

Adjusted R Squared

Test name: Adjusted R Squared

Dataset: D:\DATA ANALYSIS\H Regression Tests and Models\Adjusted R Squared\dataset.csv

Rows in original dataset: 649

Columns in original dataset: 33

Rows used in adjusted R squared model: 649

Outcome variable: G3

Numeric predictors considered: G1, G2, studytime, failures, absences, age, Medu, Fedu, traveltime

Categorical predictors considered: school, sex, address, famsize

Number of models compared: 5

Best model by adjusted R squared: Model 5: full adjusted model

Best model R squared: 0.8534

Best model adjusted R squared: 0.8504

R squared adjustment penalty: 0.0030

Best model F statistic: 284.4254

Best model F-test p-value: 1.45411e-254

Adjusted $R^2 = 1 - (1 - R^2) * (n - 1) / (n - p - 1)$.

Interpretation: adjusted R squared rewards useful explanatory power while penalizing unnecessary predictors.

Decision guide: prefer the model with the highest adjusted R squared when models use the same outcome and sample.

Reporting guide: state R^2 , adjusted R^2 , n, predictor df, F test, p-value, RMSE, and the final model predictors.

Charts generated: 7

Adjusted R squared model comparison

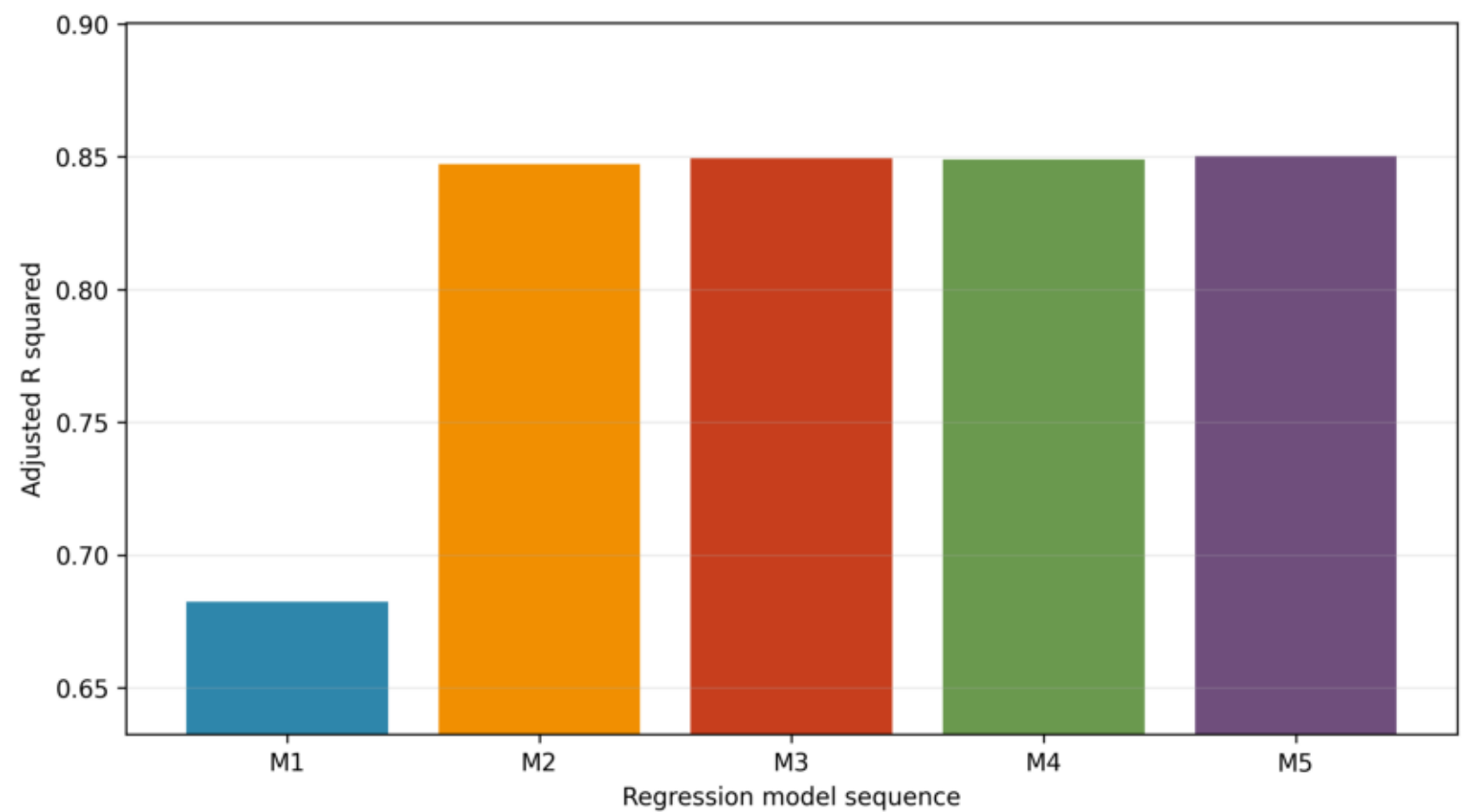
model	n	predictor_df	r_squared	adjusted_r_s	adjustment	f_statistic	f_p_value	rmse	mae	aic	bic	decision_alpha
Model 1: bas	649.0000	1.0000	0.6829	0.6824	0.0005	1393.4667	0.0000	1.8178	1.2030	2621.4936	2630.4445	Overall model
Model 2: pri	649.0000	2.0000	0.8478	0.8473	0.0005	1798.6714	0.0000	1.2596	0.7851	2147.3096	2160.7359	Overall model
Model 3: acc	649.0000	5.0000	0.8506	0.8494	0.0012	731.9659	0.0000	1.2479	0.7830	2141.2537	2168.1063	Overall model
Model 4: acc	649.0000	9.0000	0.8511	0.8490	0.0021	405.9294	0.0000	1.2455	0.7788	2146.7840	2191.5384	Overall model
Model 5: ful	649.0000	13.0000	0.8534	0.8504	0.0030	284.4254	0.0000	1.2359	0.7777	2144.6640	2207.3200	Overall model

Final model coefficients

model	term	coefficient	standard_error	t_value	p_value	ci95_lower	ci95_upper	decision_alpha_0.05
Model 5: full adju	Intercept	-0.5202	0.8029	-0.6479	0.5173	-2.0970	1.0565	Not statistically significant
Model 5: full adju	C(school)[T.MS]	-0.2083	0.1199	-1.7369	0.0829	-0.4437	0.0272	Not statistically significant
Model 5: full adju	C(sex)[T.M]	-0.2165	0.1054	-2.0534	0.0404	-0.4235	-0.0095	Statistically significant
Model 5: full adju	C(address)[T.U]	0.1688	0.1195	1.4133	0.1580	-0.0657	0.4034	Not statistically significant
Model 5: full adju	C(famsize)[T.LE3]	0.0608	0.1090	0.5578	0.5772	-0.1532	0.2747	Not statistically significant
Model 5: full adju	G1	0.1330	0.0368	3.6153	0.0003	0.0607	0.2052	Statistically significant
Model 5: full adju	G2	0.8817	0.0342	25.7601	0.0000	0.8145	0.9489	Statistically significant
Model 5: full adju	studytime	0.0630	0.0633	0.9952	0.3200	-0.0613	0.1872	Not statistically significant
Model 5: full adju	failures	-0.2313	0.0952	-2.4308	0.0153	-0.4182	-0.0444	Statistically significant
Model 5: full adju	absences	0.0167	0.0112	1.4944	0.1356	-0.0052	0.0386	Not statistically significant
Model 5: full adju	age	0.0238	0.0436	0.5461	0.5852	-0.0618	0.1094	Not statistically significant
Model 5: full adju	Medu	-0.0341	0.0591	-0.5763	0.5646	-0.1501	0.0820	Not statistically significant

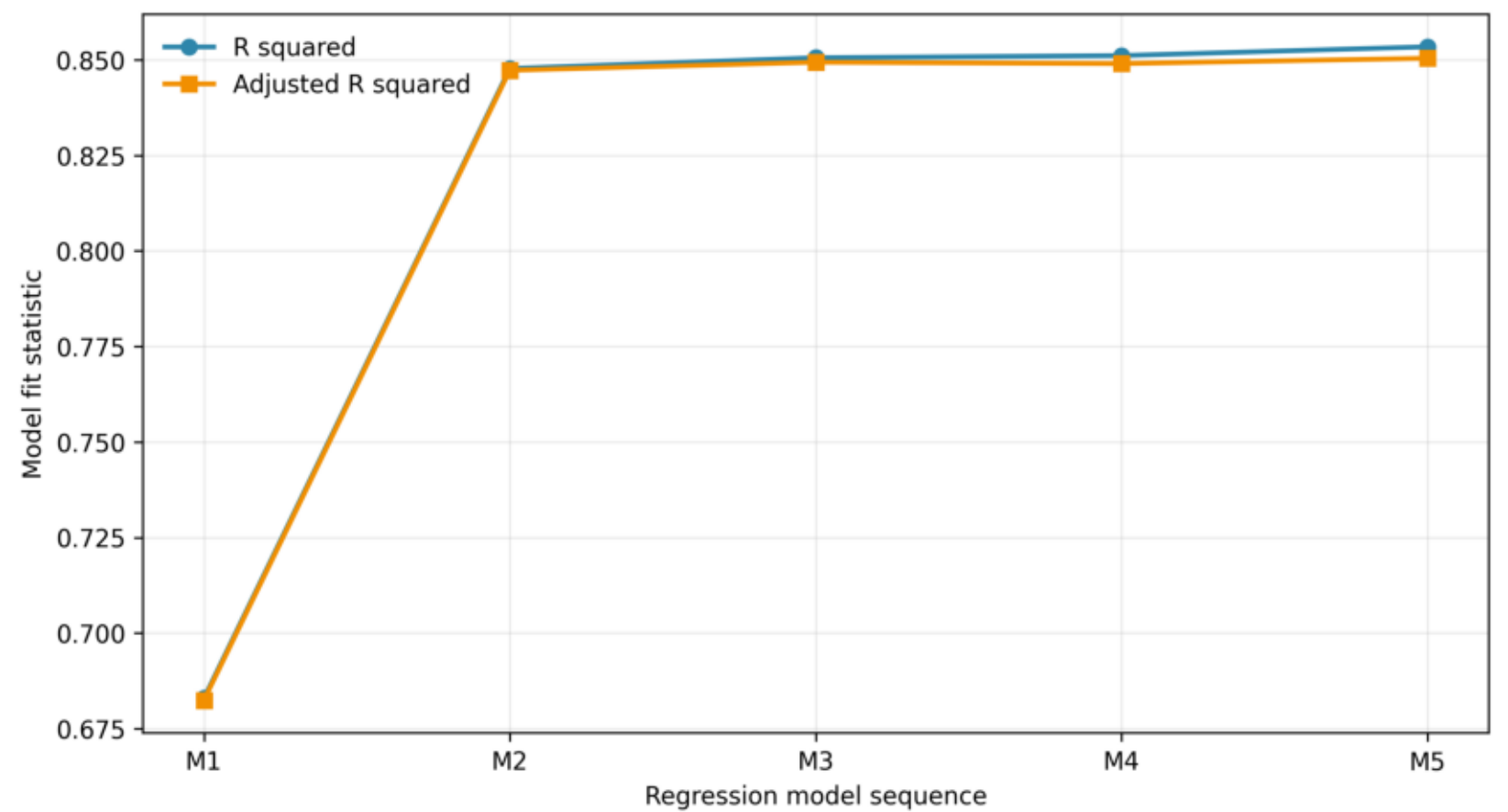
Adjusted R squared model comparison

Higher adjusted R squared indicates better fit after penalizing unnecessary predictors.



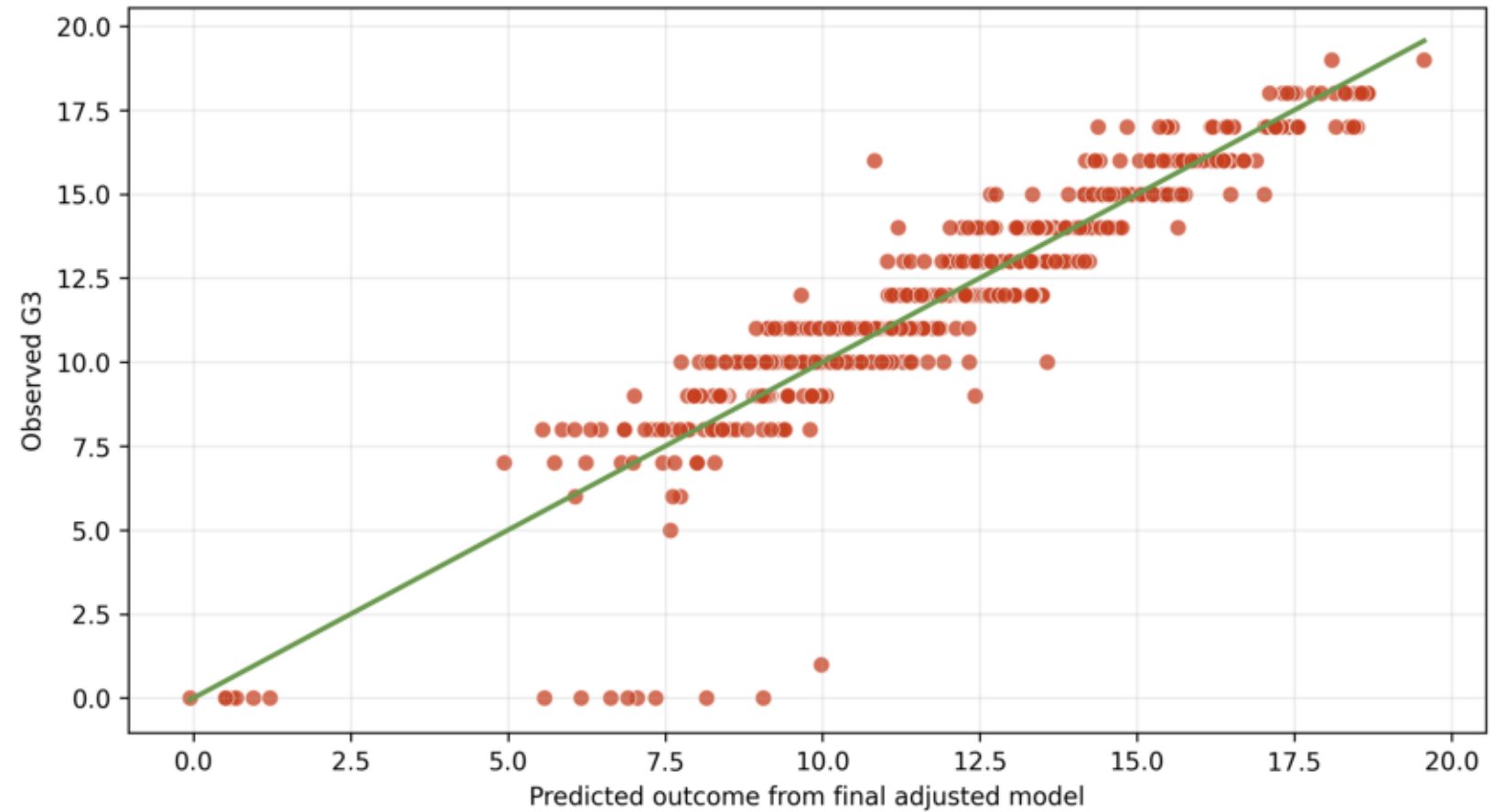
R squared vs adjusted R squared

The gap shows the penalty applied for predictor complexity.



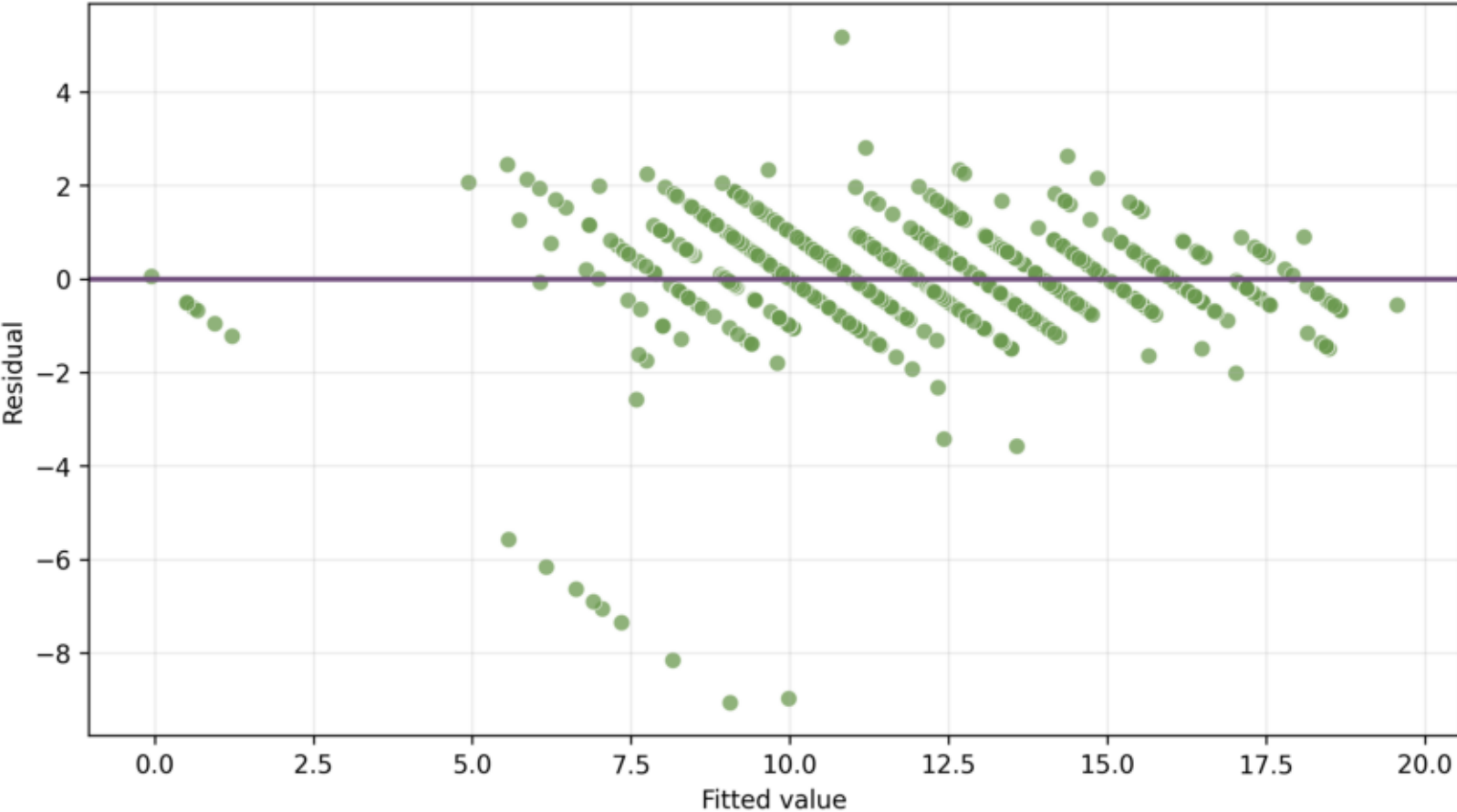
Observed vs predicted final model

Points near the diagonal show stronger prediction agreement.



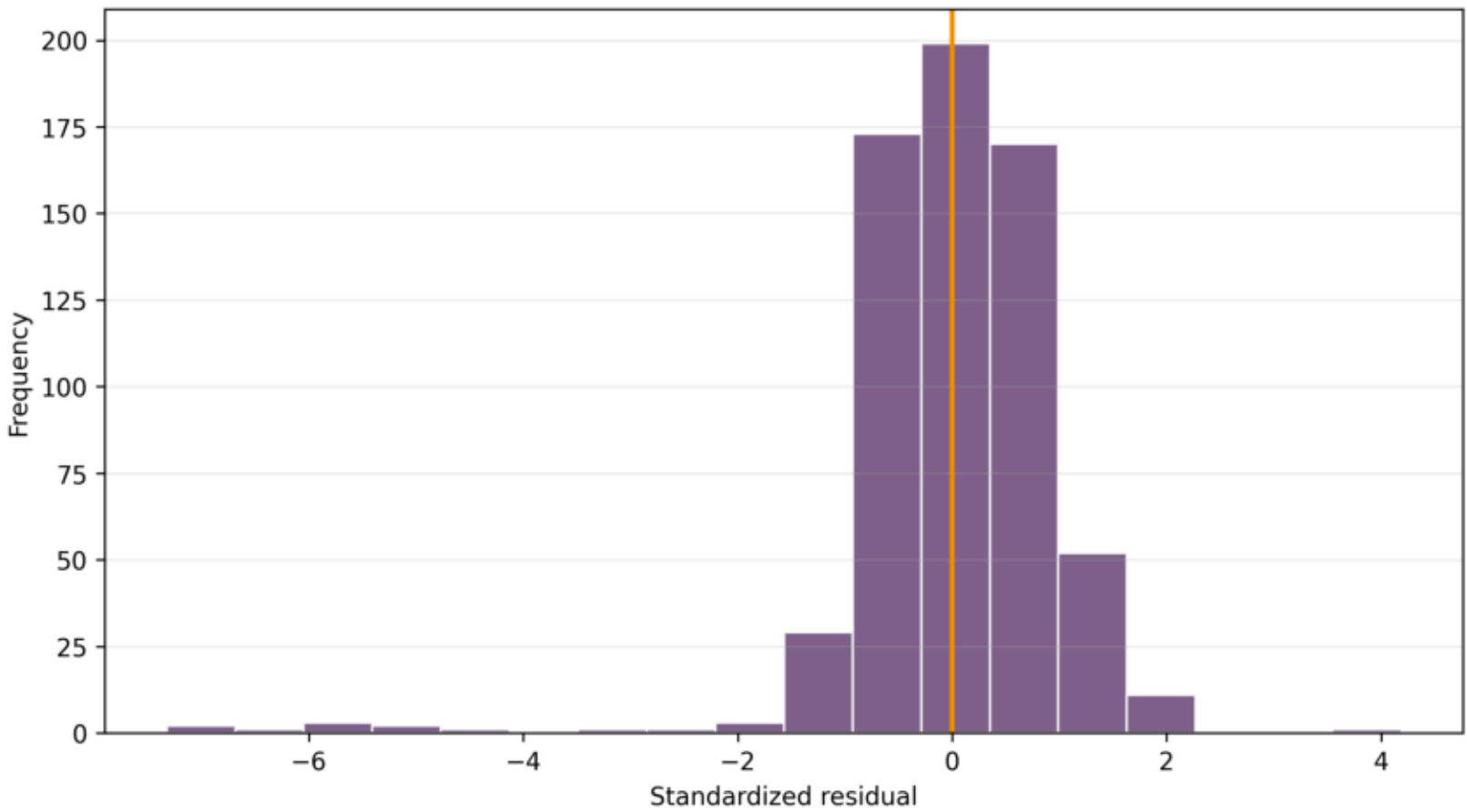
Residuals vs fitted values

A random band around zero supports a reasonable linear-regression fit.



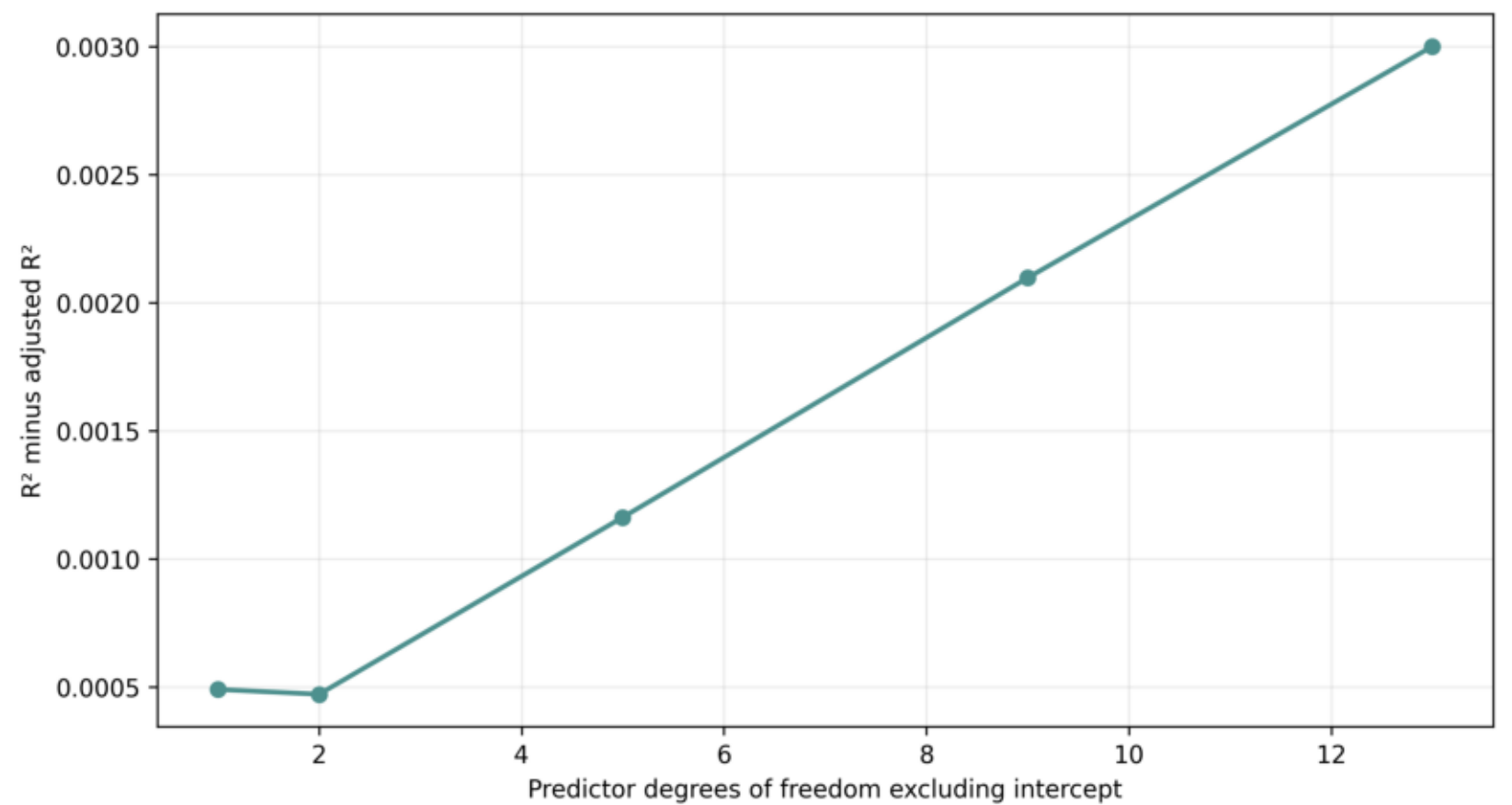
Standardized residual distribution

This chart checks whether the final model errors are centered near zero.



Adjusted R squared complexity penalty

The penalty grows as the model uses more predictor degrees of freedom.



Top coefficient evidence in final adjusted model

Larger absolute t values indicate stronger evidence for individual predictors.

