

Multilevel Regression

Test name: Multilevel Regression

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

Rows in original dataset: 649

Columns in original dataset: 33

Rows used in model: 649

Outcome variable: G3

Grouping variable: school

Number of groups: 2

Numeric predictors: G1, G2, studytime, failures, absences, age

Categorical predictors: sex, address, famsize

Centered predictors used: c_G1, c_G2, c_studytime, c_failures, c_absences, c_age

Full model formula: $G3 \sim c_G1 + c_G2 + c_studytime + c_failures + c_absences + c_age + C(sex) + C(address) + C(famsize)$

Null random-intercept model fitted with method=lbfgs

Full random-intercept model fitted with method=lbfgs

Null model ICC: 0.0000

Full model ICC: 0.0006

The ICC is small, so between-group clustering is limited in this dataset.

OLS comparison R-squared: 0.8518

Full multilevel log-likelihood: -1061.9294

Interpretation: fixed effects describe predictor associations, while the random intercept captures group-to-group baseline differences.

Reporting guide: state outcome, grouping variable, fixed predictors, random intercept, variance components, ICC, and key fixed effects with confidence intervals.

Charts generated: 8

Fixed effects

model	term	coefficient	standard_error	z_value	p_value	ci95_lower	ci95_upper	decision_alpha_0.05
Full random-intercept	Intercept	11.8672	0.0944	125.6656	0.0000	11.6821	12.0523	Statistically significant
Full random-intercept	C(sex)[T.M]	-0.1927	0.0868	-2.2203	0.0264	-0.3628	-0.0226	Statistically significant
Full random-intercept	C(address)[T.U]	0.1409						Not statistically significant
Full random-intercept	C(famsize)[T.LE3]	0.0565	0.1040	0.5436	0.5867	-0.1473	0.2604	Not statistically significant
Full random-intercept	c_G1	0.1375	0.0323	4.2637	0.0000	0.0743	0.2008	Statistically significant
Full random-intercept	c_G2	0.8810	0.0337	26.1708	0.0000	0.8150	0.9470	Statistically significant
Full random-intercept	c_studytime	0.0707	0.0566	1.2495	0.2115	-0.0402	0.1816	Not statistically significant
Full random-intercept	c_failures	-0.2226	0.0943	-2.3603	0.0183	-0.4074	-0.0378	Statistically significant
Full random-intercept	c_absences	0.0206	0.0054	3.7921	0.0001	0.0100	0.0313	Statistically significant
Full random-intercept	c_age	0.0193	0.0417	0.4623	0.6439	-0.0624	0.1010	Not statistically significant

Variance components and ICC

model	group_random_	residual_within_	total_variance	icc	icc_percent	log_likelihood	aic	bic	converged
Null random-int	0.0000	9.5788	9.5788	0.0000	0.0000	inf	-inf	-inf	1.0000
Full random-inte	0.0009	1.5436	1.5445	0.0006	0.0568	-1061.9294	2147.8588	2201.5640	1.0000

Group summary

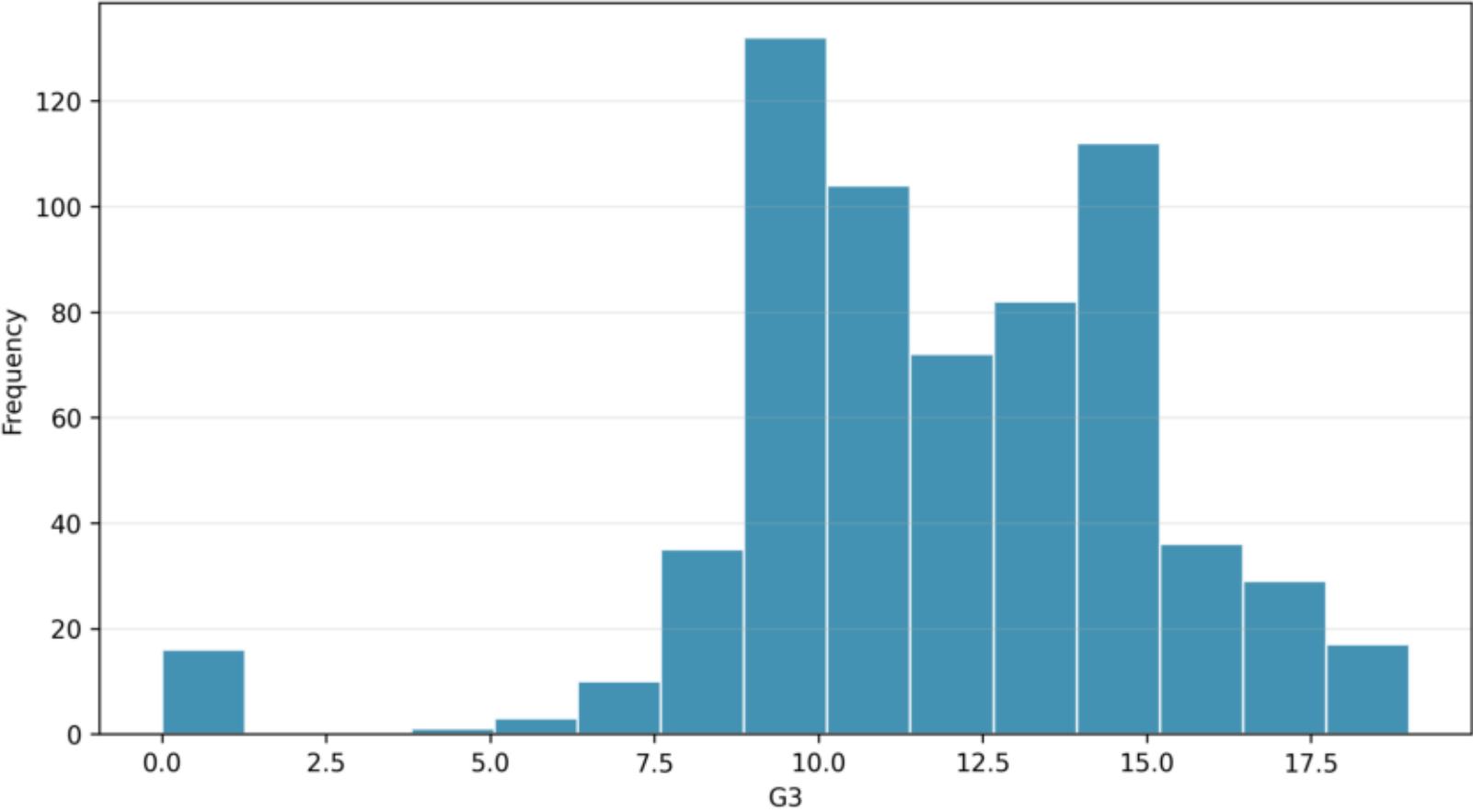
school	n	mean_outcome	sd_outcome	min_outcome	max_outcome
GP	423.0000	12.5768	2.6256	0.0000	19.0000
MS	226.0000	10.6504	3.8340	0.0000	19.0000

Random effects

school	random_intercept_effect
GP	0.0097
MS	-0.0097

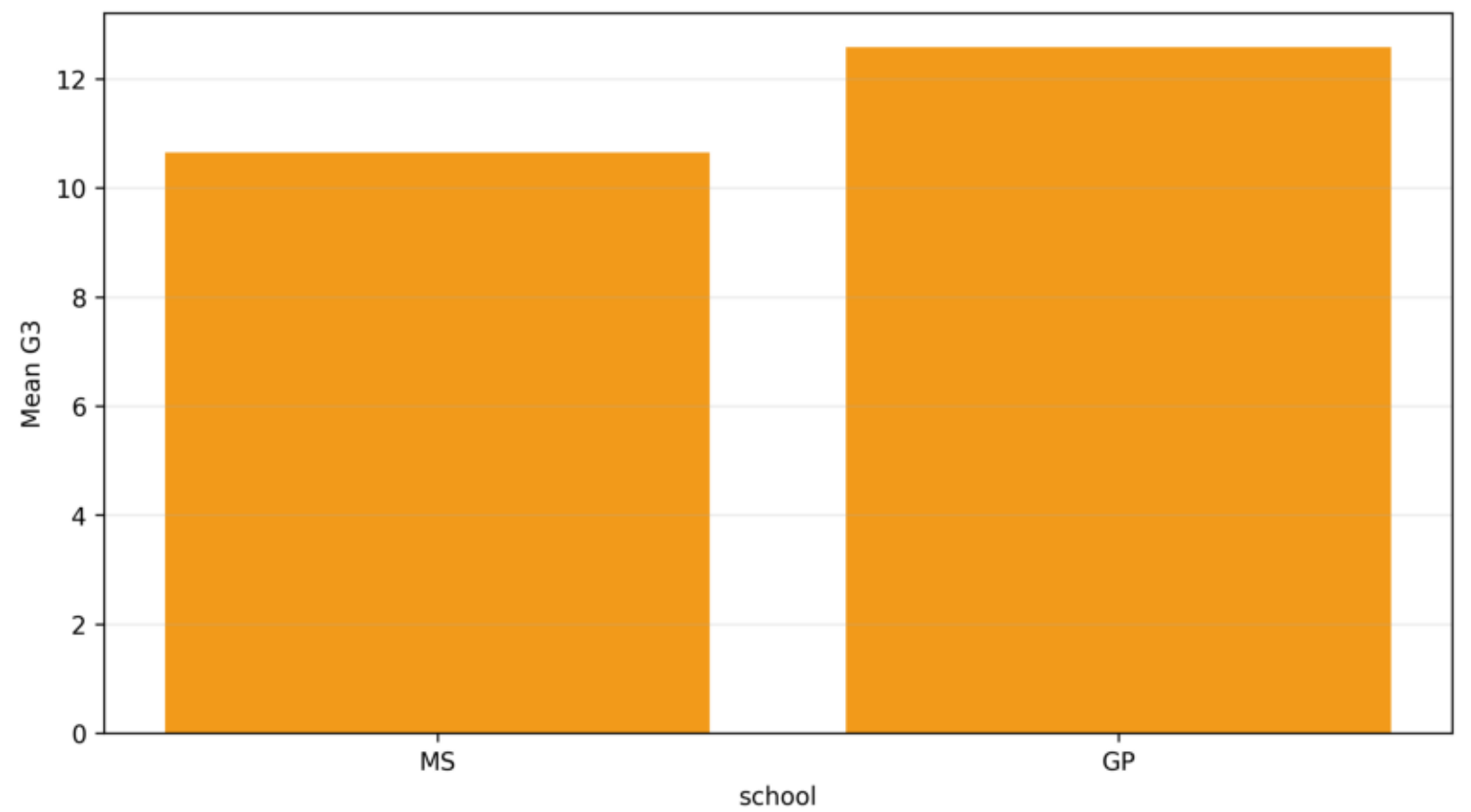
Outcome distribution

Distribution of the continuous outcome used in the multilevel model.



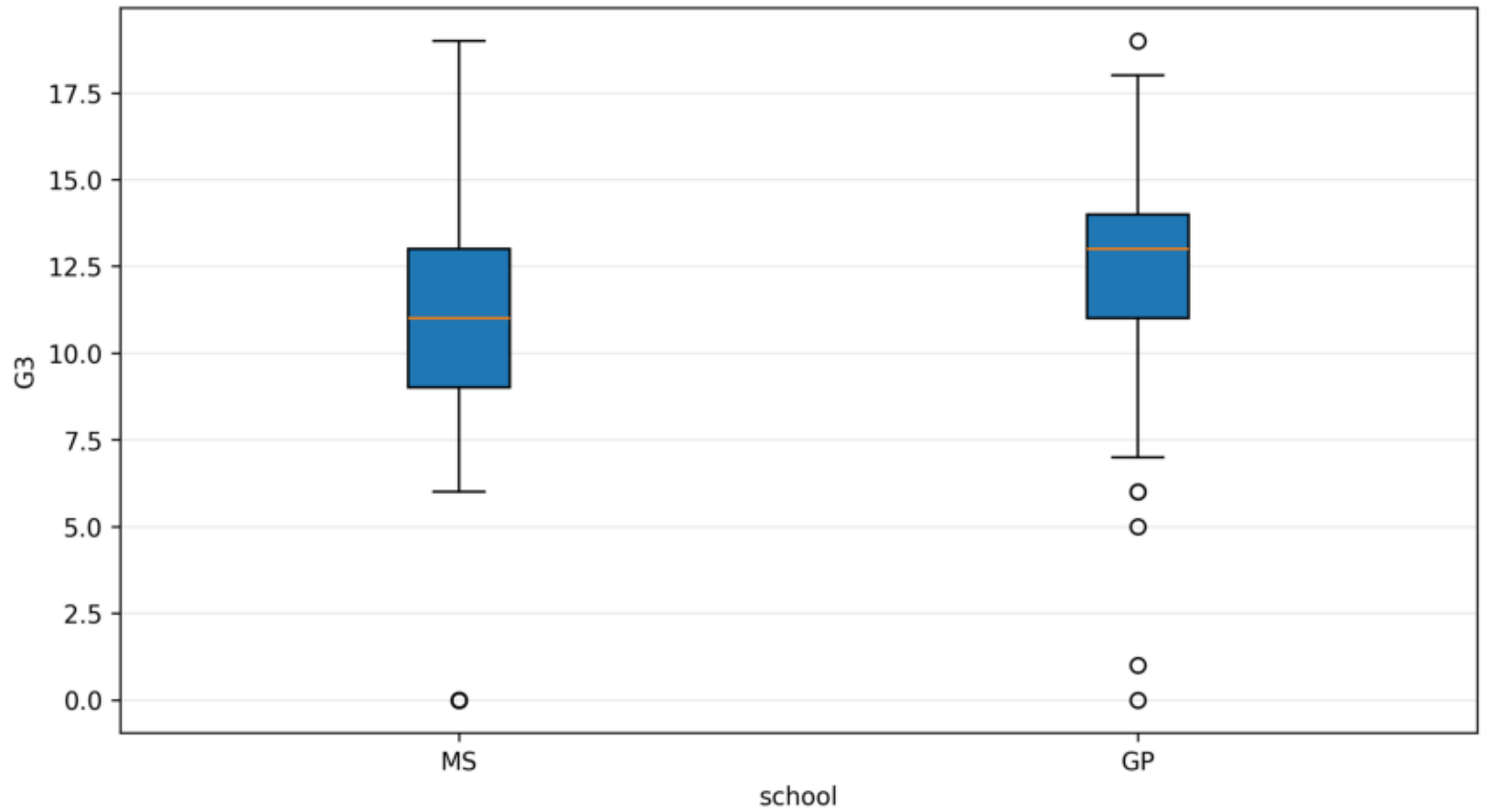
Mean G3 by school

Between-group mean differences are the reason for checking a multilevel random intercept.



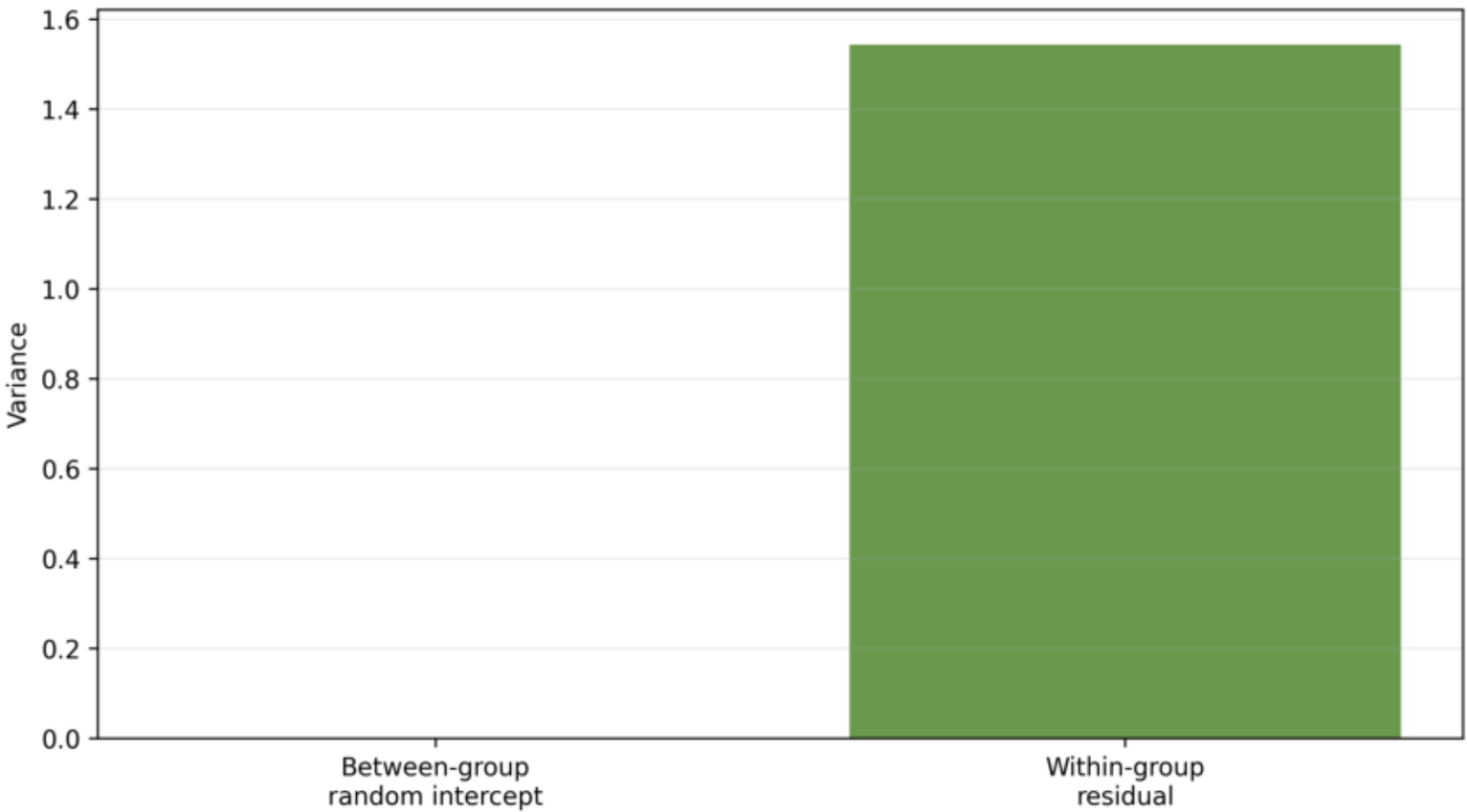
G3 spread by school

Boxplots show within-group variability and between-group differences.



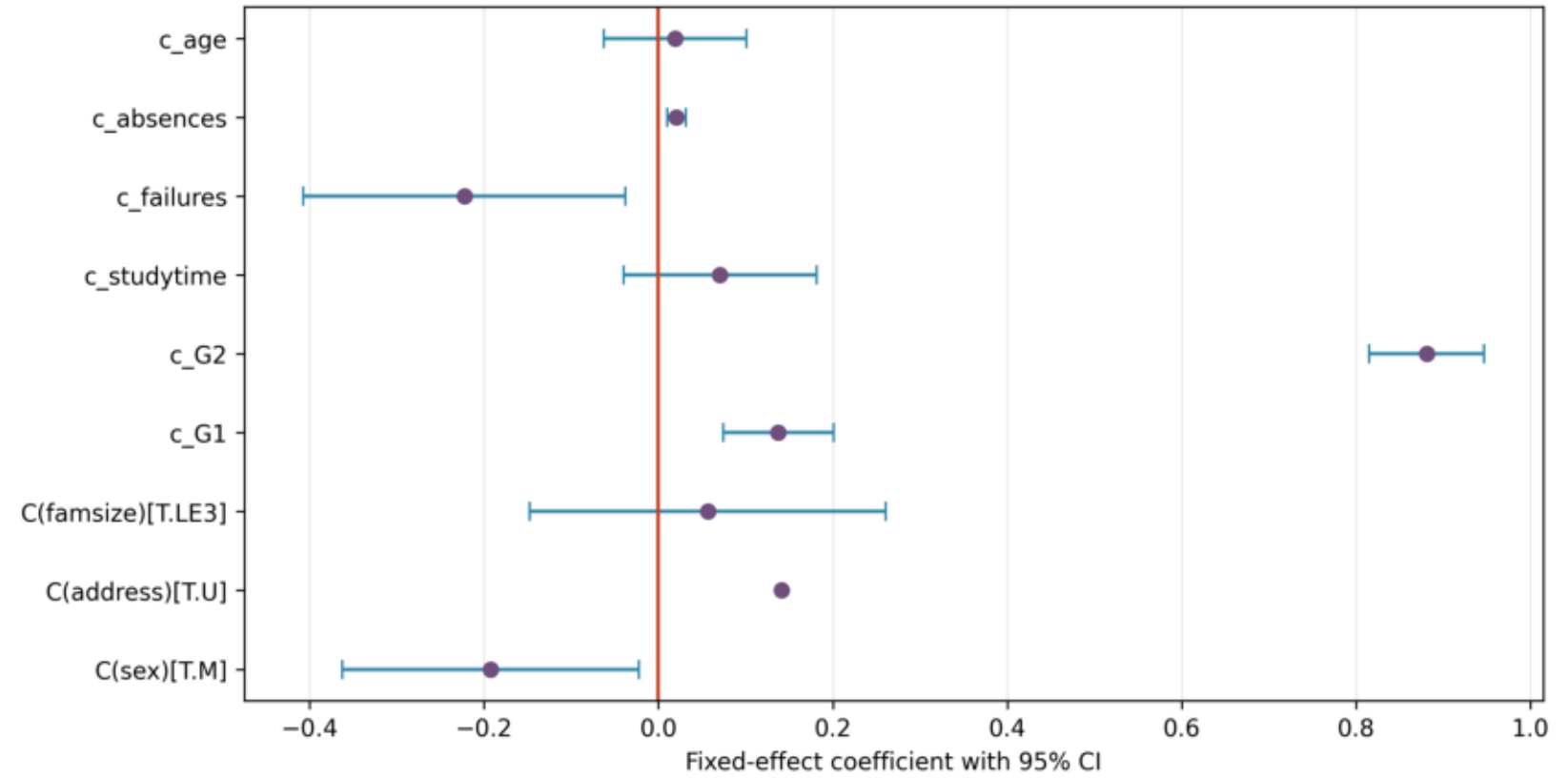
Variance components

ICC comes from the ratio of group-level variance to total model variance.



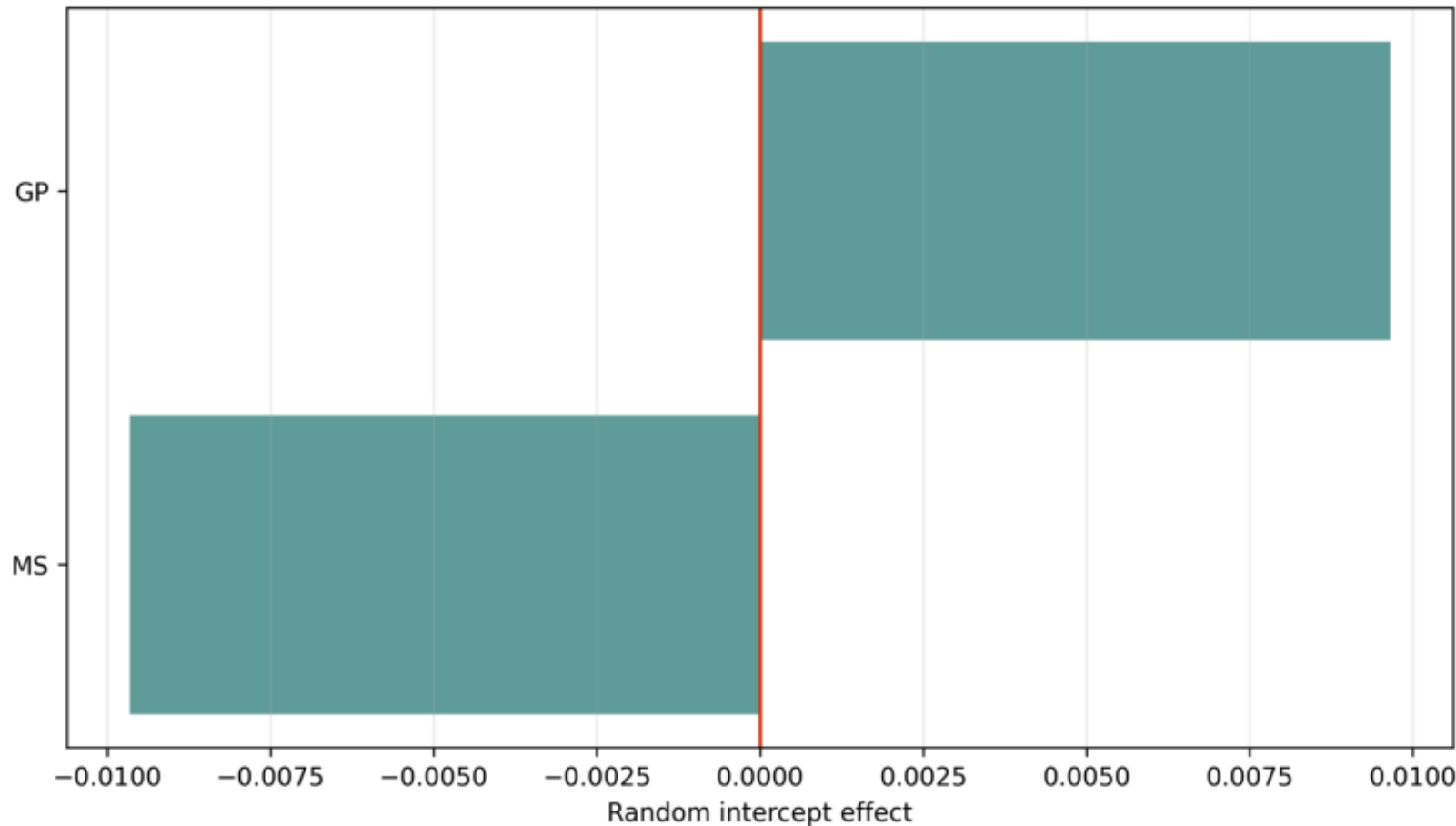
Fixed effects from multilevel model

Intervals not crossing zero show stronger evidence for predictor effects.



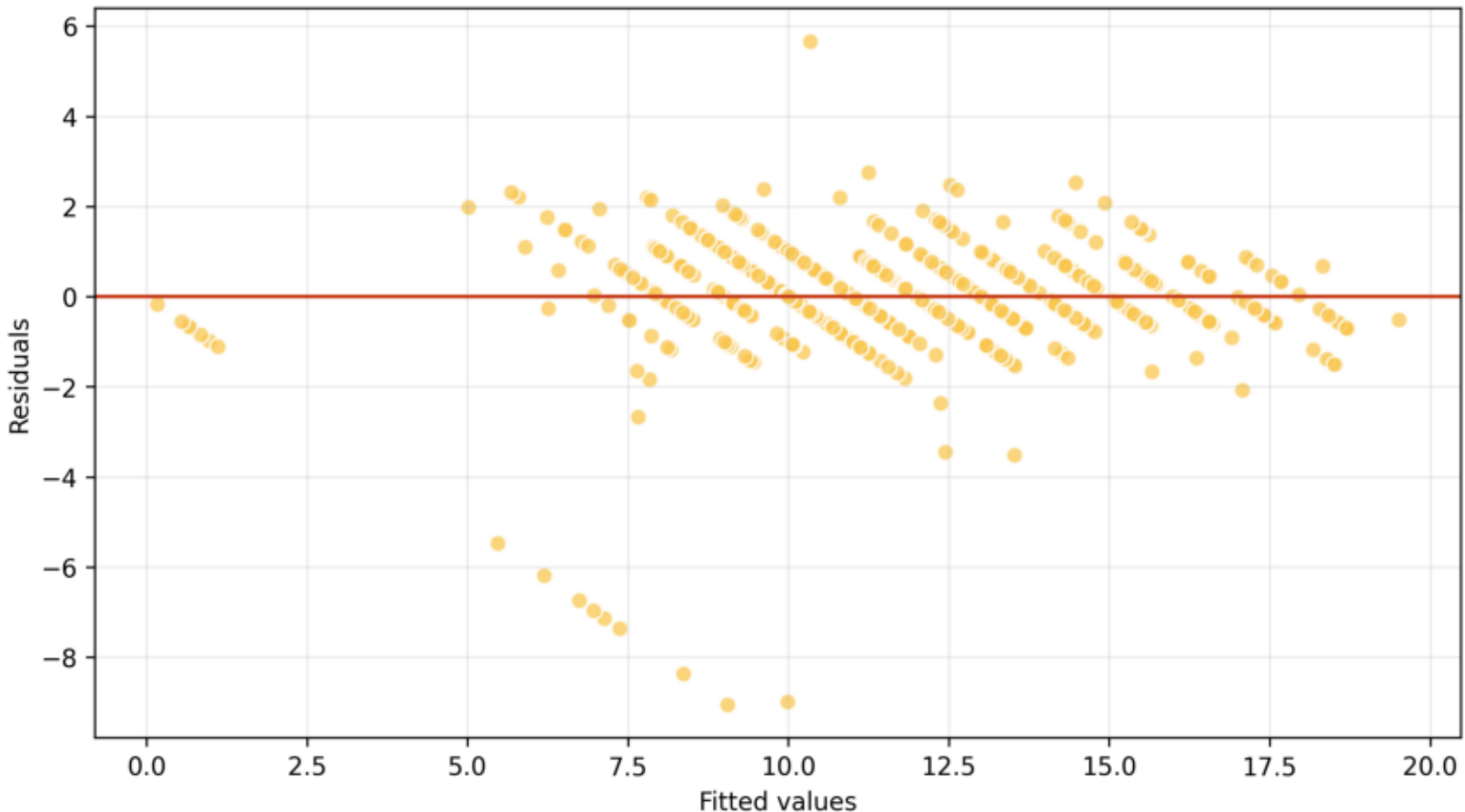
Group random intercept effects

Positive values indicate groups above the average intercept after fixed effects.



Multilevel residuals vs fitted values

Random scatter around zero supports the fitted model structure.



Observed vs fitted values

Points near the diagonal indicate stronger predictive agreement.

