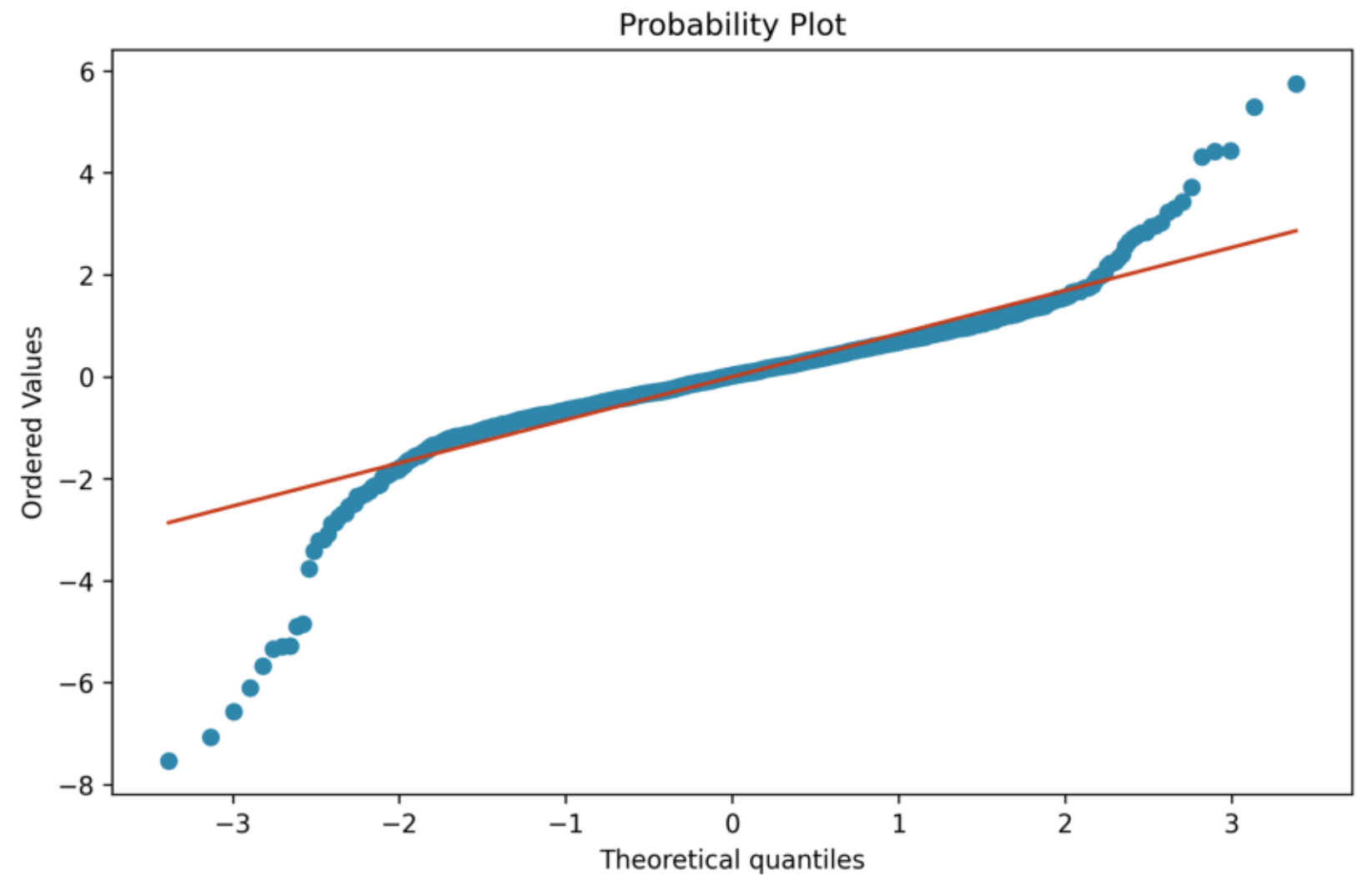
Residual distribution

A roughly symmetric distribution is desirable



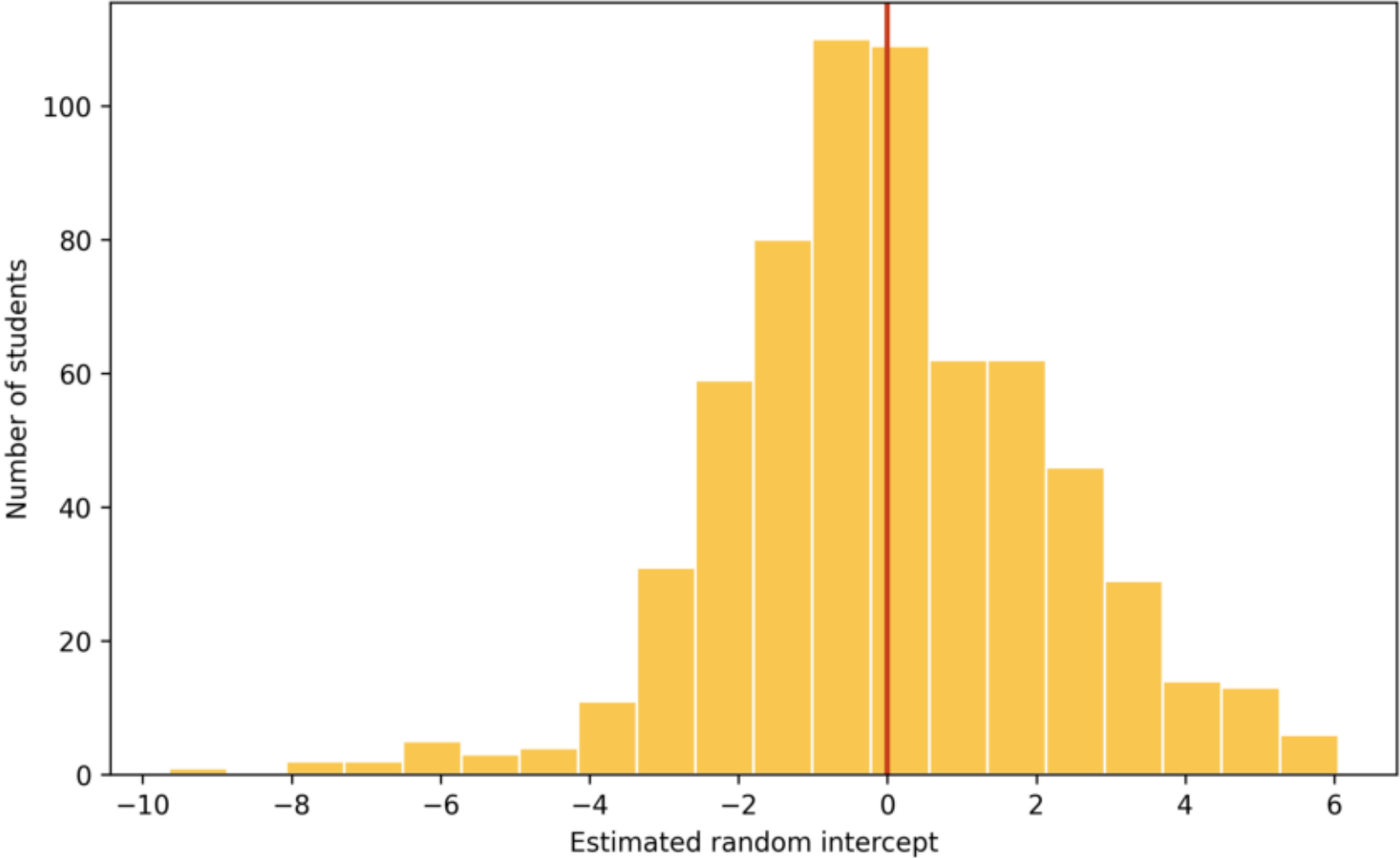
Normal Q-Q plot of residuals

Closer alignment with the reference line indicates stronger normality



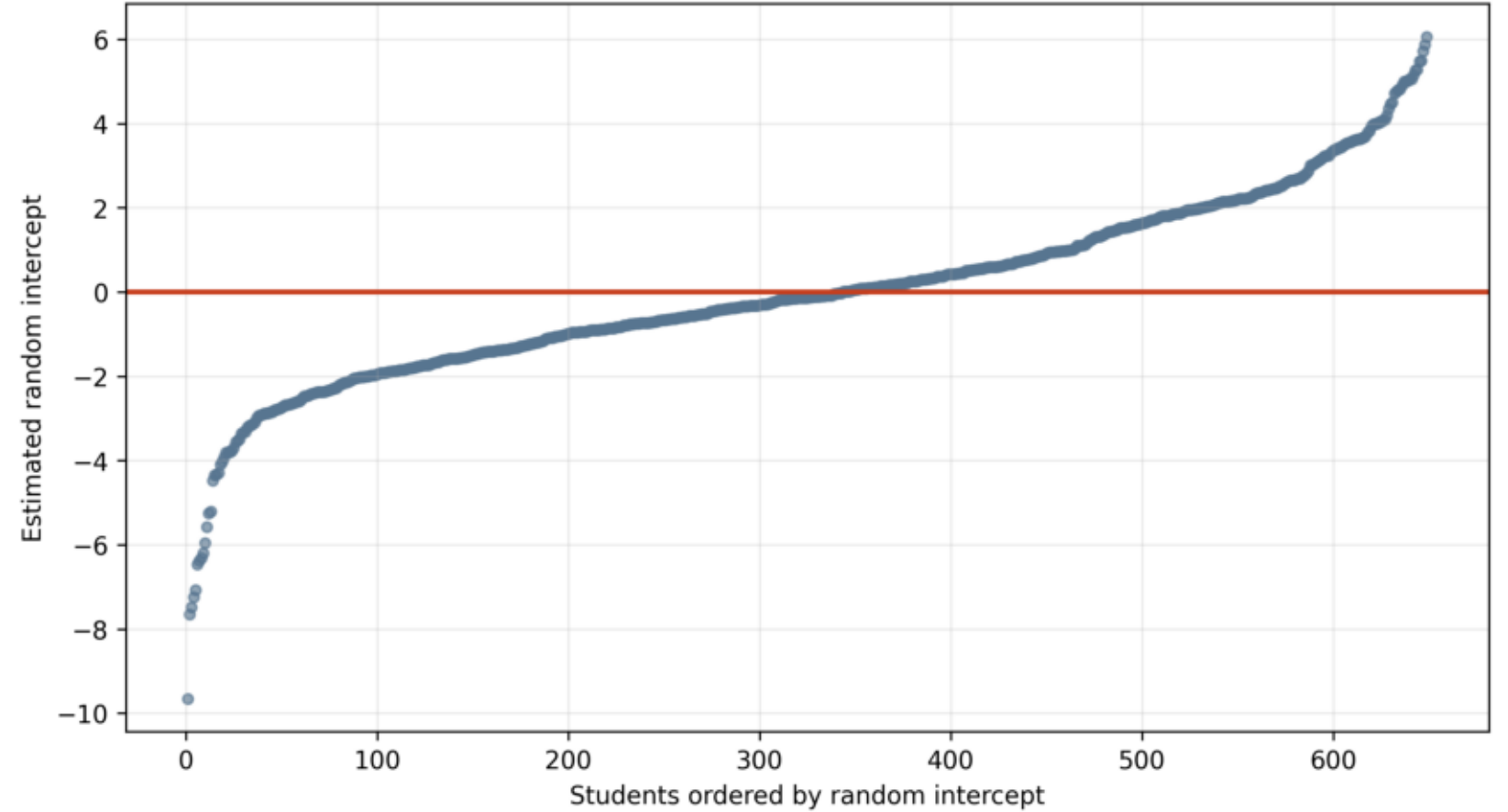
Random-intercept distribution

Positive values indicate above-average persistent student effects



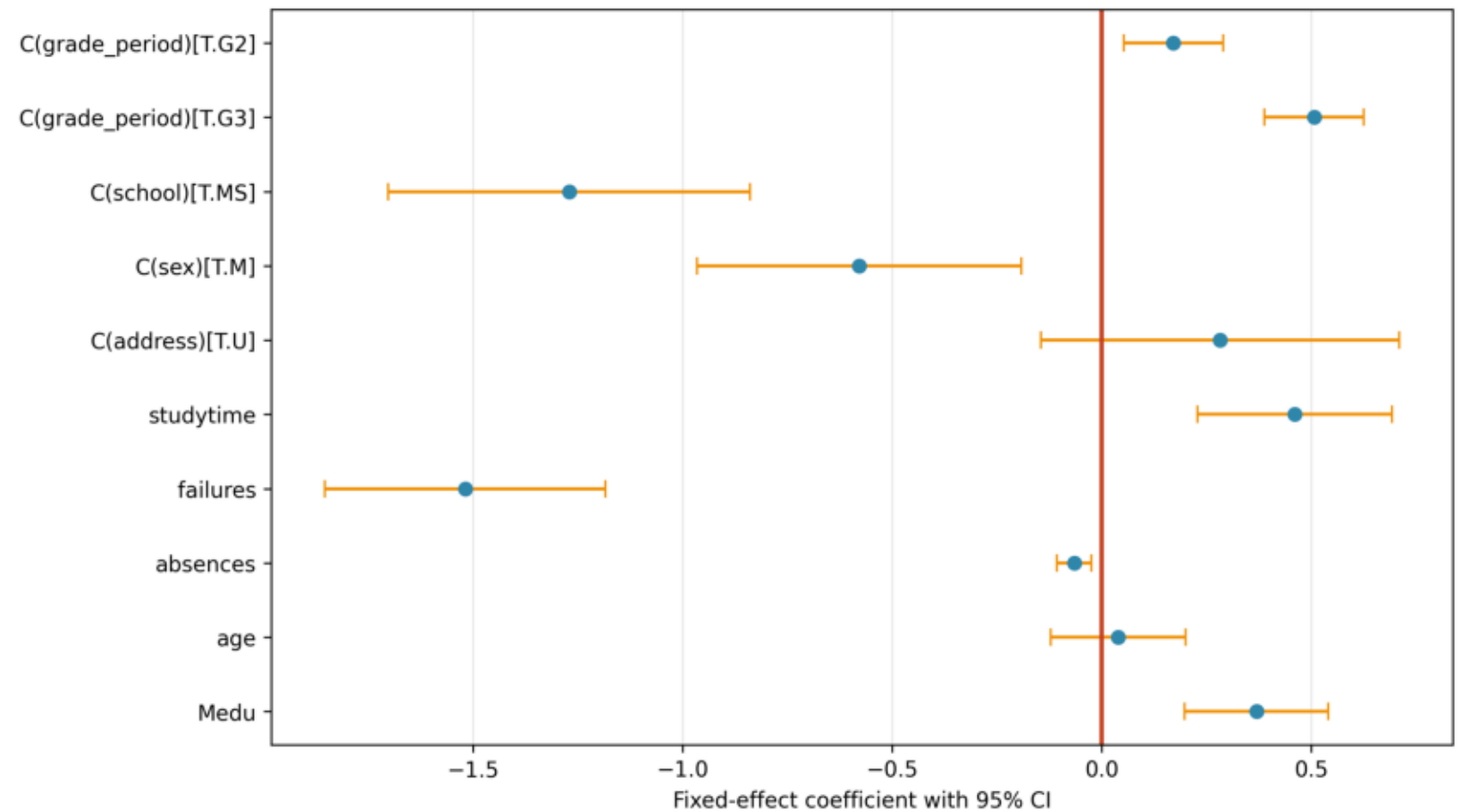
Random-intercept caterpillar

The chart displays between-student heterogeneity



Fixed-effect coefficient plot

Intervals crossing zero indicate weaker evidence at the 5% level



Variance components

Random-intercept and residual variances determine the ICC

