

Appendix

Appendix A: Further Simulation Results

A.1 Further Simulation Results for Linear Base-learners

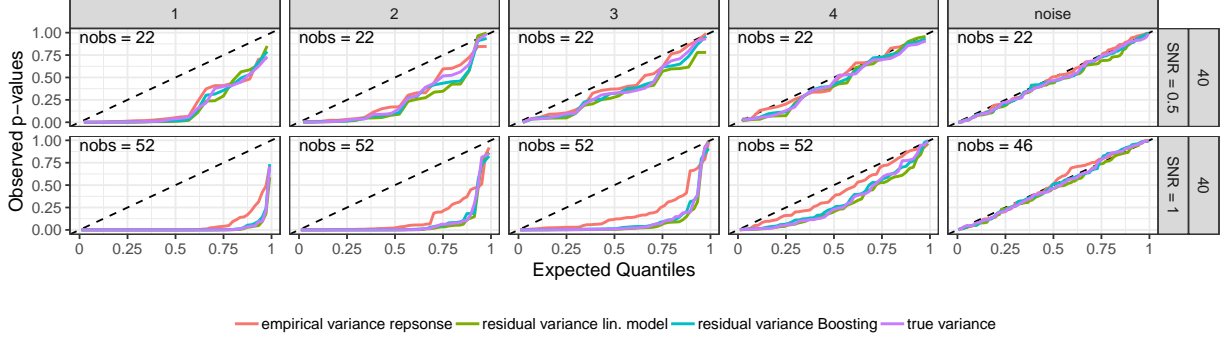


Figure 1: Observed p-values vs. expected quantiles across different covariates (columns) as well as different SNR (rows) for boosting with different variance values / estimates (colours), 26 variables including 22 noise variables, $B = 1000$, a total of 100 simulation iterations and $n = 100$ (in contrast to $n = 25$ in the main article). p-values are shown for simulation iterations, in which either the true model or a model larger than the true model is selected. For each setting, the number of iterations (nobs) is noted in the left upper corner.

A.2 Further Simulation Results for P-spline Base-learners

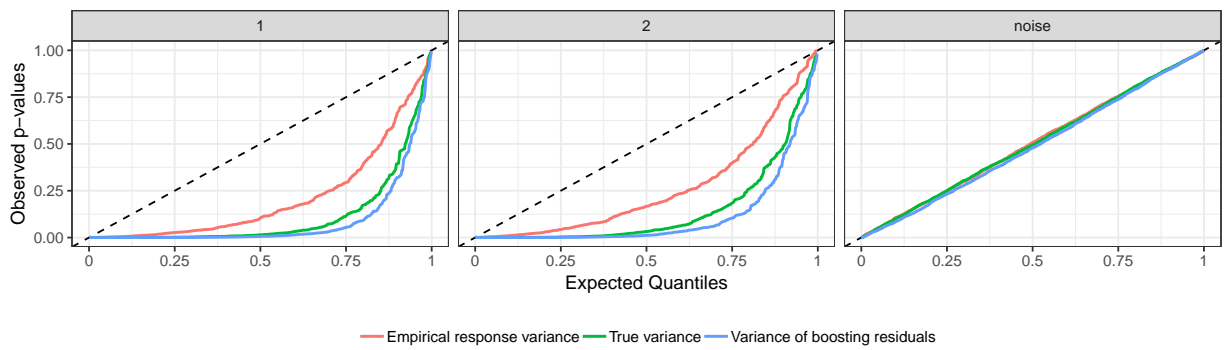


Figure 2: Observed p-values vs. expected quantiles across different covariates (rows) as well as different variance values / estimates (columns) for two different SNR (colours) for testing boosted P-spline baselearners after 50 iterations and a step-length of 0.1, using a total of 500 simulation iterations. p-values are shown for simulation iterations, in which either the true model or a model larger than the true model is selected. All plots are based on 500 simulation iterations as the selection procedure always selected a model with both truly non-linear effects and further noise variables.

Appendix B: Further Application Results

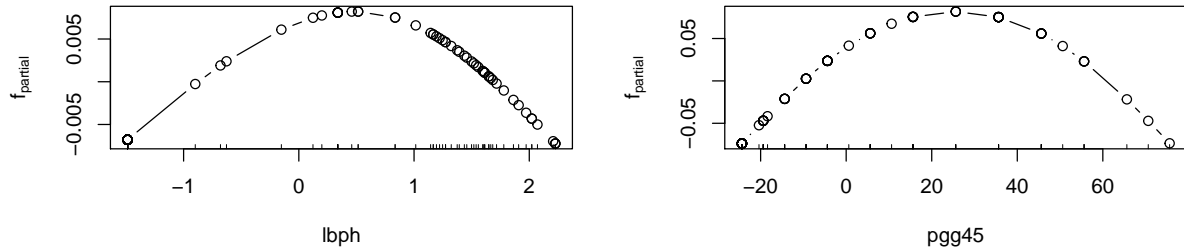


Figure 3: Partial effects of estimated non-linear deviations for the covariates $lbph$ and $pgg45$.

Appendix C: Simulation Code

The R-code and link to the software to reproduce simulation application results can be found at https://github.com/davidruegamer/inference_boosting.