

Supplement to “Extended Comparisons of Best Subset Selection, Forward Stepwise Selection, and the Lasso”

Trevor Hastie Robert Tibshirani Ryan J. Tibshirani

This supplementary document contains plots from the simulation suite described in the paper “Extended Comparisons of Best Subset Selection, Forward Stepwise Selection, and the Lasso”. The plots in Section 1 precisely follow the simulation format described in the paper. Those in Section 2 follow an analogous format, except that the tuning has been done using an “oracle”, rather than a validation set as in Section 1. Specifically, the tuning parameter for each method in each scenario is chosen to minimize the average risk over all of the repetitions.

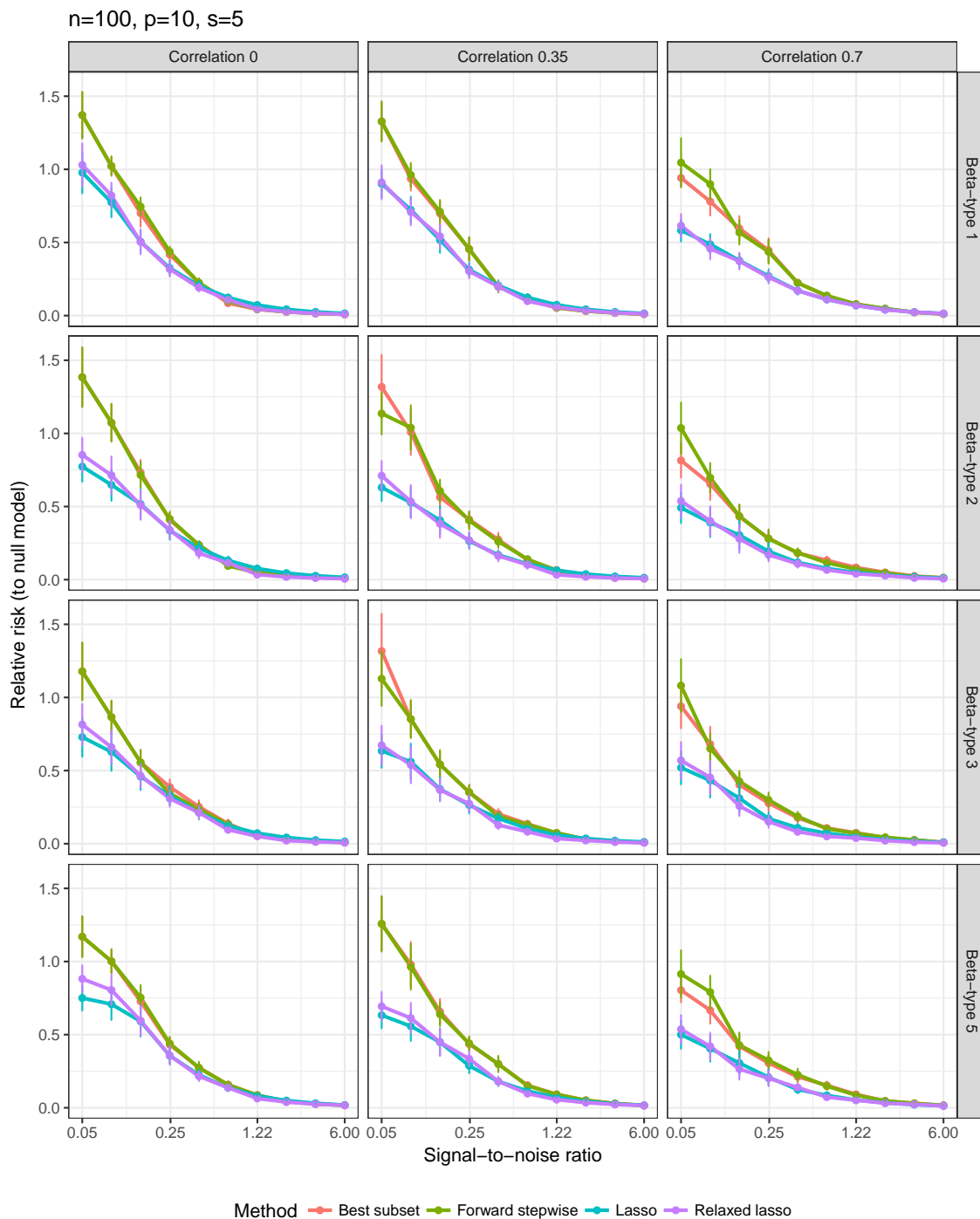
Contents

A	Validation tuning	3
A.1	Low setting: $n = 100, p = 10, s = 5$	3
A.1.1	Relative risk (to null model)	3
A.1.2	Relative test error (to Bayes)	4
A.1.3	Proportion of variance explained	5
A.1.4	Number of nonzero coefficients	6
A.2	Medium setting: $n = 500, p = 100, s = 5$	7
A.2.1	Relative risk (to null model)	7
A.2.2	Relative test error (to Bayes)	8
A.2.3	Proportion of variance explained	9
A.2.4	Number of nonzero coefficients	10
A.3	High-5 setting: $n = 50, p = 1000, s = 5$	11
A.3.1	Relative risk (to null model)	11
A.3.2	Relative test error (to Bayes)	12
A.3.3	Proportion of variance explained	13
A.3.4	Number of nonzero coefficients	14
A.4	High-10 setting: $n = 100, p = 1000, s = 10$	15
A.4.1	Relative risk (to null model)	15
A.4.2	Relative test error (to Bayes)	16
A.4.3	Proportion of variance explained	17
A.4.4	Number of nonzero coefficients	18
B	Oracle tuning	19
B.1	Low setting: $n = 100, p = 10, s = 5$	19
B.1.1	Relative risk (to null model)	19
B.1.2	Relative test error (to Bayes)	20
B.1.3	Proportion of variance explained	21
B.1.4	Number of nonzero coefficients	22
B.2	Medium setting: $n = 500, p = 100, s = 5$	23
B.2.1	Relative risk (to null model)	23
B.2.2	Relative test error (to Bayes)	24
B.2.3	Proportion of variance explained	25
B.2.4	Number of nonzero coefficients	26
B.3	High-5 setting: $n = 50, p = 1000, s = 5$	27
B.3.1	Relative risk (to null model)	27
B.3.2	Relative test error (to Bayes)	28
B.3.3	Proportion of variance explained	29
B.3.4	Number of nonzero coefficients	30
B.4	High-10 setting: $n = 100, p = 1000, s = 10$	31
B.4.1	Relative risk (to null model)	31
B.4.2	Relative test error (to Bayes)	32
B.4.3	Proportion of variance explained	33
B.4.4	Number of nonzero coefficients	34

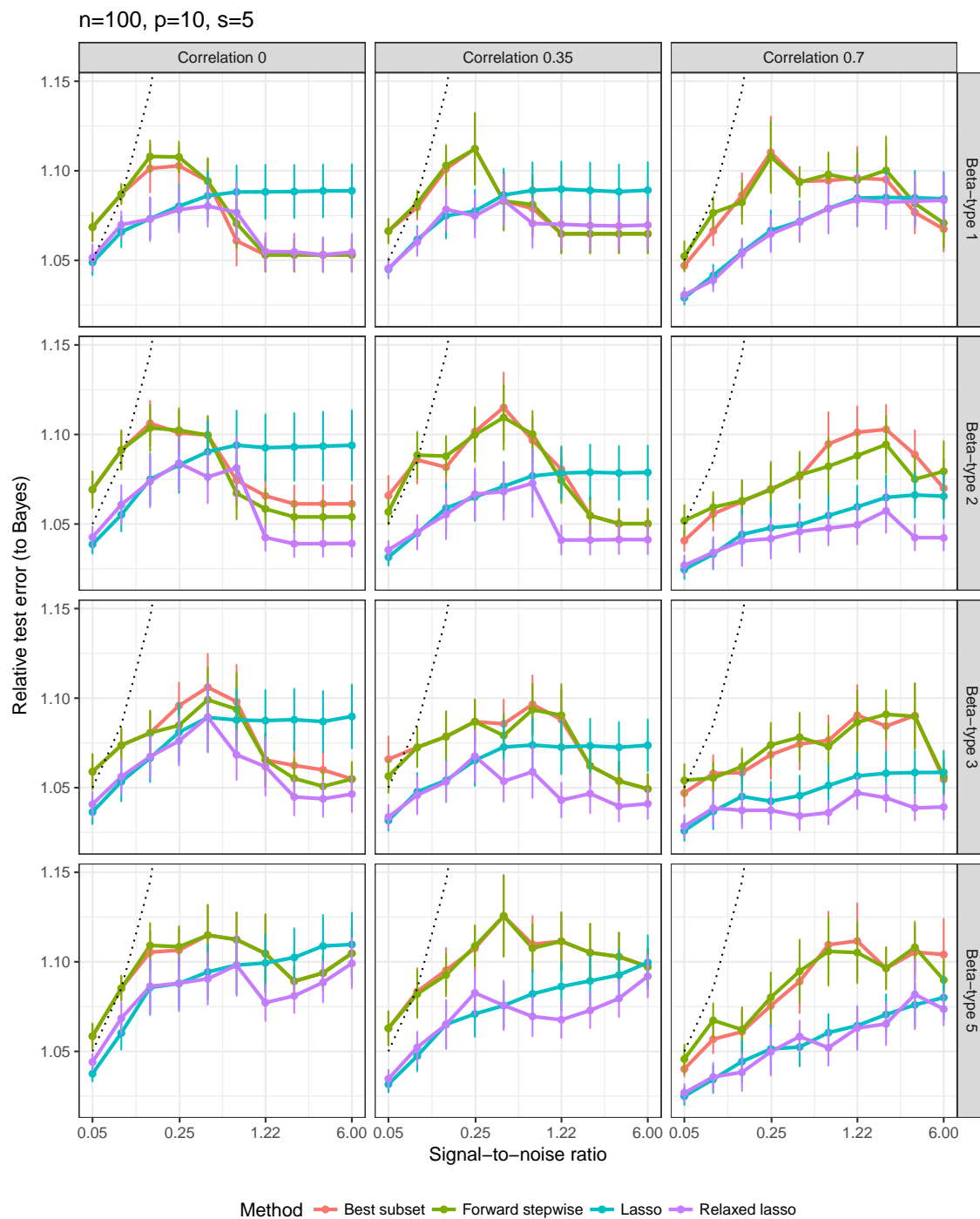
A Validation tuning

A.1 Low setting: $n = 100, p = 10, s = 5$

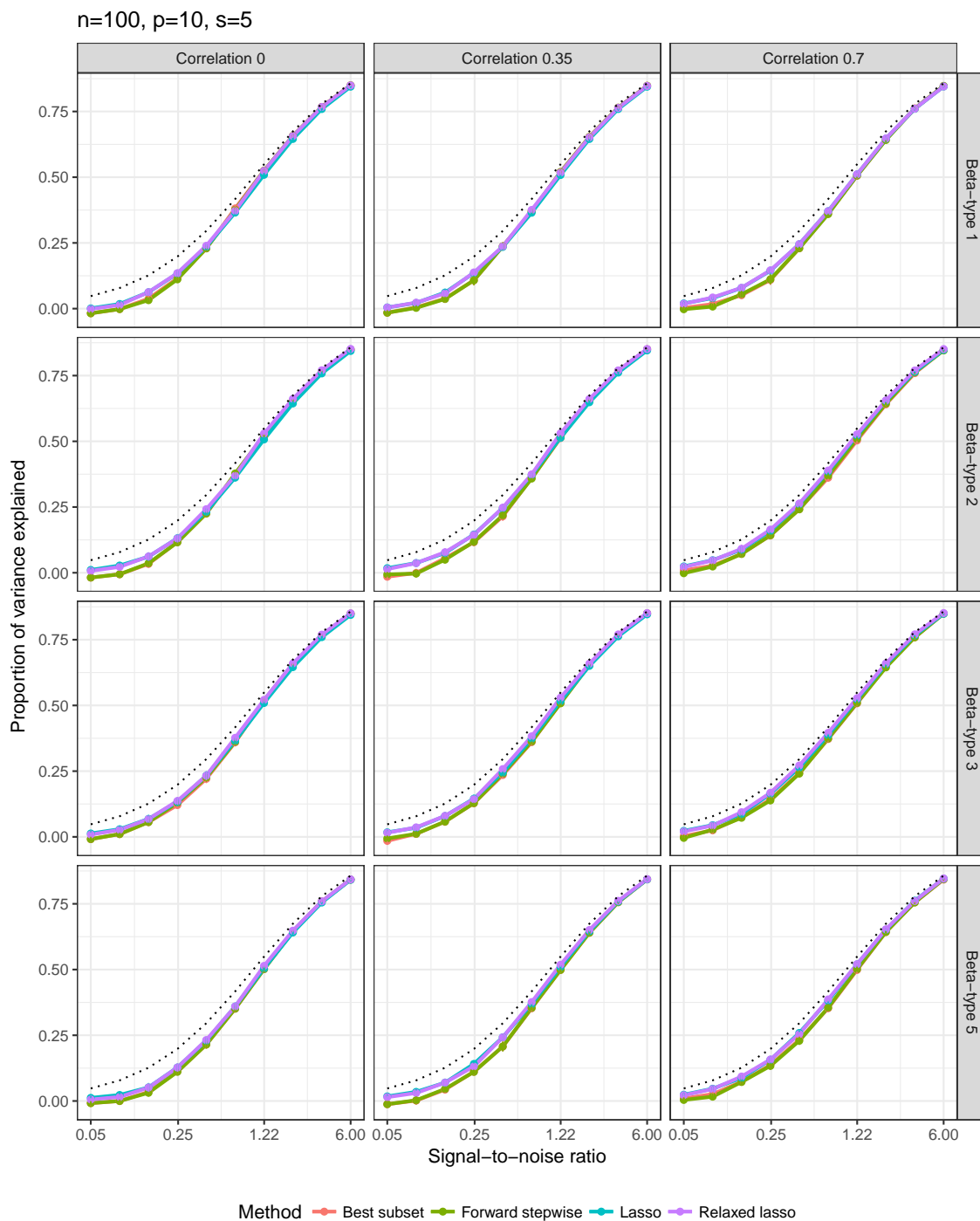
A.1.1 Relative risk (to null model)



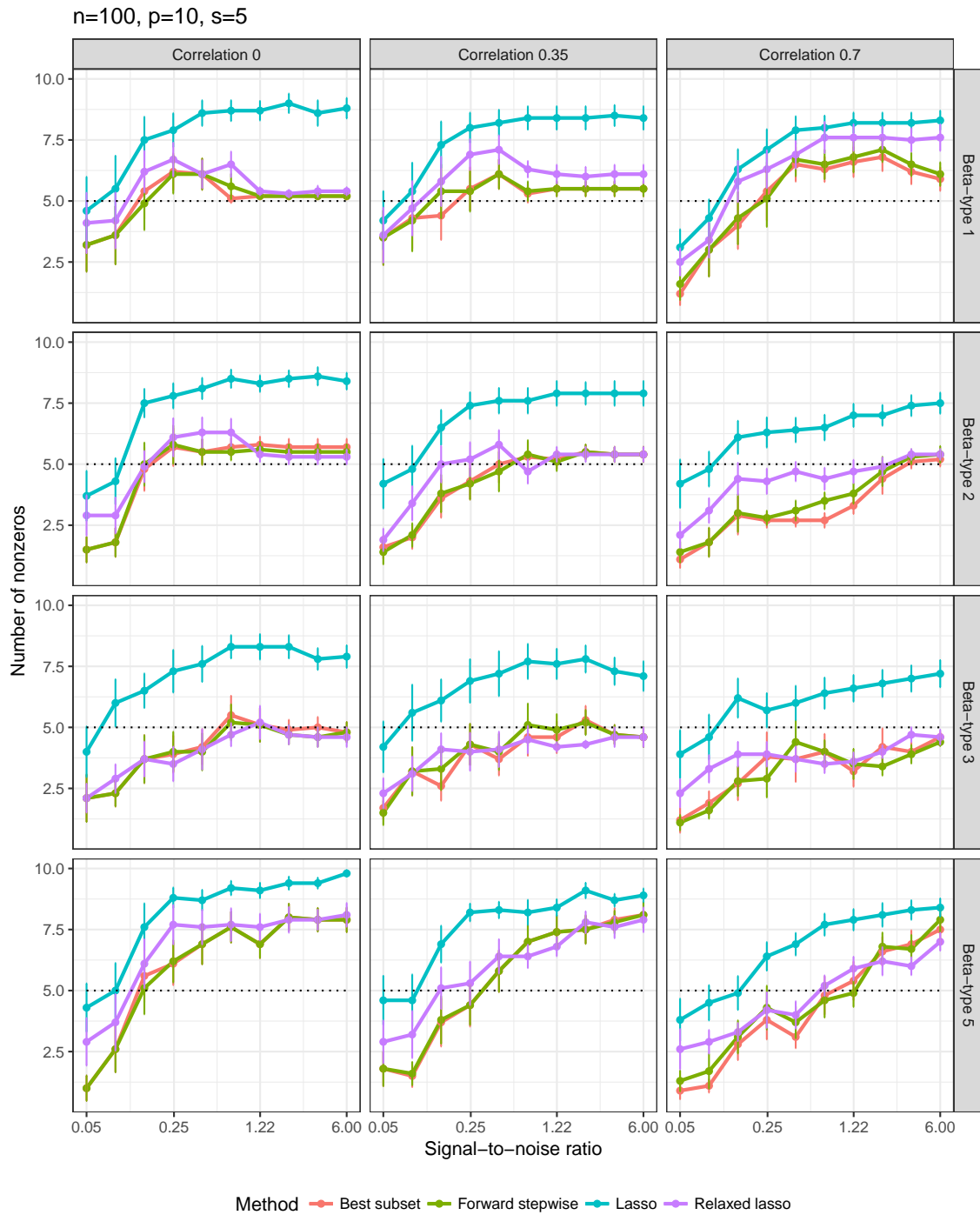
A.1.2 Relative test error (to Bayes)



A.1.3 Proportion of variance explained

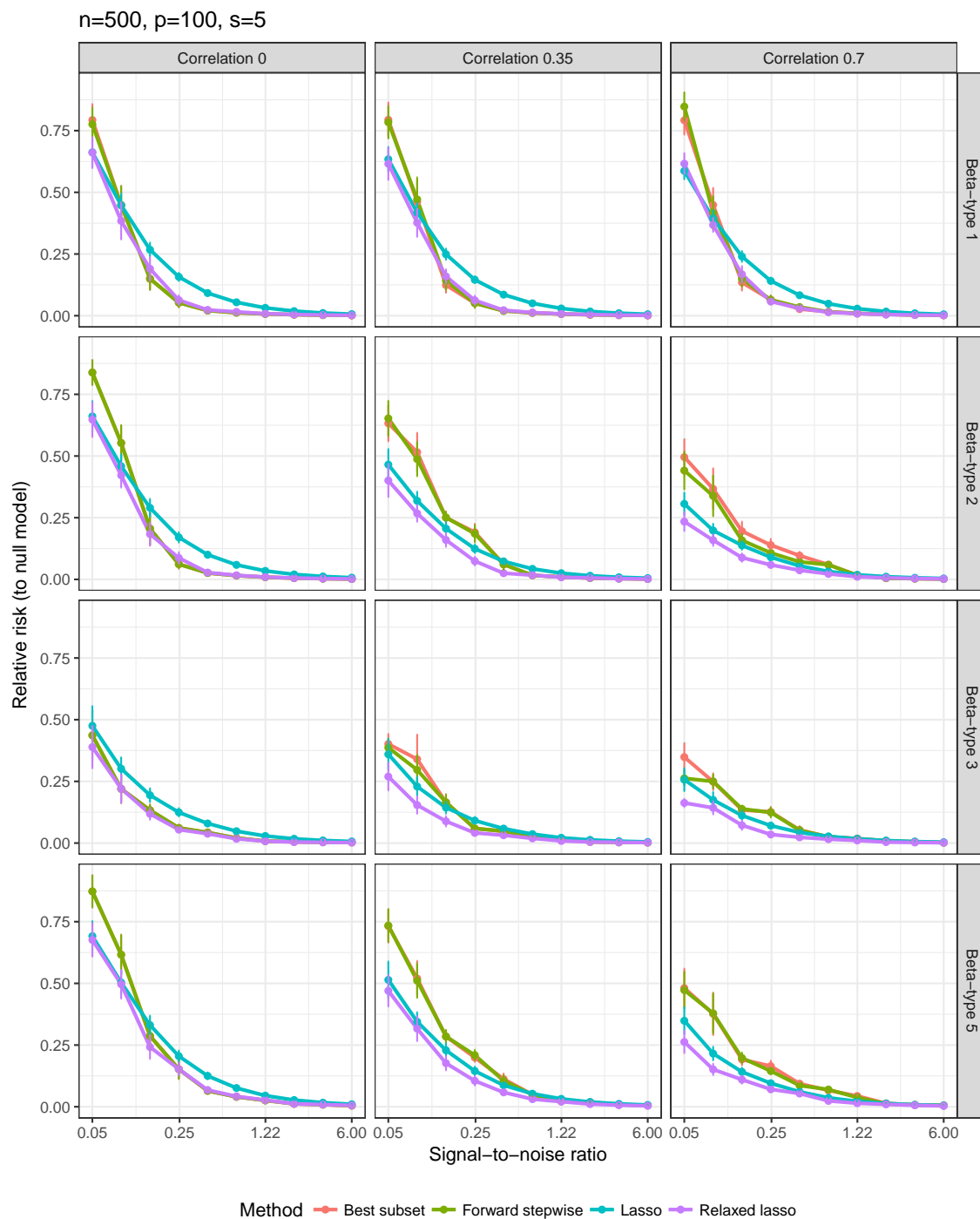


A.1.4 Number of nonzero coefficients

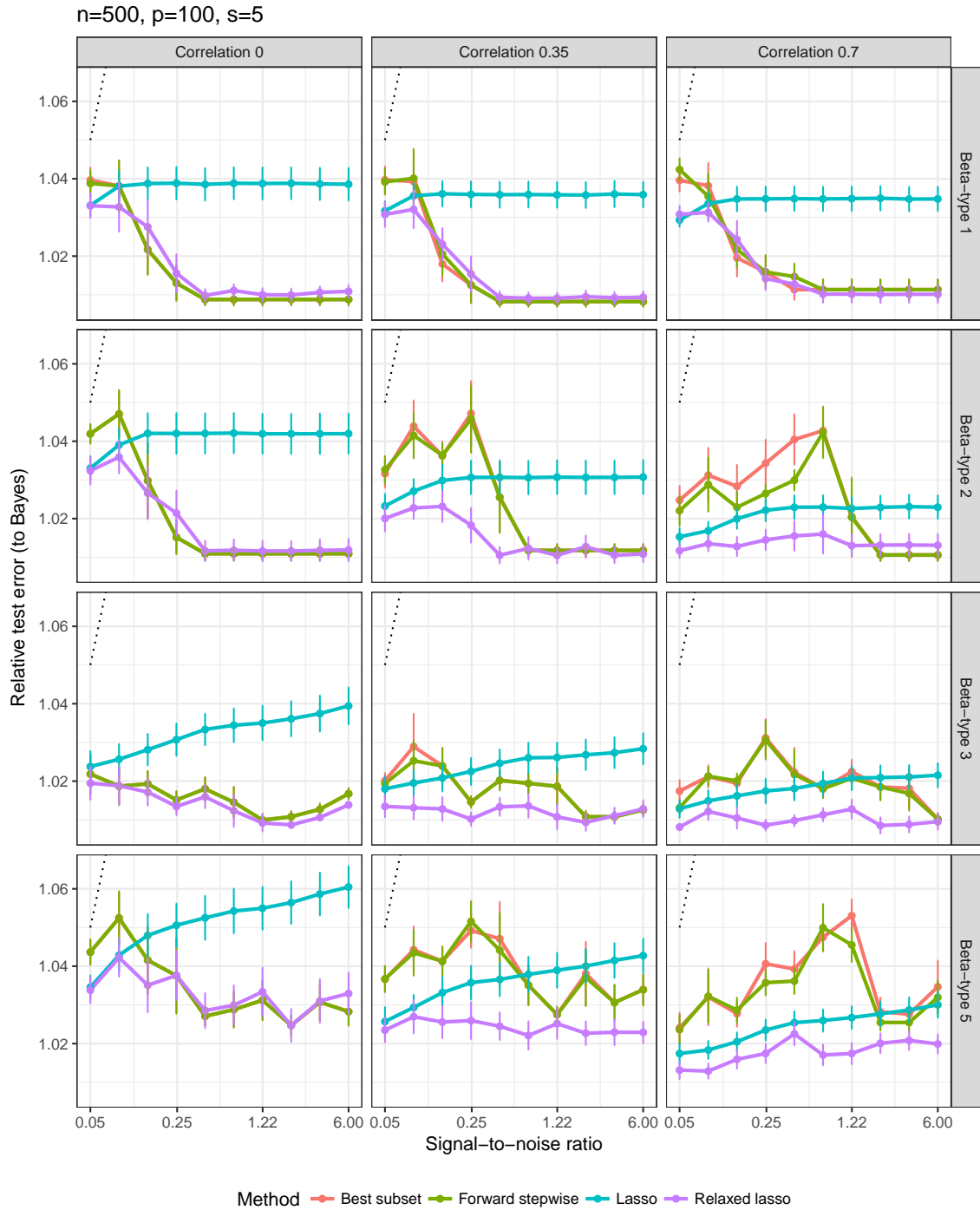


A.2 Medium setting: $n = 500, p = 100, s = 5$

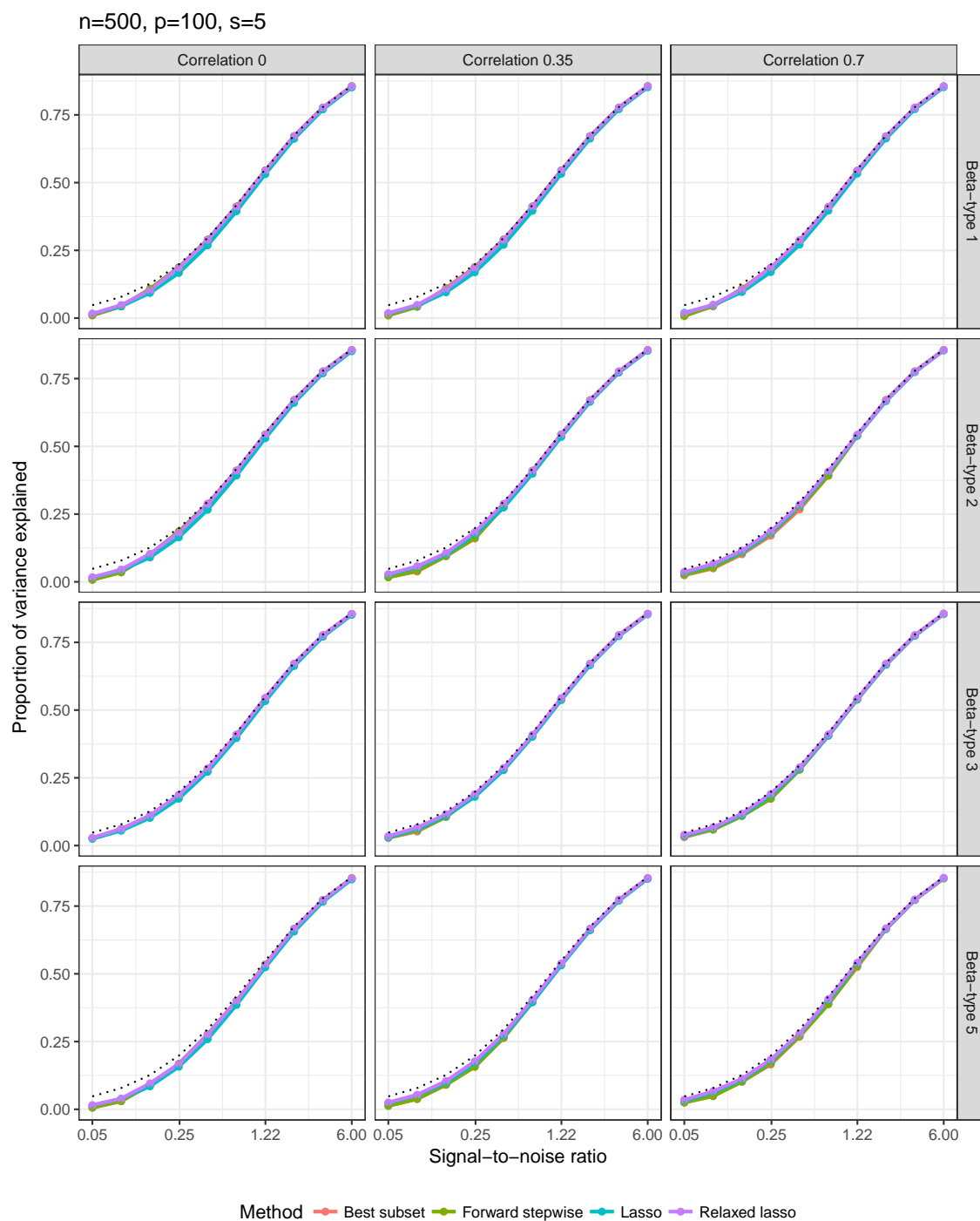
A.2.1 Relative risk (to null model)



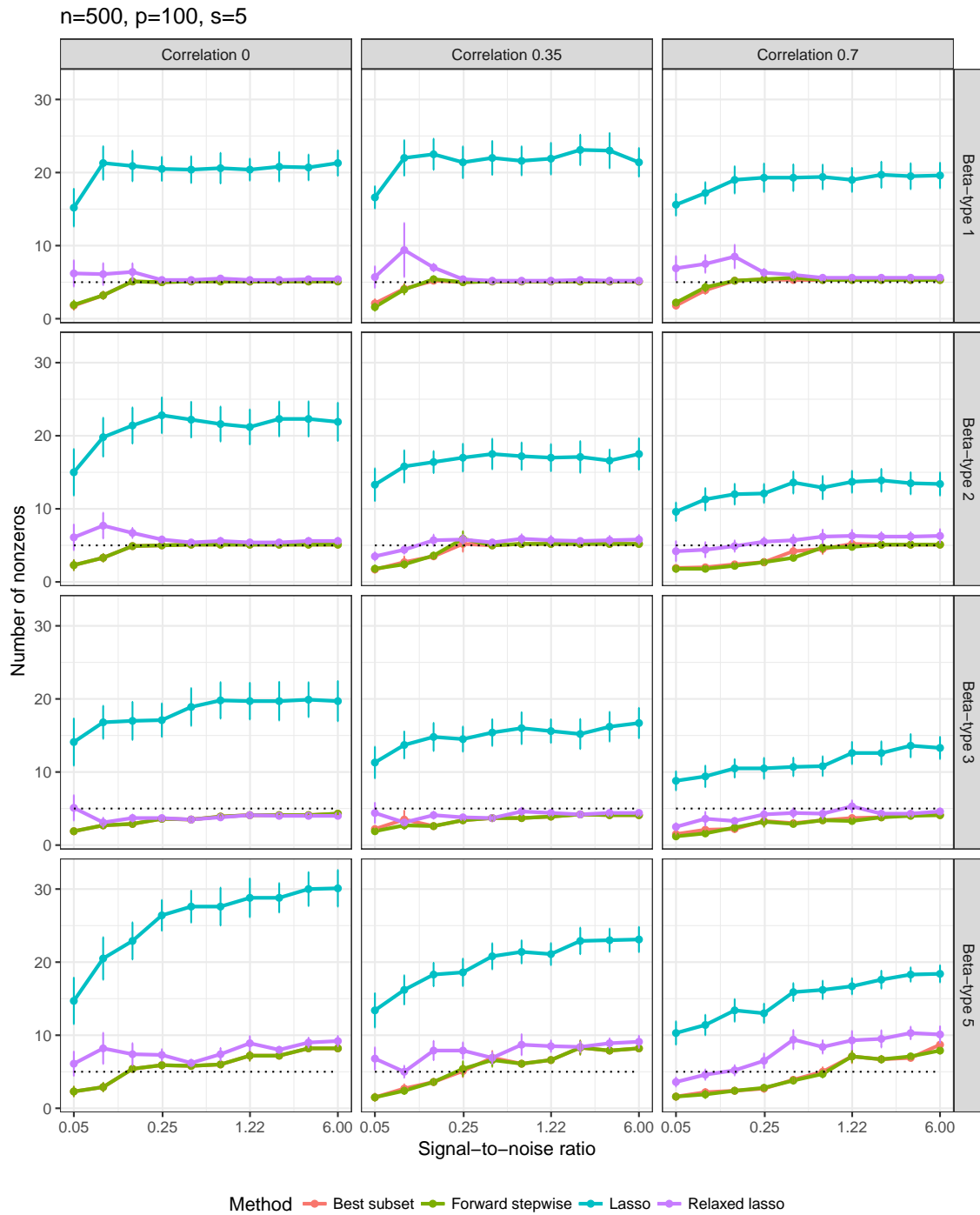
A.2.2 Relative test error (to Bayes)



A.2.3 Proportion of variance explained

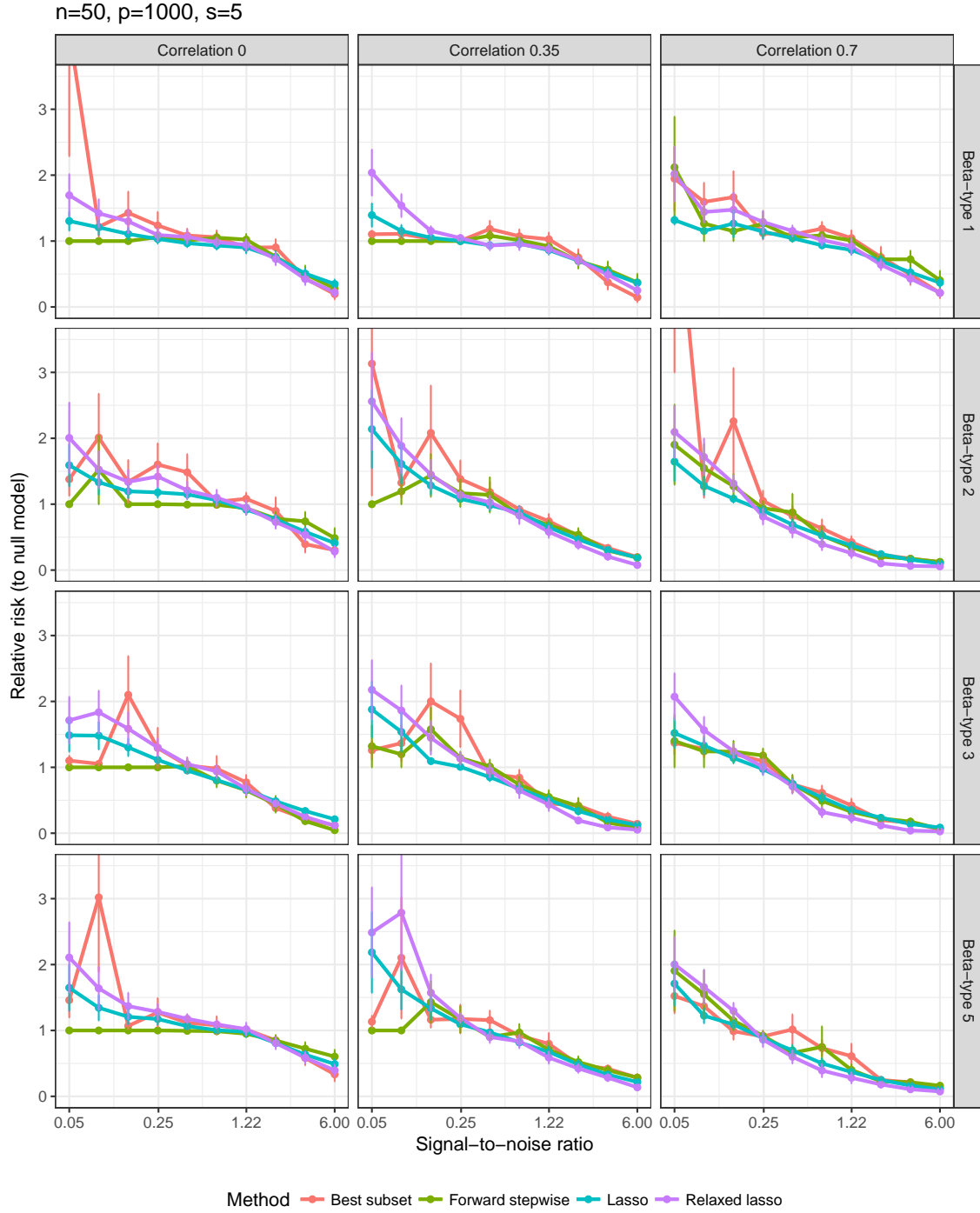


A.2.4 Number of nonzero coefficients

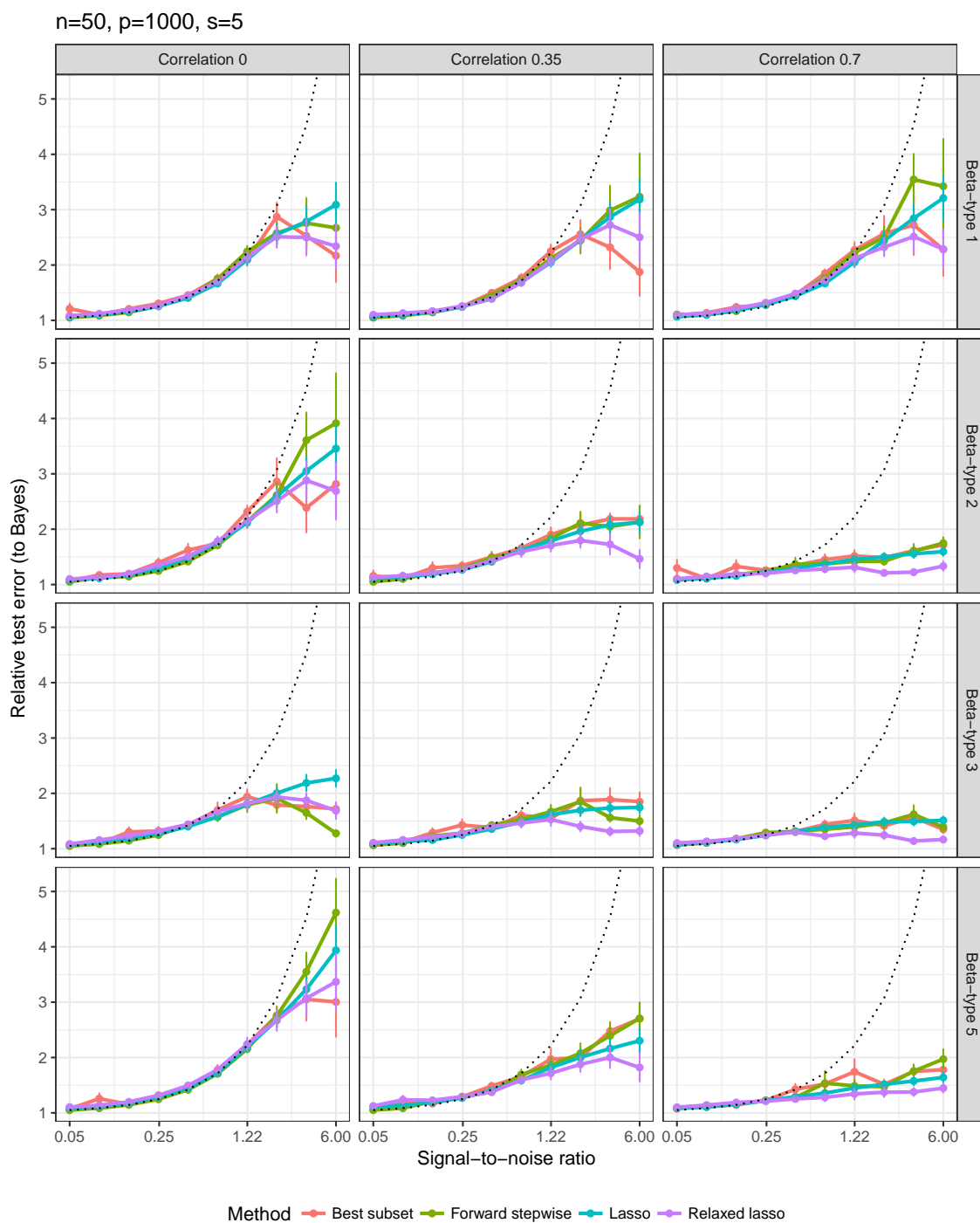


A.3 High-5 setting: $n = 50, p = 1000, s = 5$

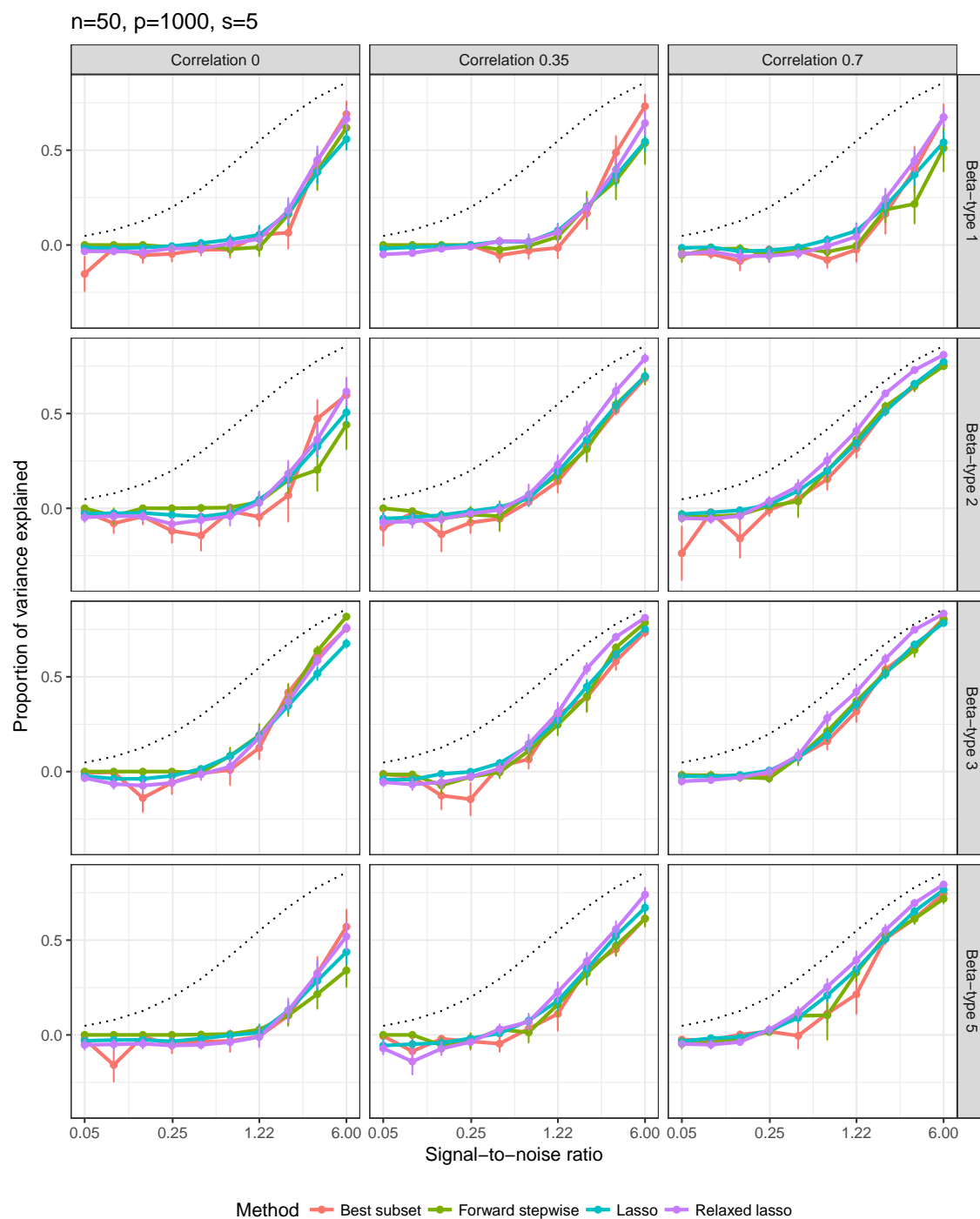
A.3.1 Relative risk (to null model)



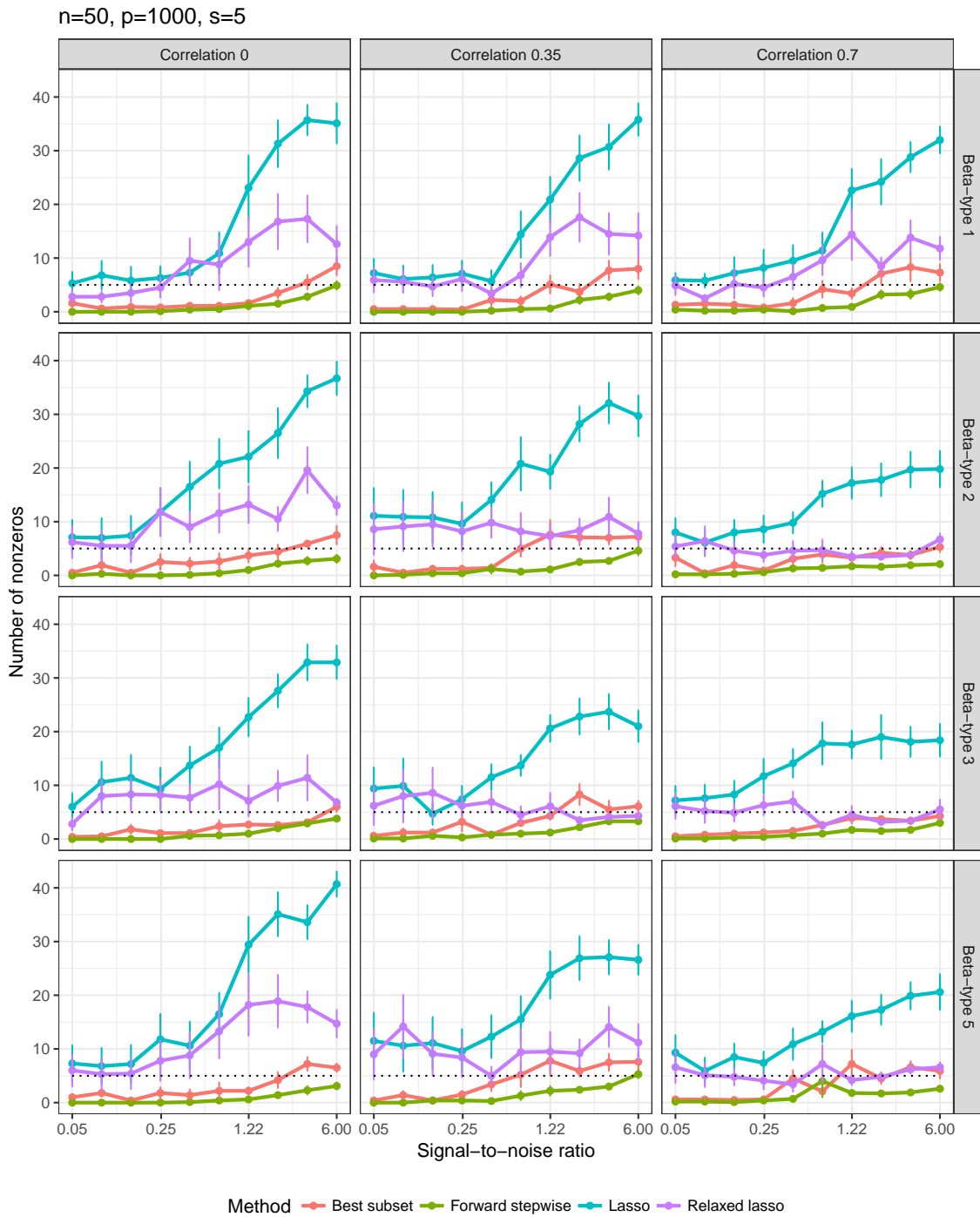
A.3.2 Relative test error (to Bayes)



A.3.3 Proportion of variance explained

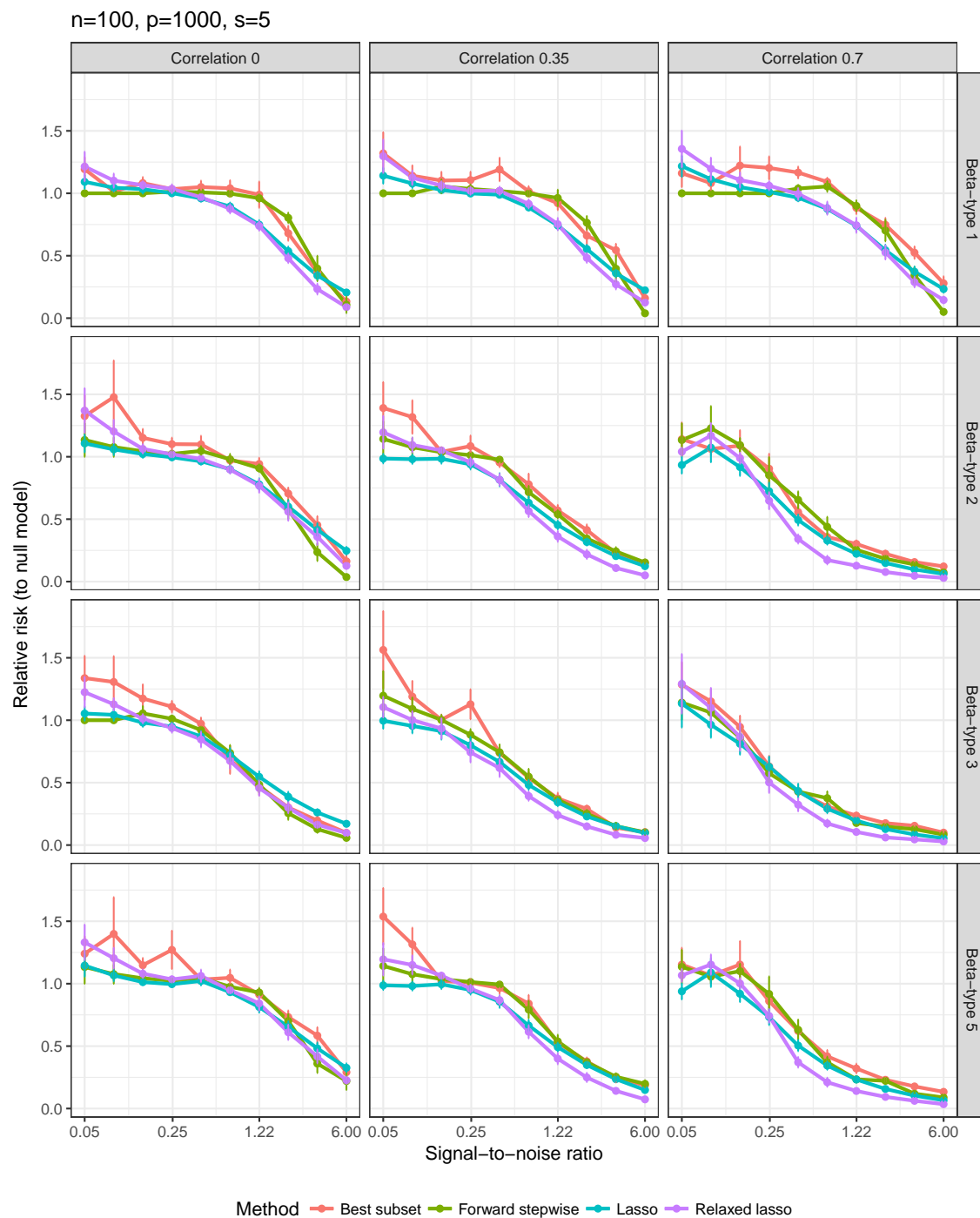


A.3.4 Number of nonzero coefficients

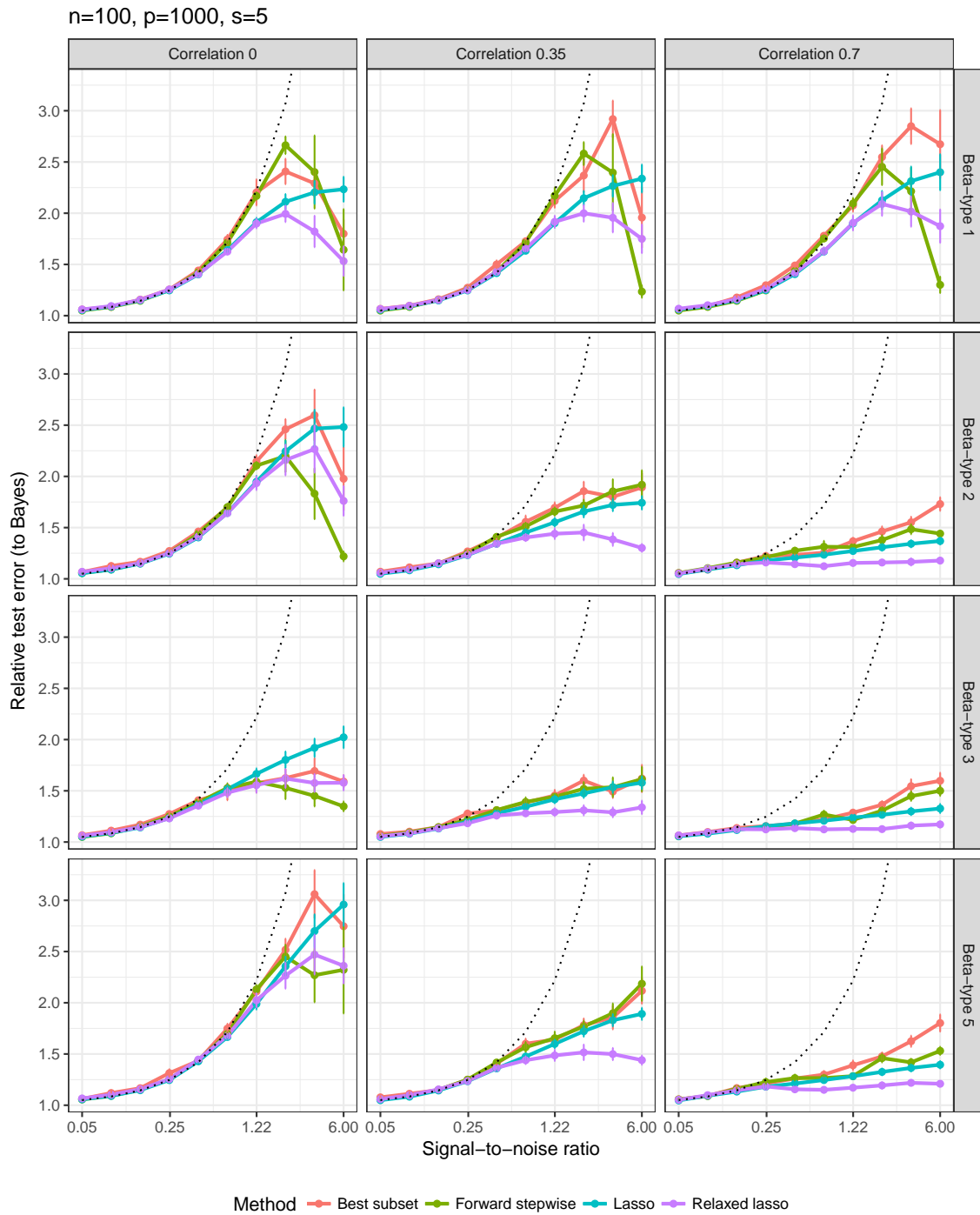


A.4 High-10 setting: $n = 100, p = 1000, s = 10$

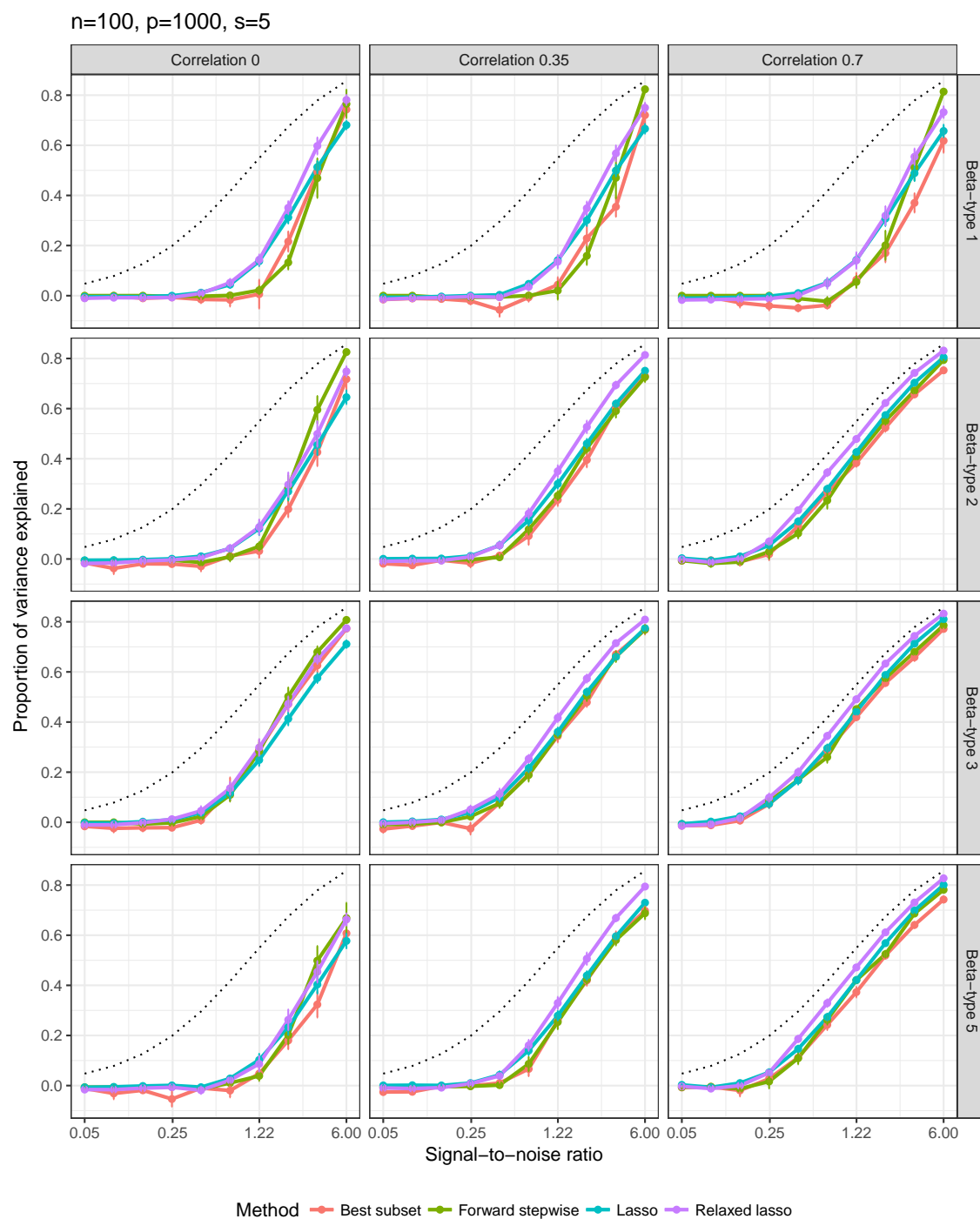
A.4.1 Relative risk (to null model)



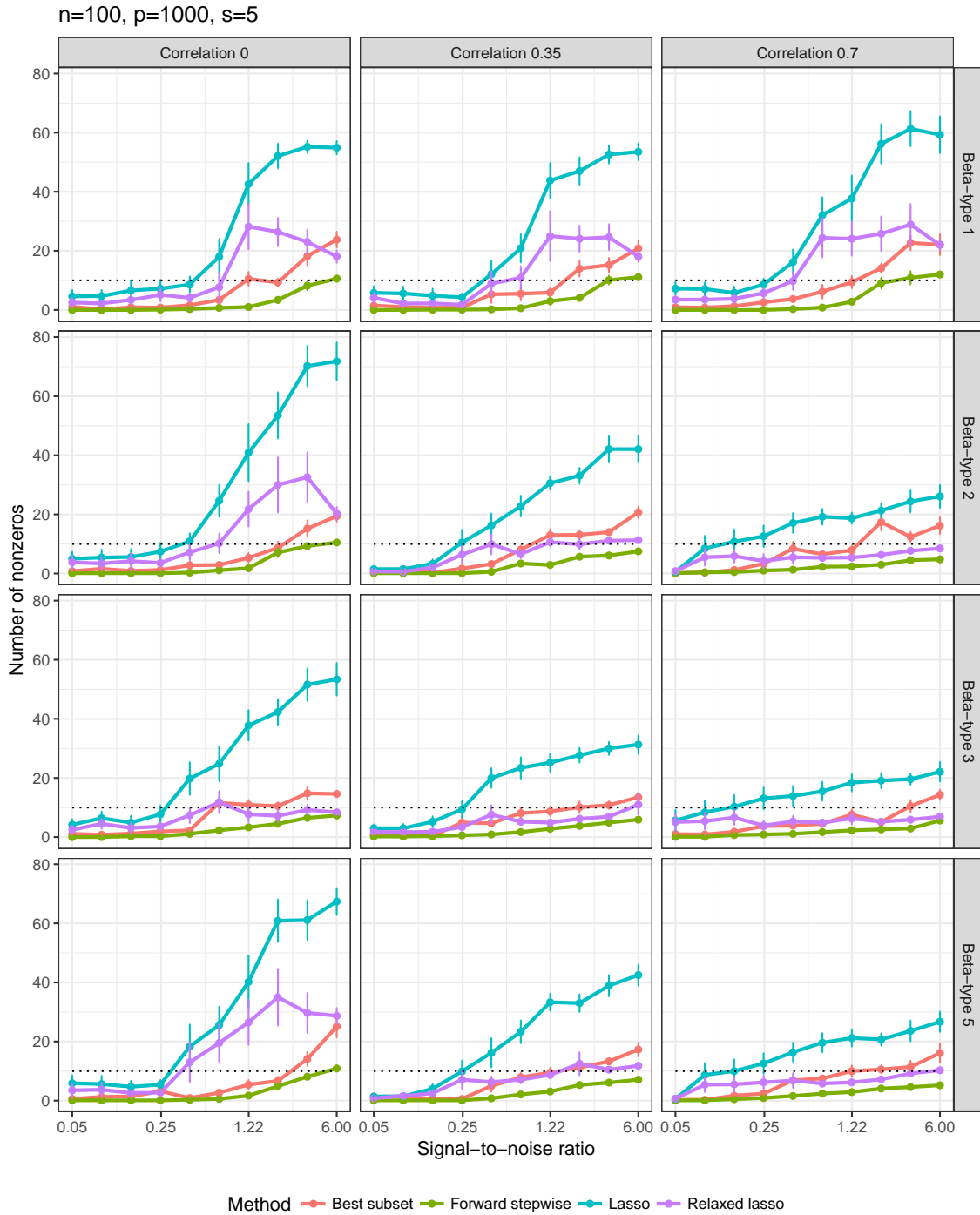
A.4.2 Relative test error (to Bayes)



A.4.3 Proportion of variance explained



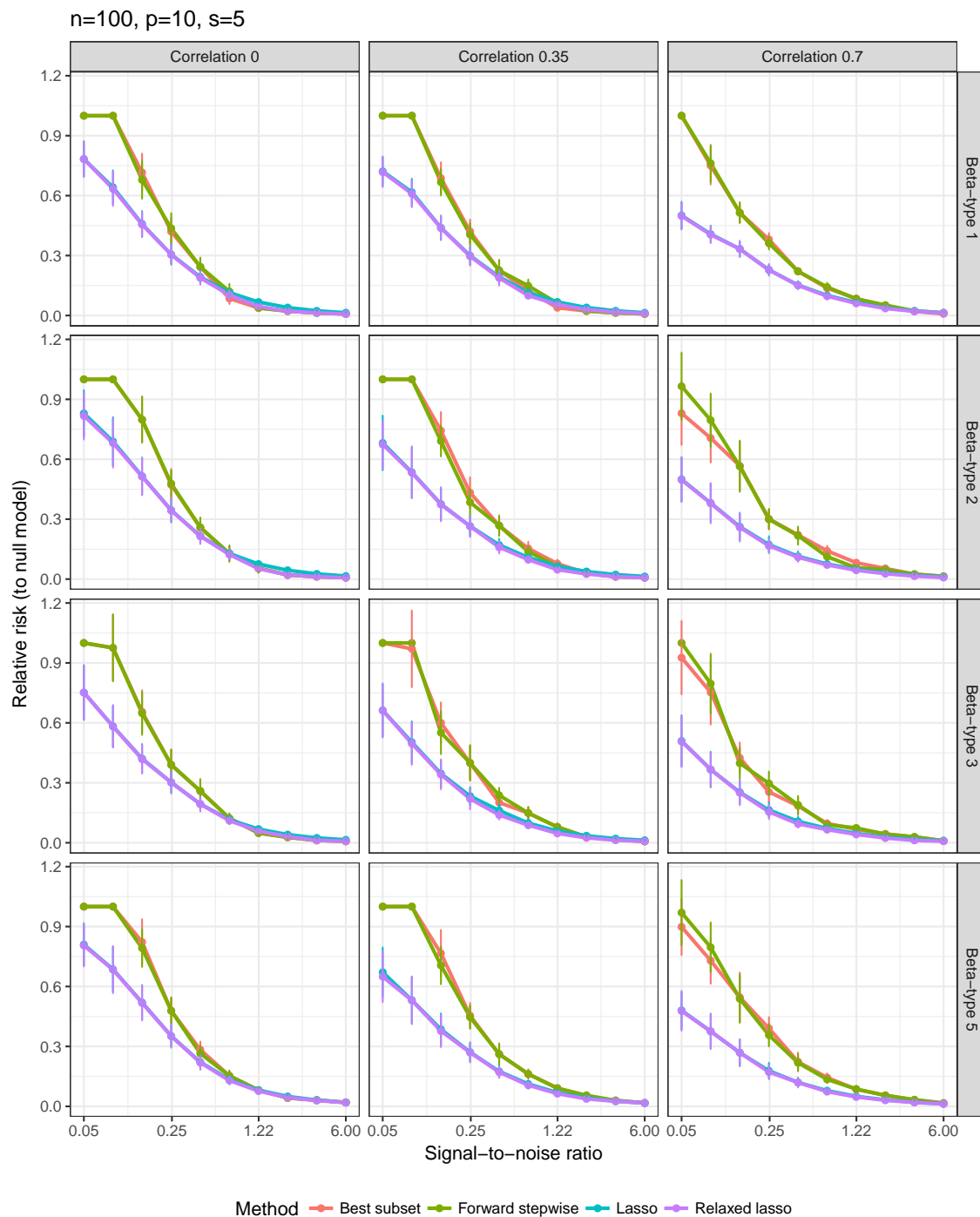
A.4.4 Number of nonzero coefficients



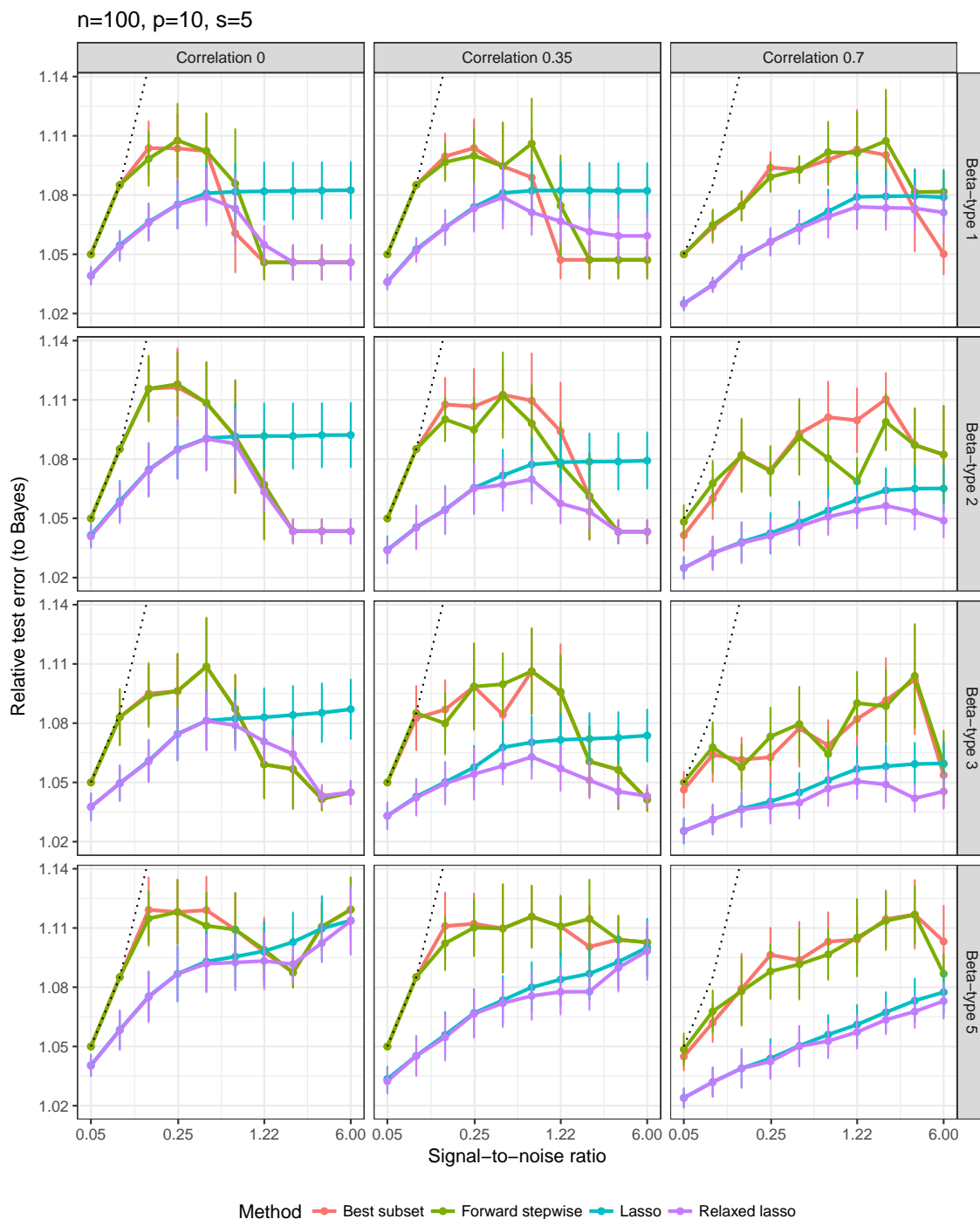
B Oracle tuning

B.1 Low setting: $n = 100, p = 10, s = 5$

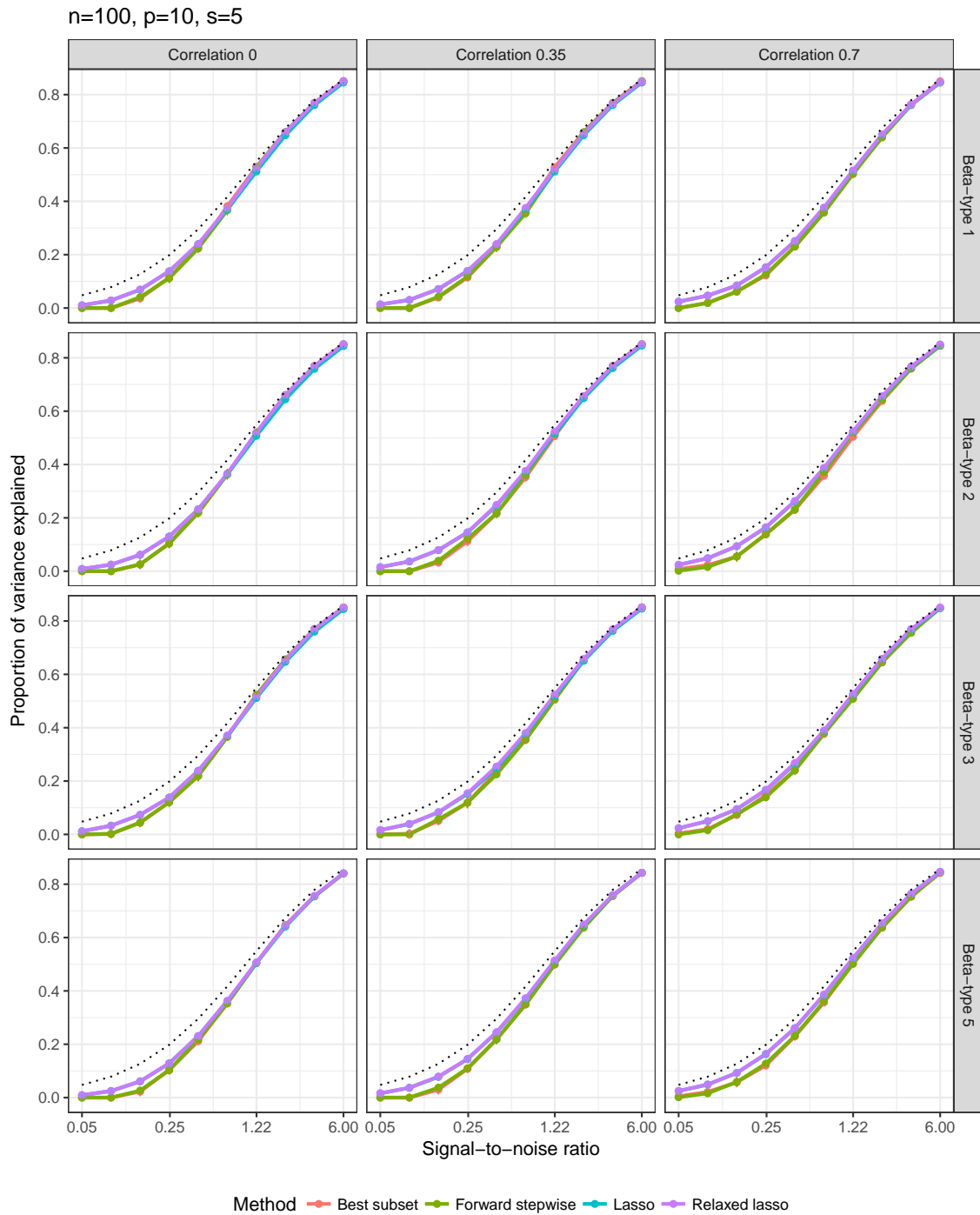
B.1.1 Relative risk (to null model)



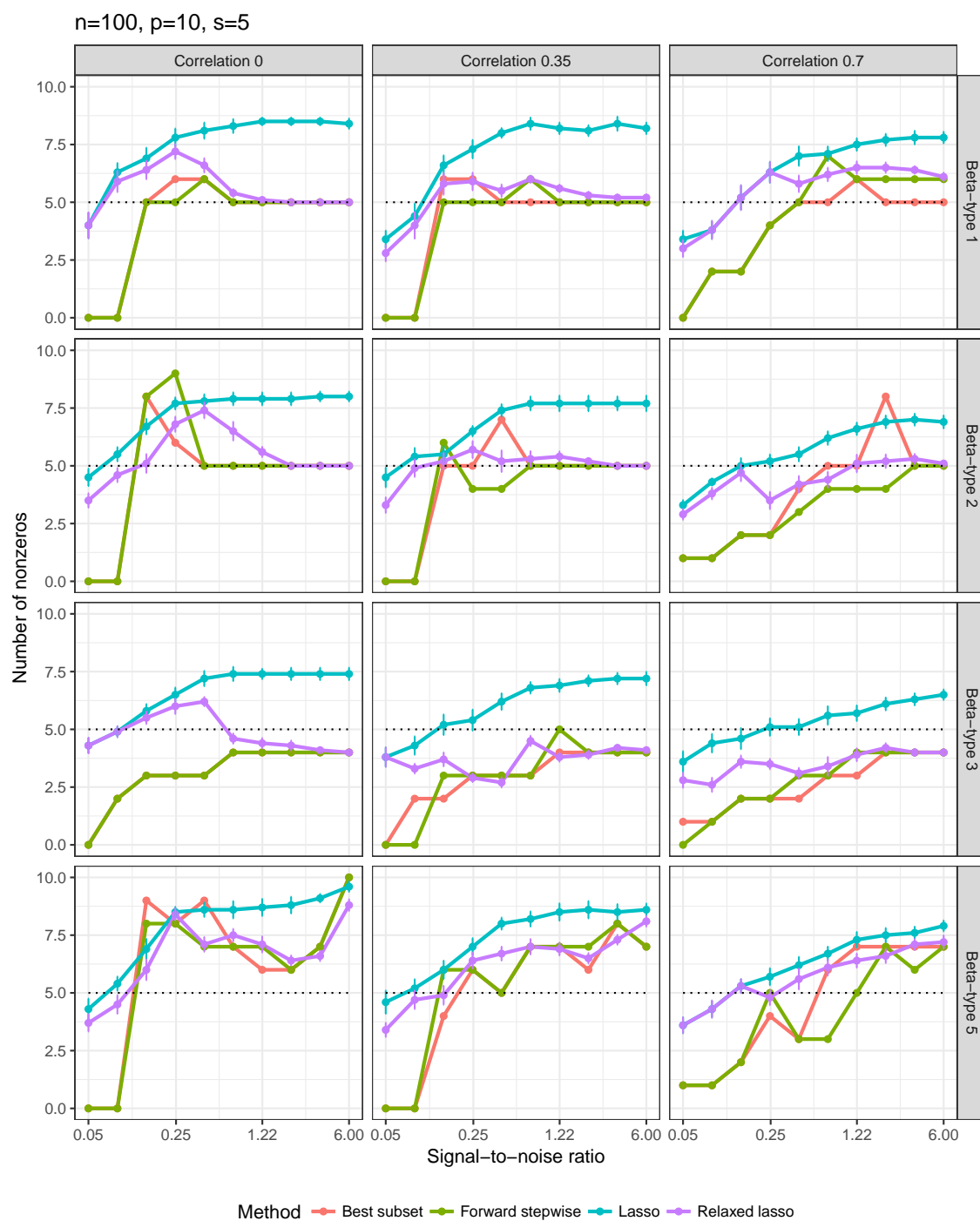
B.1.2 Relative test error (to Bayes)



B.1.3 Proportion of variance explained

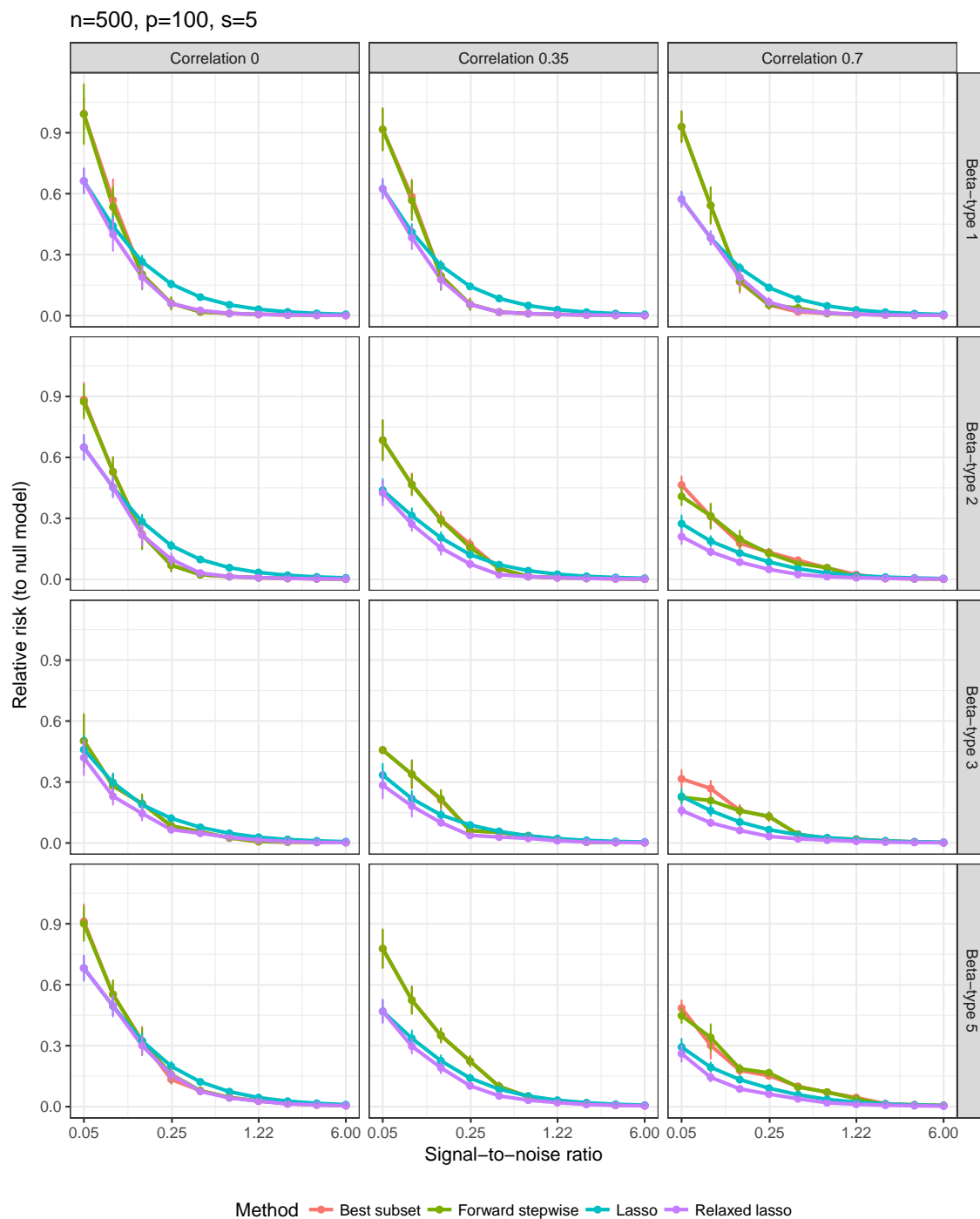


B.1.4 Number of nonzero coefficients

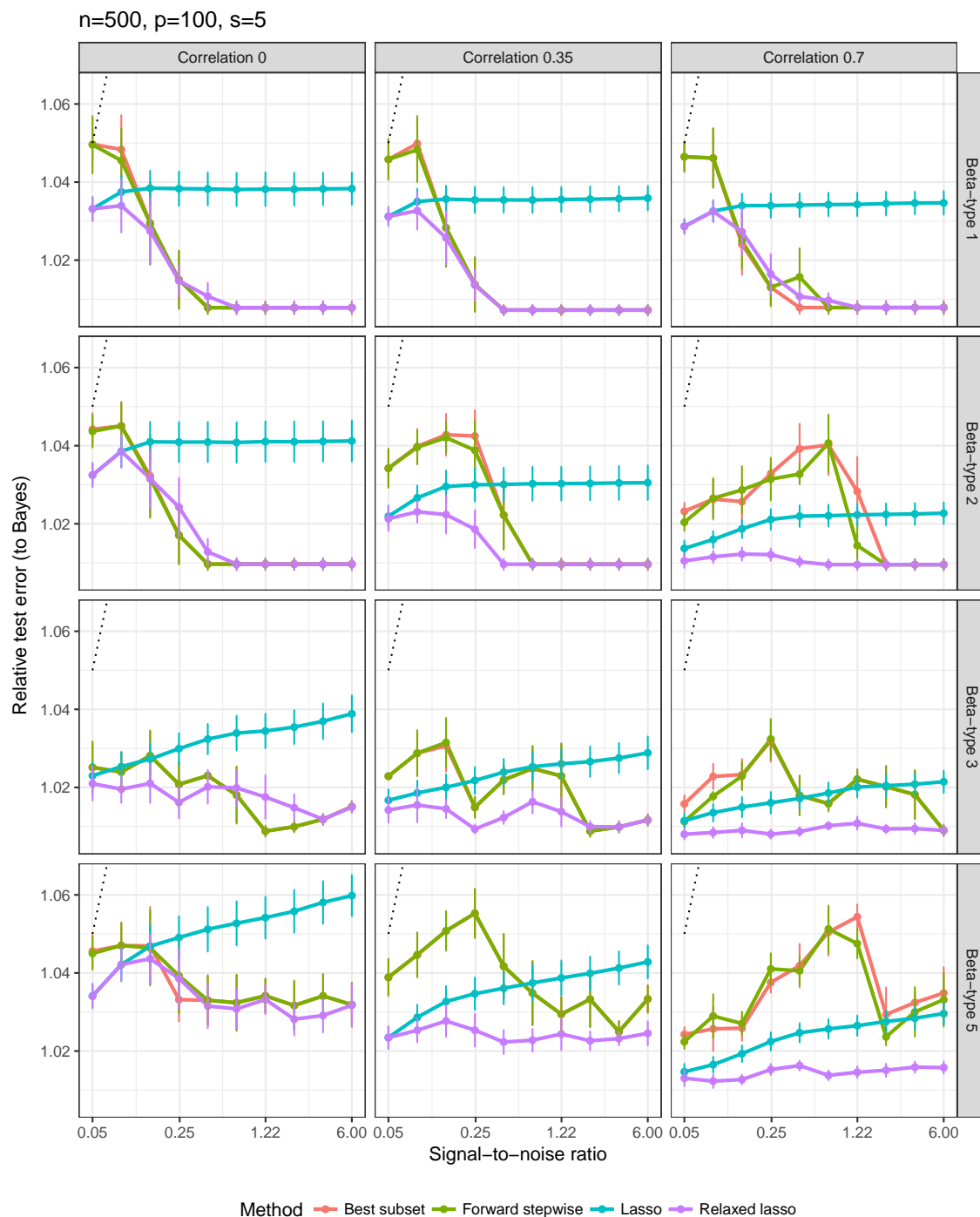


B.2 Medium setting: $n = 500, p = 100, s = 5$

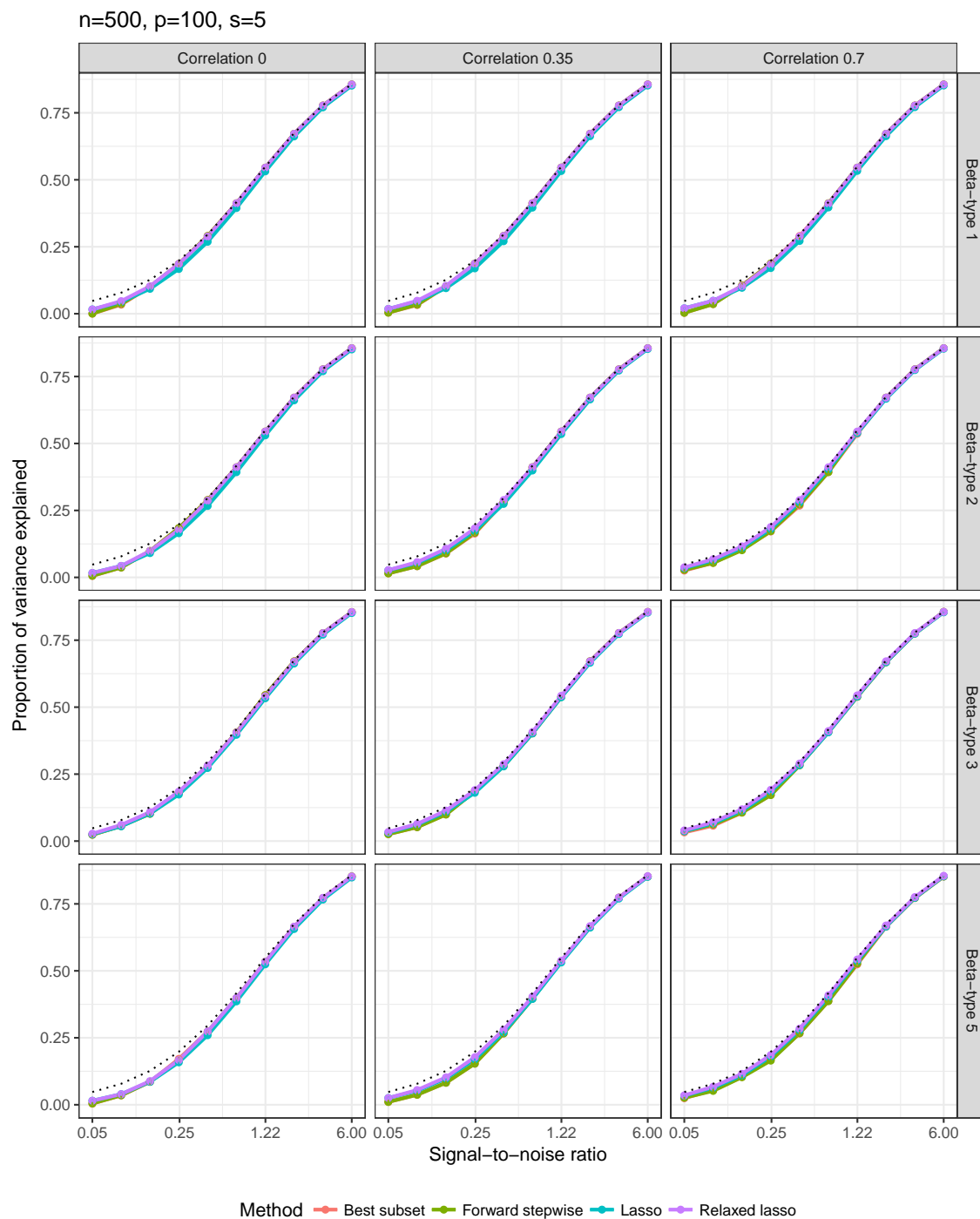
B.2.1 Relative risk (to null model)



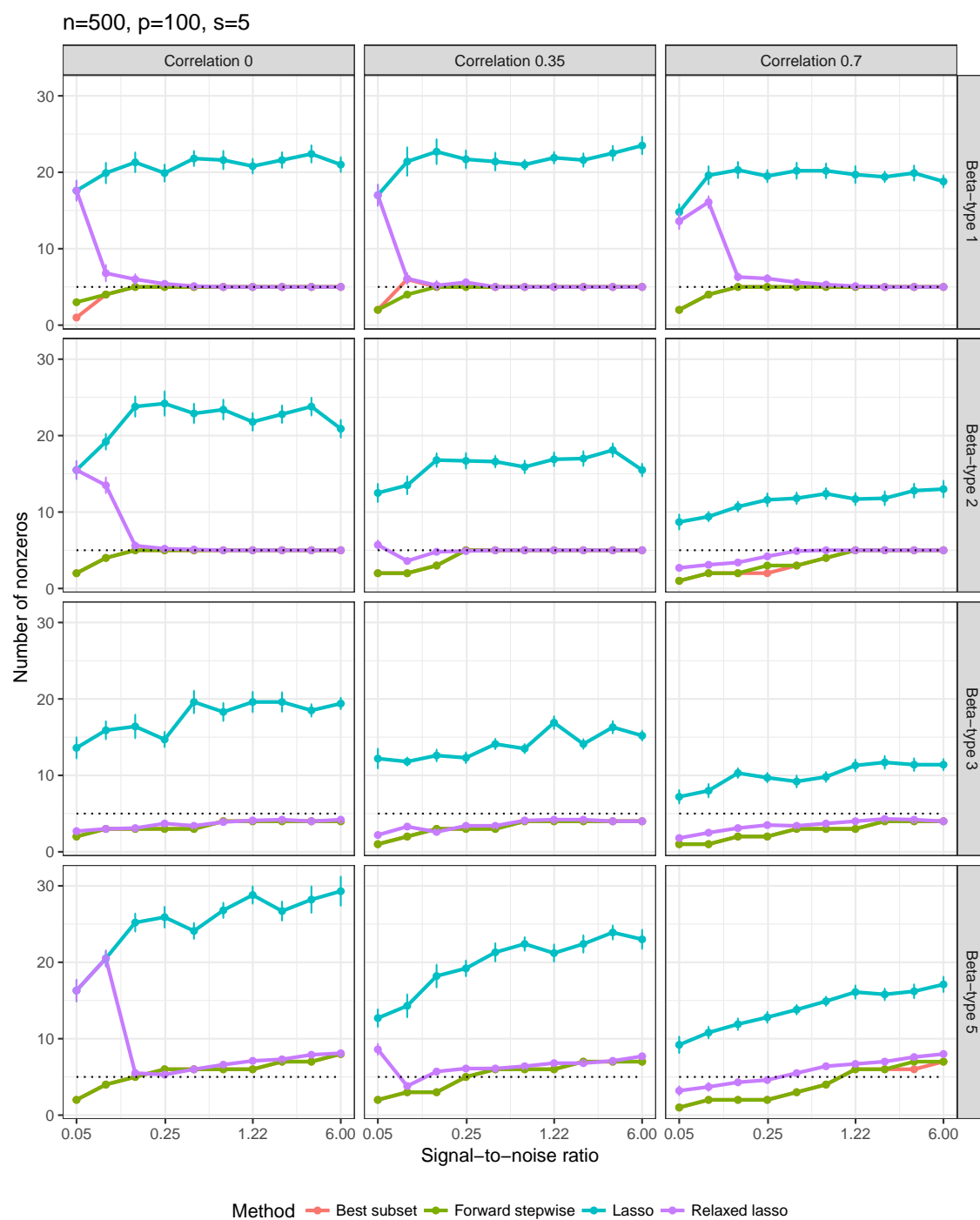
B.2.2 Relative test error (to Bayes)



B.2.3 Proportion of variance explained

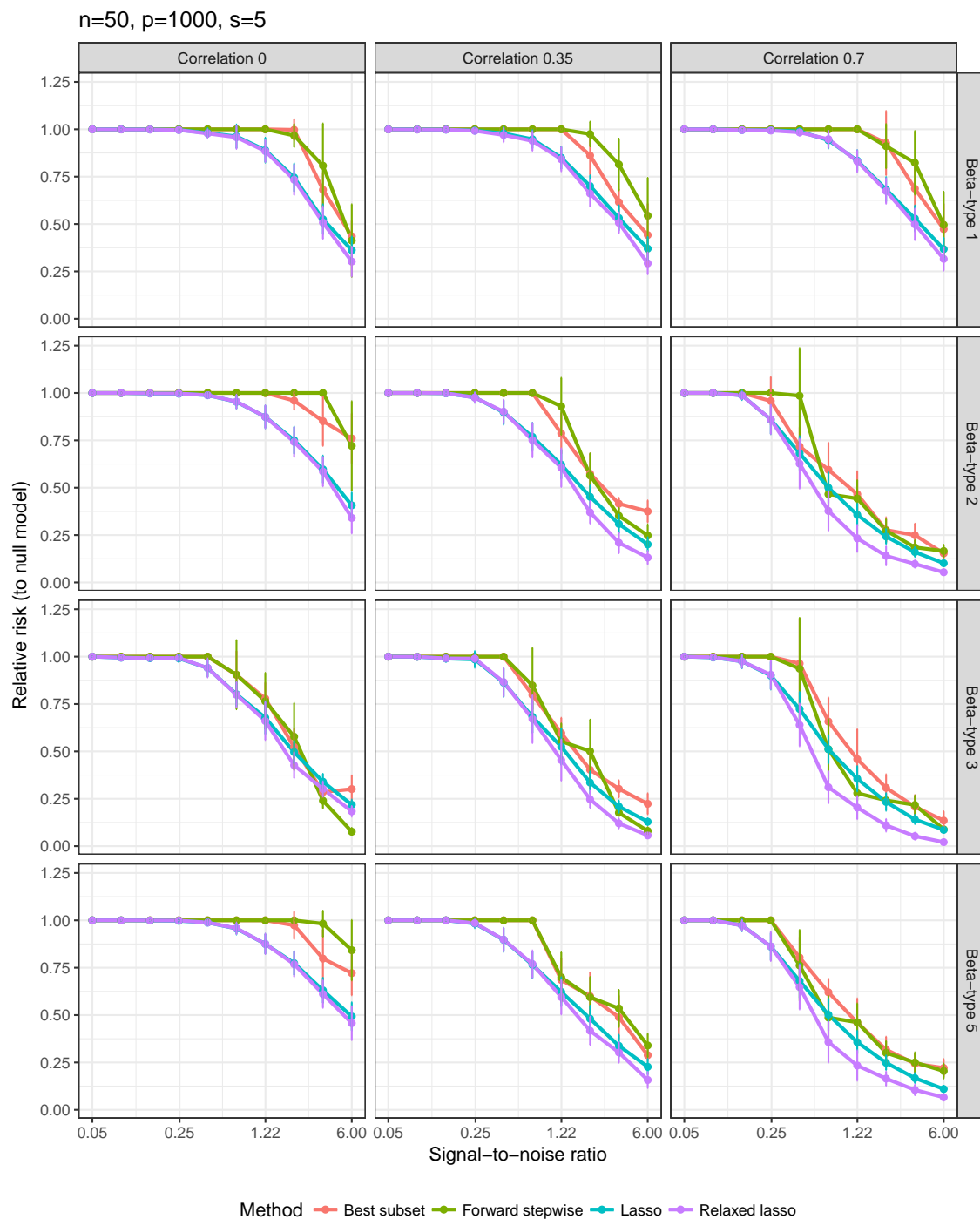


B.2.4 Number of nonzero coefficients

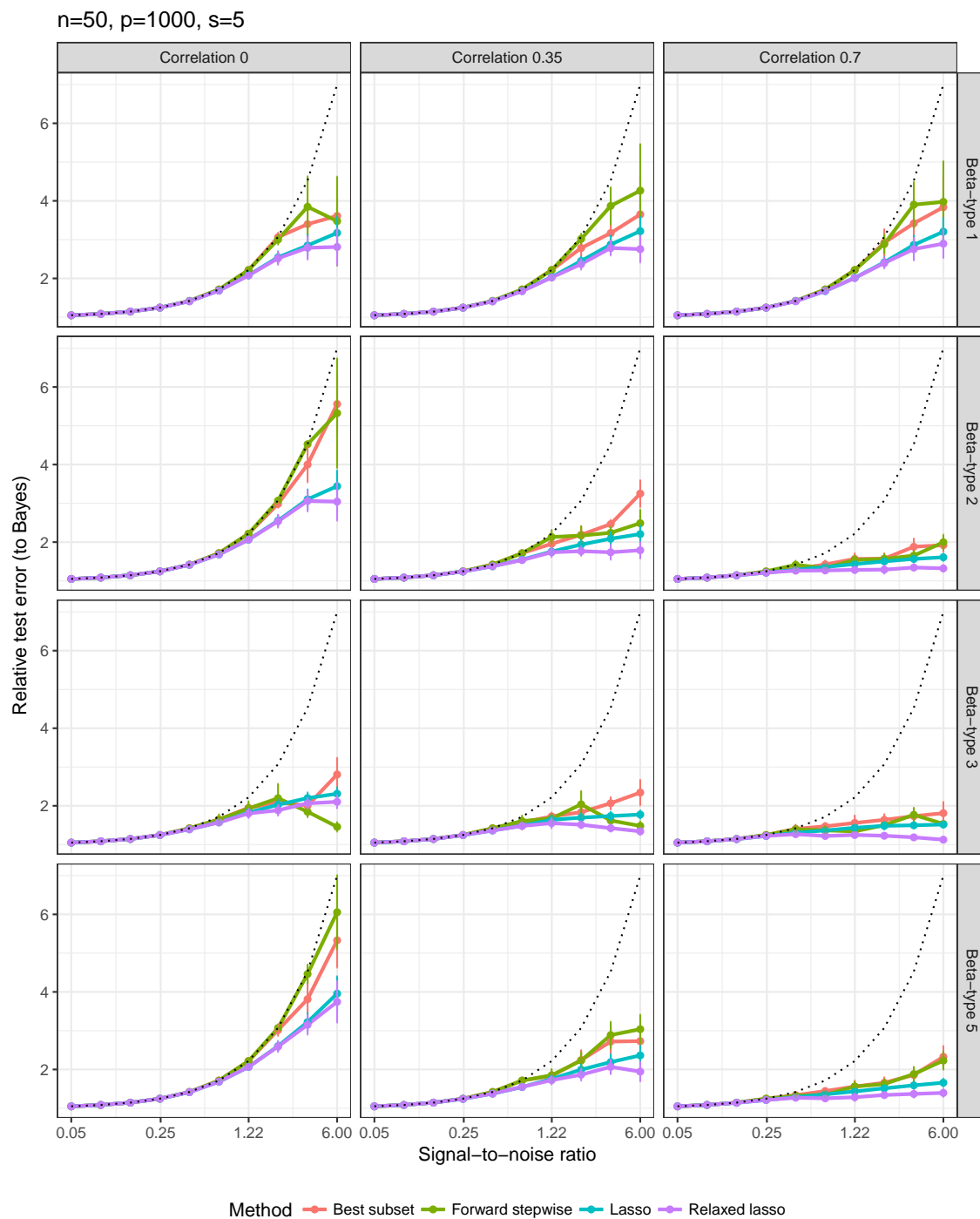


B.3 High-5 setting: $n = 50, p = 1000, s = 5$

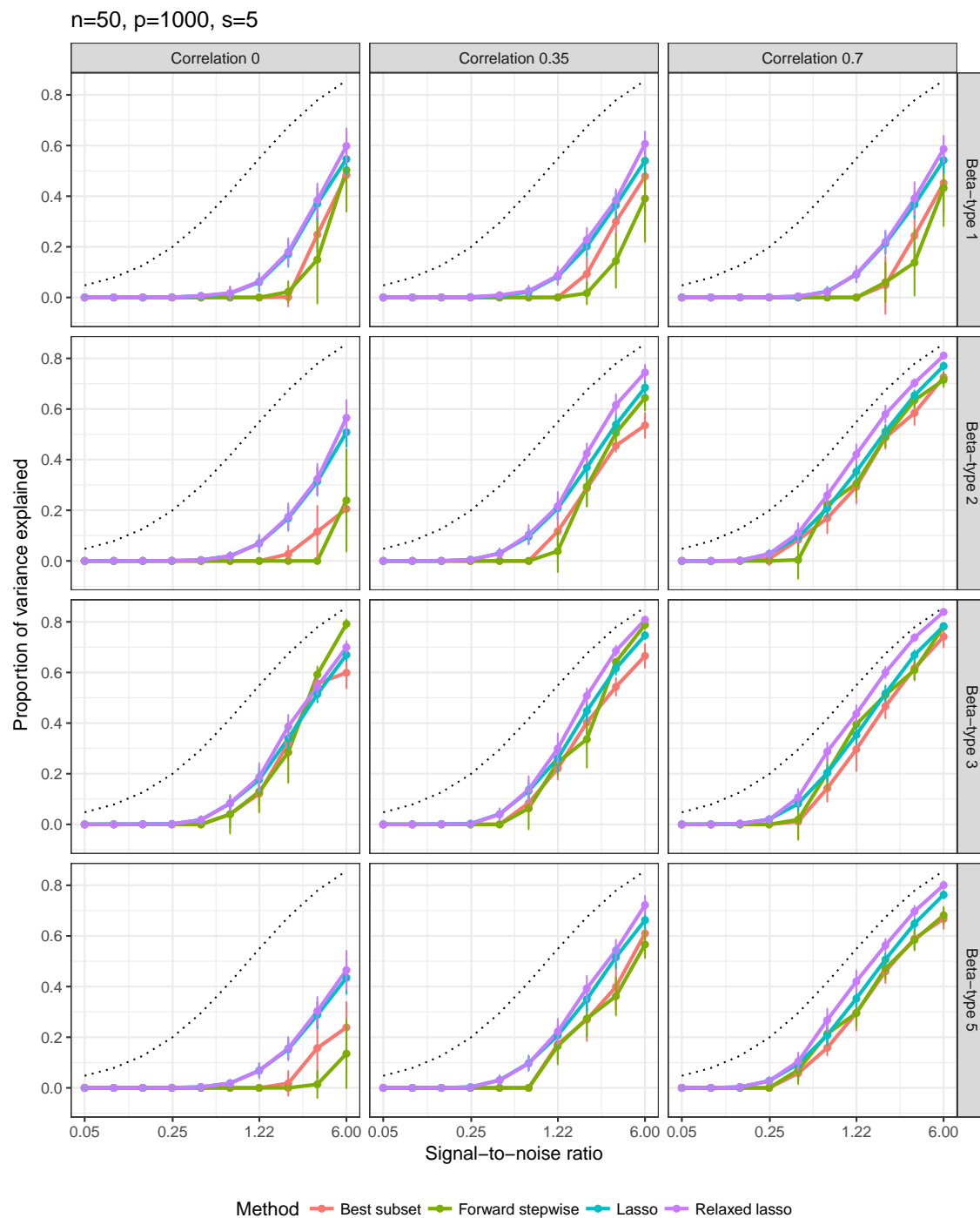
B.3.1 Relative risk (to null model)



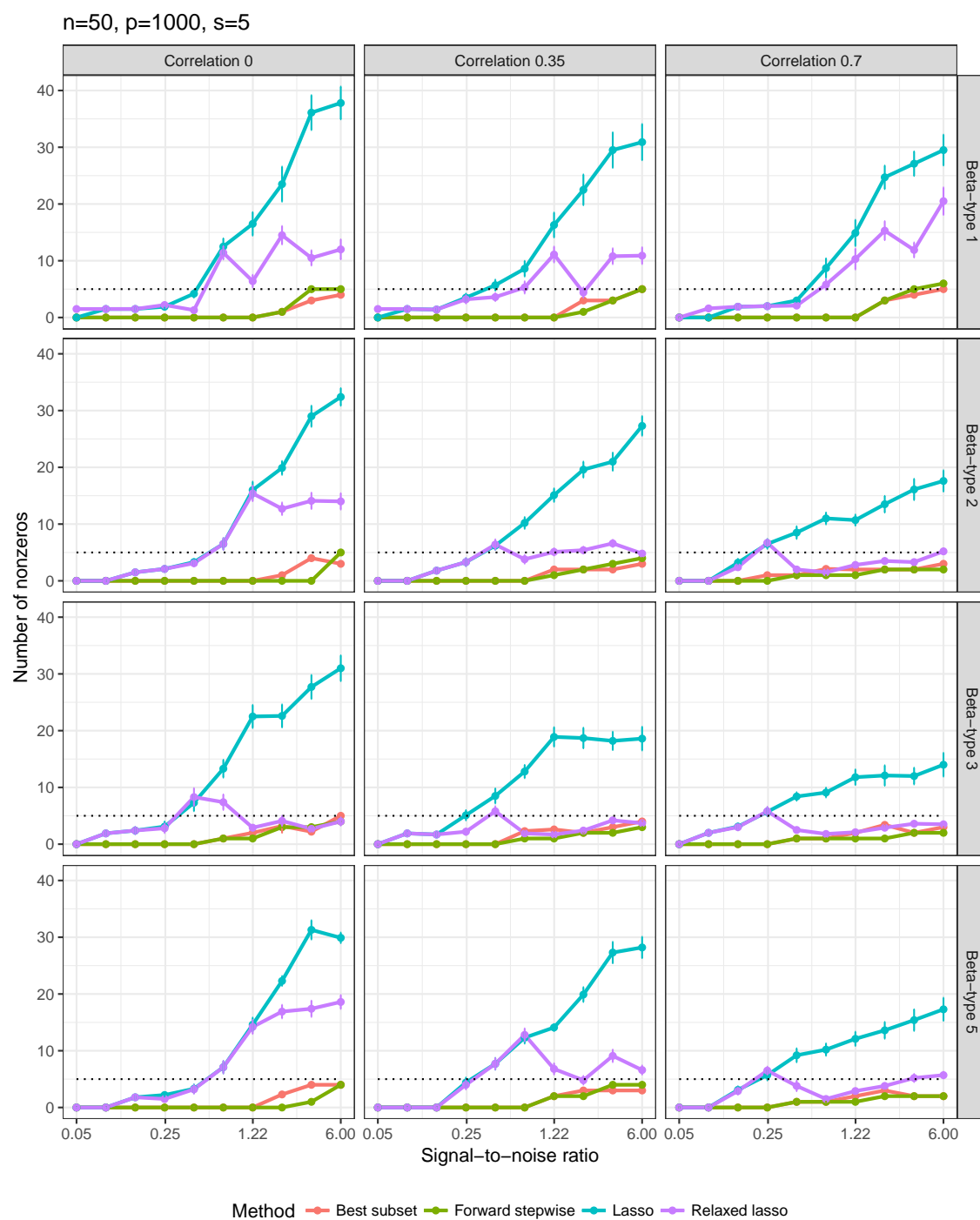
B.3.2 Relative test error (to Bayes)



B.3.3 Proportion of variance explained

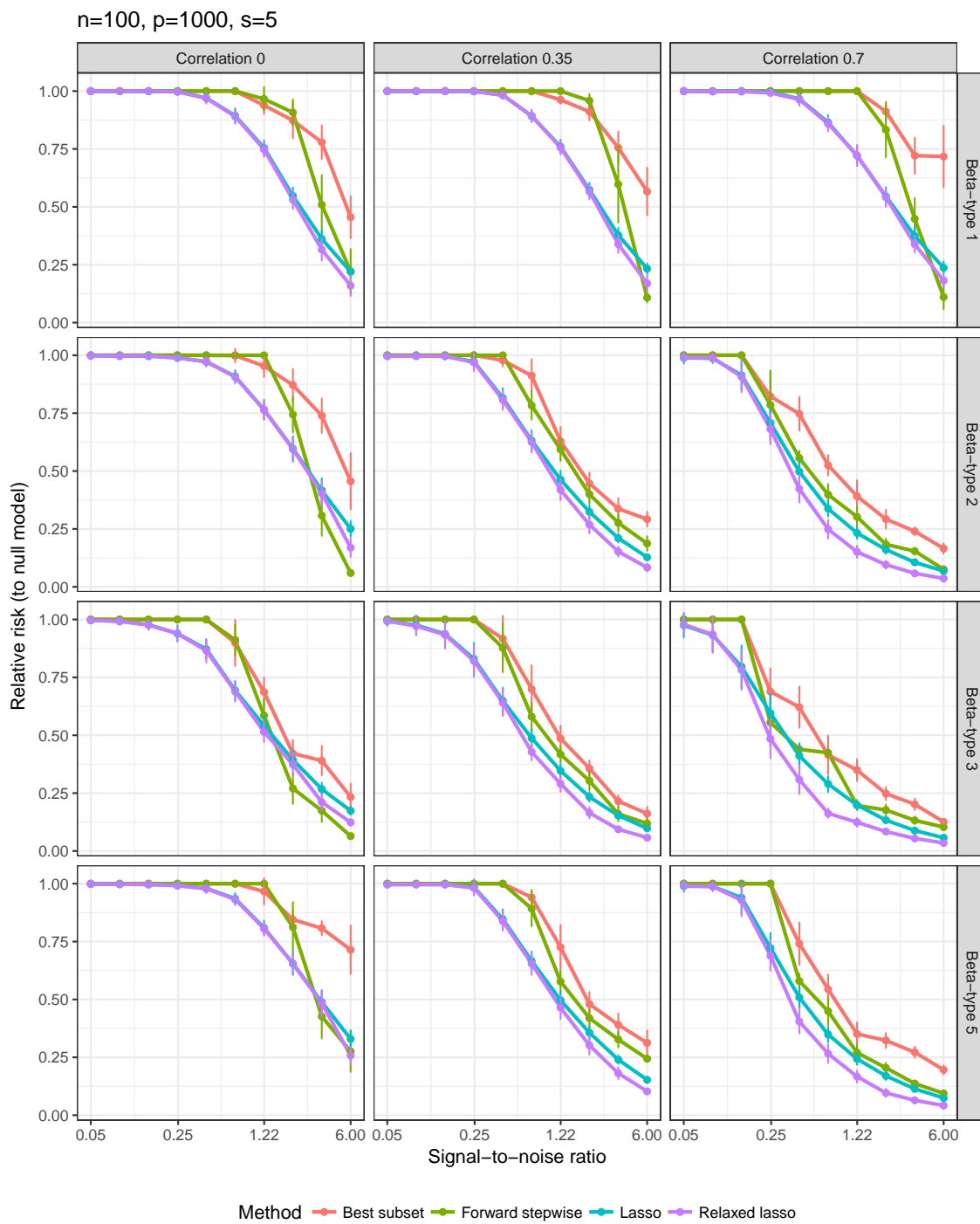


B.3.4 Number of nonzero coefficients

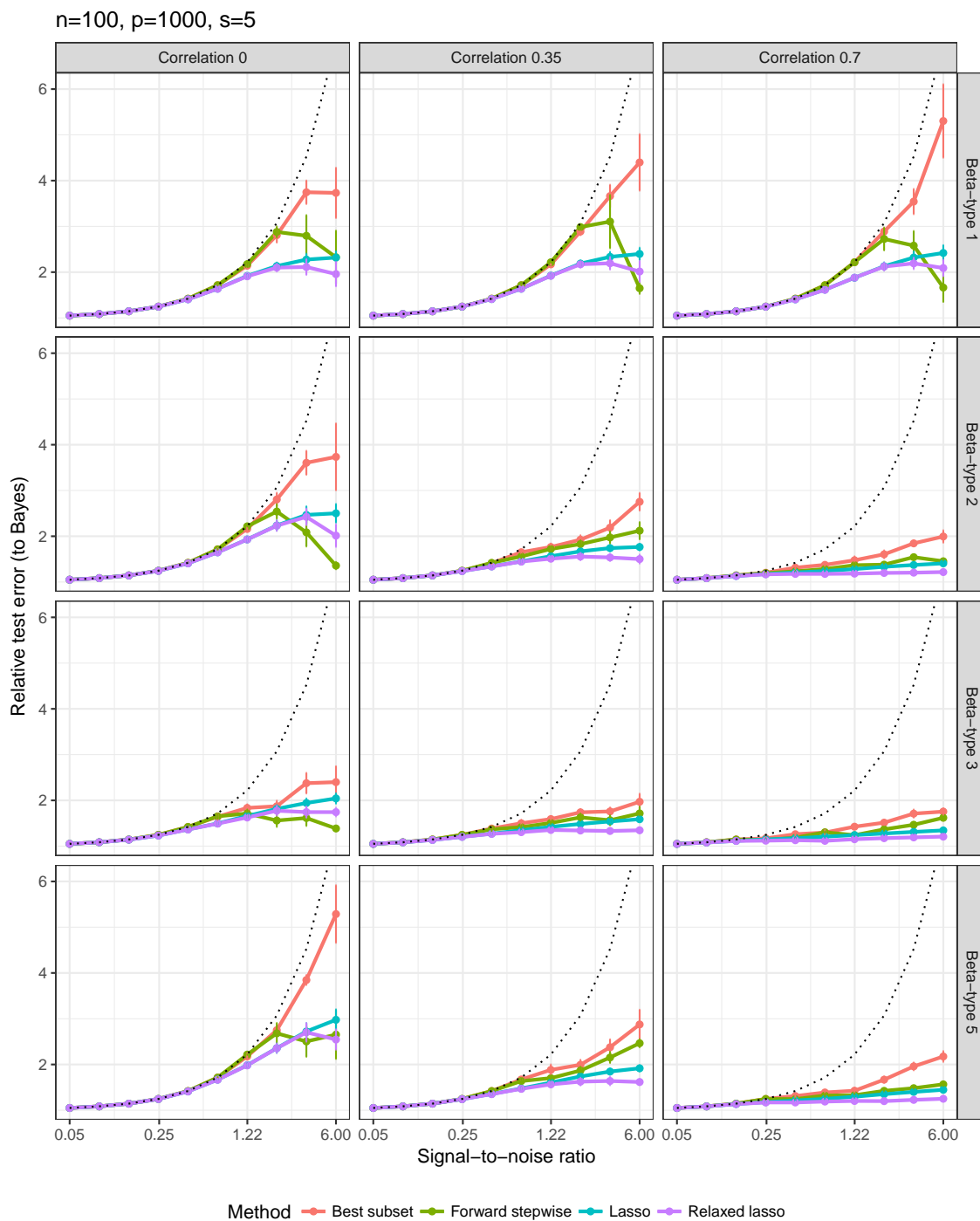


B.4 High-10 setting: $n = 100, p = 1000, s = 10$

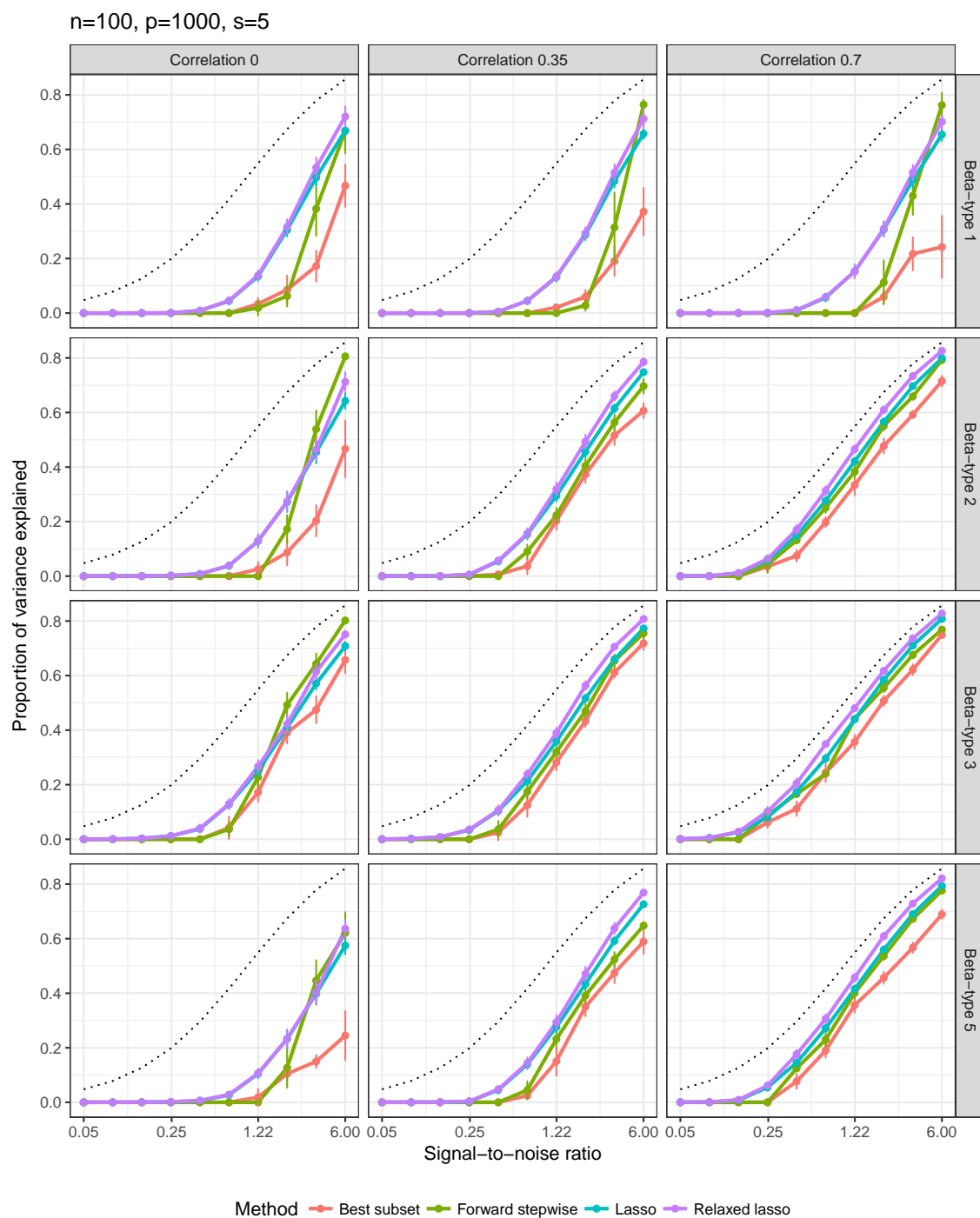
B.4.1 Relative risk (to null model)



B.4.2 Relative test error (to Bayes)



B.4.3 Proportion of variance explained



B.4.4 Number of nonzero coefficients

