Supplementary material part A: Simulation study

Frank Schaarschmidt

March 27, 2015

1 Parameter settings

1.1 Models

Two types of models were investigated: 1) The model

$$y_{ij} = \alpha_i + \beta_i x_{ij} + e_{ij}, \quad e_{ij} \sim N(0, \sigma^2),$$

allowing treatment specific intercepts and slopes is referred to as **lin.reg**. The values of x_{ij} were drawn from a uniform distribution $U(10, 30)$, parameter settings shown in Table 1 for $I = 3$ groups and Table 2 for $I = 6$ groups were used with an residual variance $e_{ij} \sim N(0, \sigma^2 = 2)$ to simulate the y_{ij}s.

Table 1: Parameter settings for intercepts and slopes in the lin.reg. model for the simulations with $I = 3$ treatment groups.

α_1	α_2	α_3	β_1	β_2	β_3
20	20	20	0.0	0.0	0.0
20	20	20	1.0	1.0	2.0
20	20	20	1.0	2.0	0.5
20	20	25	2.0	2.0	2.0
20	23	15	1.0	2.0	3.0

Table 2: Parameter settings for intercepts and slopes in the lin.reg. model for the simulations with $I = 6$ treatment groups.

α_1	α_2	α_3	α_4	α_5	α_6	β_1	β_2	β_3	β_4	β_5	β_6
20	20	20	20	20	20	0.0	0.0	0.0	0.0	0.0	0.0
20	20	20	20	20	20	1.0	1.0	1.0	1.0	2.0	0.5
20	20	20	20	20	20	1.0	2.0	0.5	0.0	-0.5	3.0
20	20	20	20	25	28	3.0	3.0	3.0	3.0	3.0	3.0
20	23	15	18	21	30	1.0	2.0	3.0	-0.2	-0.5	0.0

2) The model

$$y_{ij} = \alpha_i + \beta_{1i} x_{ij} + \beta_{2i} x_{ij}^2 + e_{ij},$$

allowing treatment specific intercepts, slopes, and also treatment specific curvatures β_{2i}, i.e. allowing treatment-interaction with the linear and quadratic term, is referred to as **quad.reg** in the results. The values of x_{ij} were drawn from a uniform distribution $U(0, 10)$, parameter settings shown in Table 3 for $I = 3$ groups and Table 4 for $I = 6$ groups were used with an residual variance $e_{ij} \sim N(0, \sigma^2 = 3)$ to simulate the y_{ij}s.

Table 3: Parameter settings for intercepts and slopes in the quad.reg. model for the simulations with $I = 3$ treatment groups.

α_1	α_2	α_3	β_{11}	β_{12}	β_{13}	β_{21}	β_{22}	β_{23}
80	80	80	10	10	10	-0.5	-0.5	-0.5
80	90	80	10	10	10	-0.5	-0.5	-0.5
20	20	20	1	1	3	0.3	0.3	0.3
80	80	80	5	10	20	0	-0.5	-1
20	20	20	5	10	15	0	-0.5	-1.2
20	20	20	5	5	-5	0.25	0.25	1.5

Table 4: Parameter settings for intercepts and slopes in the quad.reg. model for the simulations with $I = 6$ treatment groups.

α_1	α_2	α_3	α_4	α_5	α_6	β_{11}	β_{12}	β_{13}	β_{14}	β_{15}	β_{16}	β_{21}	β_{22}	β_{23}	β_{24}	β_{25}	β_{26}
80	80	80	80	80	80	10	10	10	10	10	10	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
80	90	80	80	80	80	10	10	10	10	10	10	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
20	20	20	20	20	20	1	1	3	1	1	1	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
80	80	80	80	80	80	5	5	10	10	20	20	0	0	-0.5	-0.5	-1	-1
20	20	20	20	20	20	5	5	10	10	15	15	0	0	-0.5	-0.5	-1.2	-1.2
20	20	20	20	20	20	5	5	5	5	-5	-5	0.25	0.25	0.25	0.25	1.5	1.5

All models and parameter settings were simulated for equal sample sizes per treatment group $n_i = 5, 10, 20, 100$, and for the unbalanced sample size settings with $n_1 = 5$, all remaining $n_i = 10$ and $n_1 = 10$, all remaining $n_i = 20$.

1.2 Between-treatment contrasts and number of covariate values Q

Three basic types of between-treatment comparisons were used in the simulations, exemplified below for differences among $I = 3$ treatment groups. Comparisons to a common control group are referred to as Dunnett-type, all pairwise comparisons are referred to as Tukey-type and comparisons of treatments to an overall mean (weighted by sample size) are referred to as GrandMean.

$$\mathbf{C}_{Dunnett} = \begin{pmatrix} \text{-1} & 1 & 0 \\ \text{-1} & 0 & 1 \end{pmatrix}, \mathbf{C}_{Tukey} = \begin{pmatrix} \text{-1} & 1 & 0 \\ \text{-1} & 0 & 1 \\ 0 & \text{-1} & 1 \end{pmatrix}, \mathbf{C}_{GrandMean} = \begin{pmatrix} 0.67 & -0.33 & -0.33 \\ -0.33 & 0.67 & -0.33 \\ -0.33 & -0.33 & 0.67 \end{pmatrix}$$

For the model referred to as lin.reg and differences of regression lines, all three contrast types were simulated for $I = 3$, and $I = 6$ treatment groups, results are shown in the main paper, Figure 10. For ratios of regression lines in lin.reg and the quad.reg model, involving interactions with the linear and/or quadratic terms, only the Dunnett-type comparisons were simulated for $I = 3, 6$, for treatment comparisons in terms of differences as well as ratios (overview shown in Figure 11 of the main paper).

In all simulation settings described so far, the covariate grid $\tilde{\mathbf{x}}$ of interest was constructed at $Q = 3$, $Q = 6$, $Q = 10$, and $Q = 20$ equidistant grid points over the predefined range of covariate values.

For each combination of the models, parameters, sample sizes and types of inference described above, 5000 data sets were drawn from the corresponding model and nominal 95% simultaneous confidence intervals were computed, and the simultaneous coverage probability was estimated. Thus, for an exact method, one can expect that such an estimated simultaneous coverage probability will fall with probability 0.95 into the range $[0.944; 0.956]$, based on the sampling error of the binomial distribution $Bin(n = 5000, \pi = 0.95)$. These limits are shown as dotted lines in the figures of the main paper and below.

2 Detailed results of the simulation study for ratios in the lin.reg and for the quad.reg model

For differences between treatment-specific model predictions the multivariate t quantile could not be calculated for up to 6% of the simulated data sets that involved sample size $n_i = 5$ in all or some treatment groups. For ratios of treatment-specific model predictions, the additional problem may occur that the quadratic equation leading to the Fieller-type intervals has no bounded solution, especially when the sample size is low. The proportion of these computational problems among the 5000 simulations for each parameter setting is shown as a colour scale in Figures 1, 2 and 3. For the computation of the simultaneous coverage probabilities, unbounded intervals have been counted as covering the true parameter vector.

For a number of settings, very low coverages probabilities occurred (as low as 0 in some cases), particularly in parameter settings where the methods are obviously inadequate (e.g. using a grid of $Q = 3$ covariate values to approximate confidence bands for treatment differences or ratios that form no straight line depending on the covariate, see Figures 2 and 3). In these cases, coverage probabilities are only depicted down to a lower limit of 0.9 (ratios), and 0.92 (difference). To preserve information on the coverage probabilities omitted from the figures by these restrictions, the minimal coverage probability of the setting is shown as a red number in parentheses, whenever at least one value 0.9 (or 0.92) occurred.

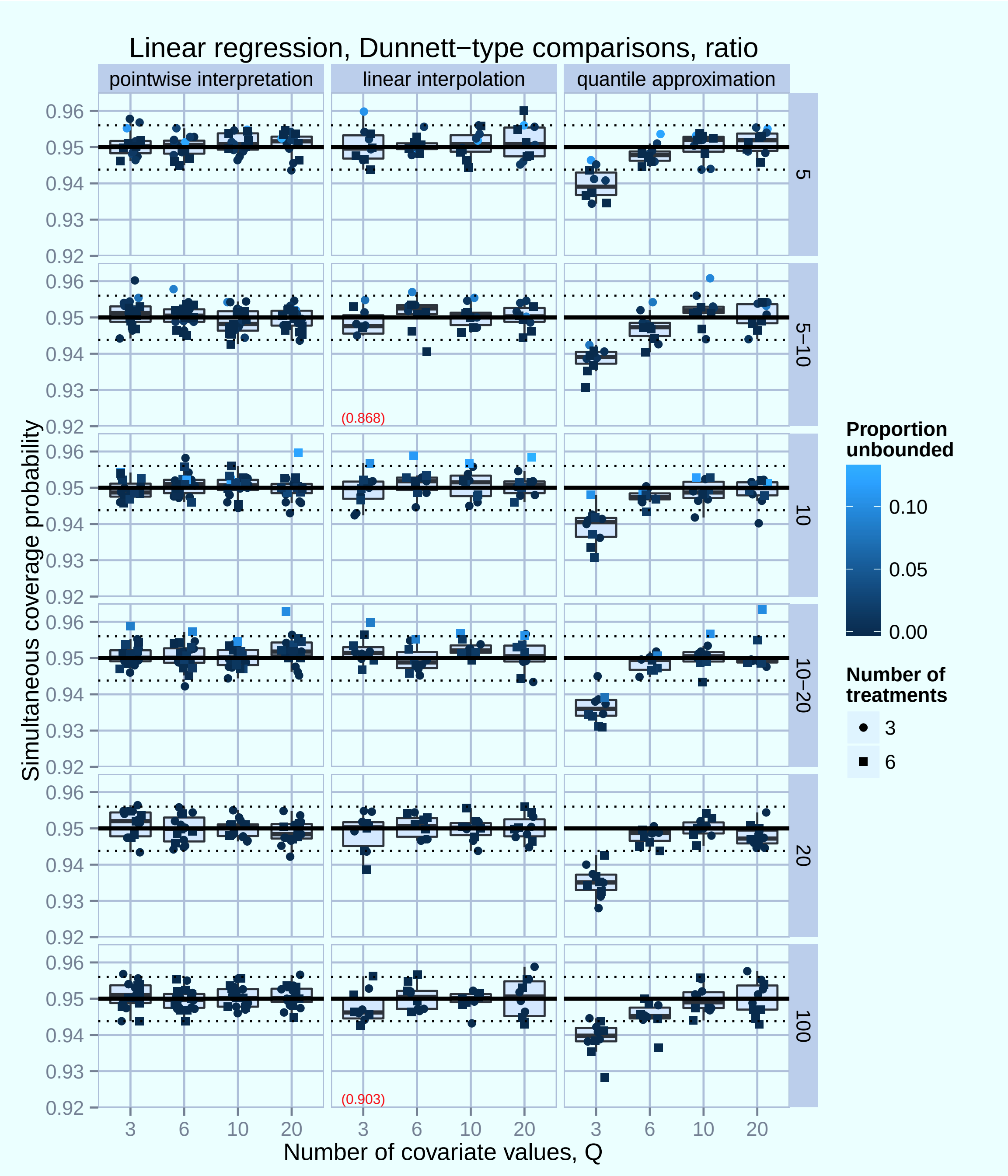

Figure 1: Estimated simultaneous coverage probabilities (5000 simulations) for Dunnett-type comparisons to a control treatment in terms of **ratios** of regression lines (model **lin.reg**). Dotted horizontal lines show the range in which simulated simultaneous coverage probability should fall in 95% of cases, if the applied method has exactly 95% simultaneous confidence level. Settings with different sample size per group are shown in separate rows, columns distinguish simultaneous coverage probability of point wise interpretation from two options to approximate simultaneous confidence bands. Symbols distinguish simulations with $I = 3$ and $I = 6$ groups, the colour scale show the proportion of simulations in which the quadratic equation for the Fieller-type confidence interval had no soulation. Where estimated coverage probabilities below 0.9 occured these are not shown in the plot, the minimal coverage probability is given as a number in parentheses, in red.

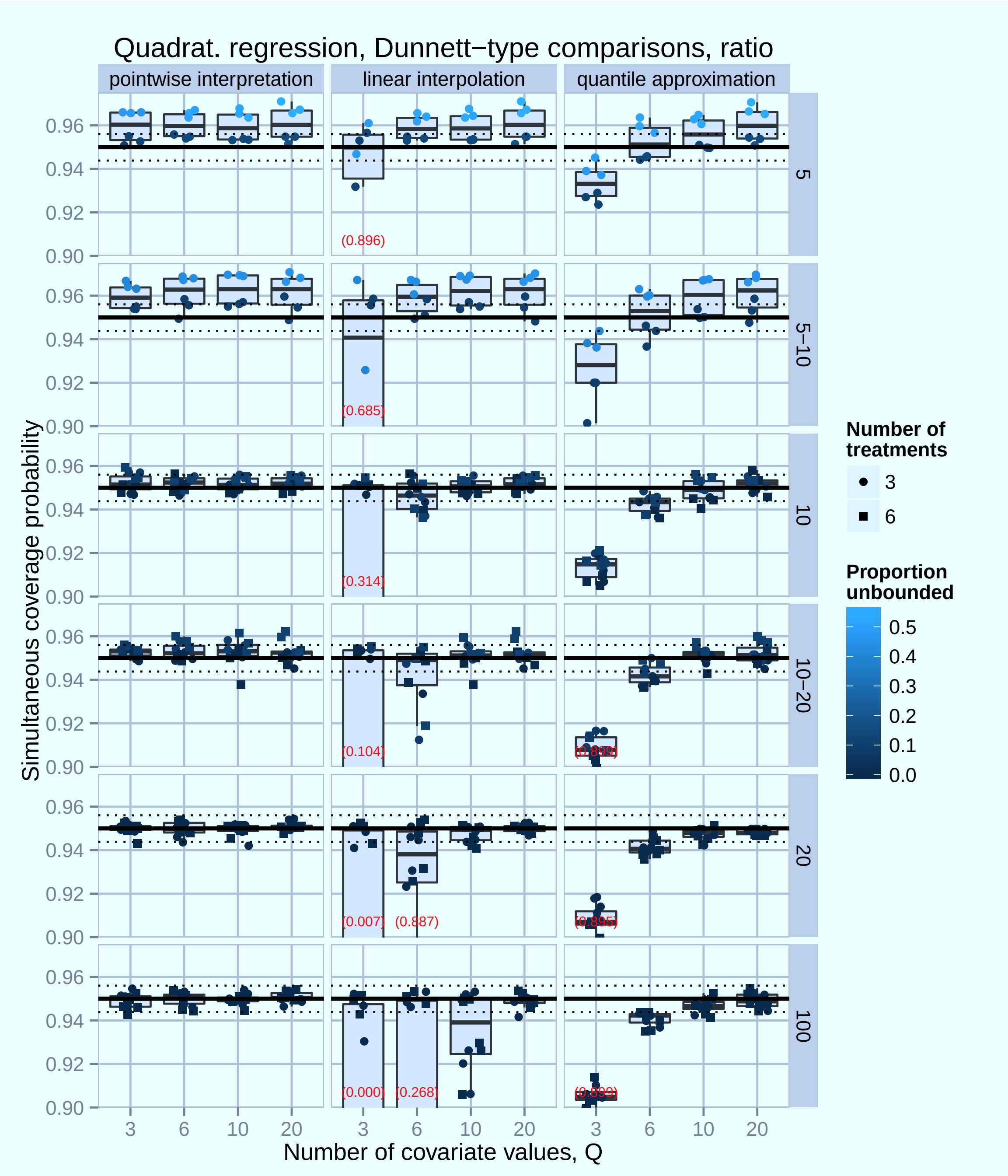

Figure 2: Estimated simultaneous coverage probabilities (5000 simulations) for Dunnett-type comparisons to a control treatment in terms of **ratios** of treatment-specific model predictions in the **quadratic** model (**quad.reg**). Dotted horizontal lines show the range in which simulated simultaneous coverage probability should fall in 95% of cases, if the applied method has exactly 95% simultaneous confidence level. Settings with different sample size per group are shown in separate rows, columns distinguish simultaneous coverage probability of point wise interpretation from two options to approximate simultaneous confidence bands. Symbols distinguish simulations with $I = 3$ and $I = 6$ groups, the colour scale shows the proportion of simulations in which the quadratic equation for the Fieller-type confidence interval had no solution, or the multivariate t-quantile could not be calculated. Where estimated coverage probabilities below 0.9 occured these are not shown in the plot, the minimal coverage probability is given as a number in parentheses, in red.

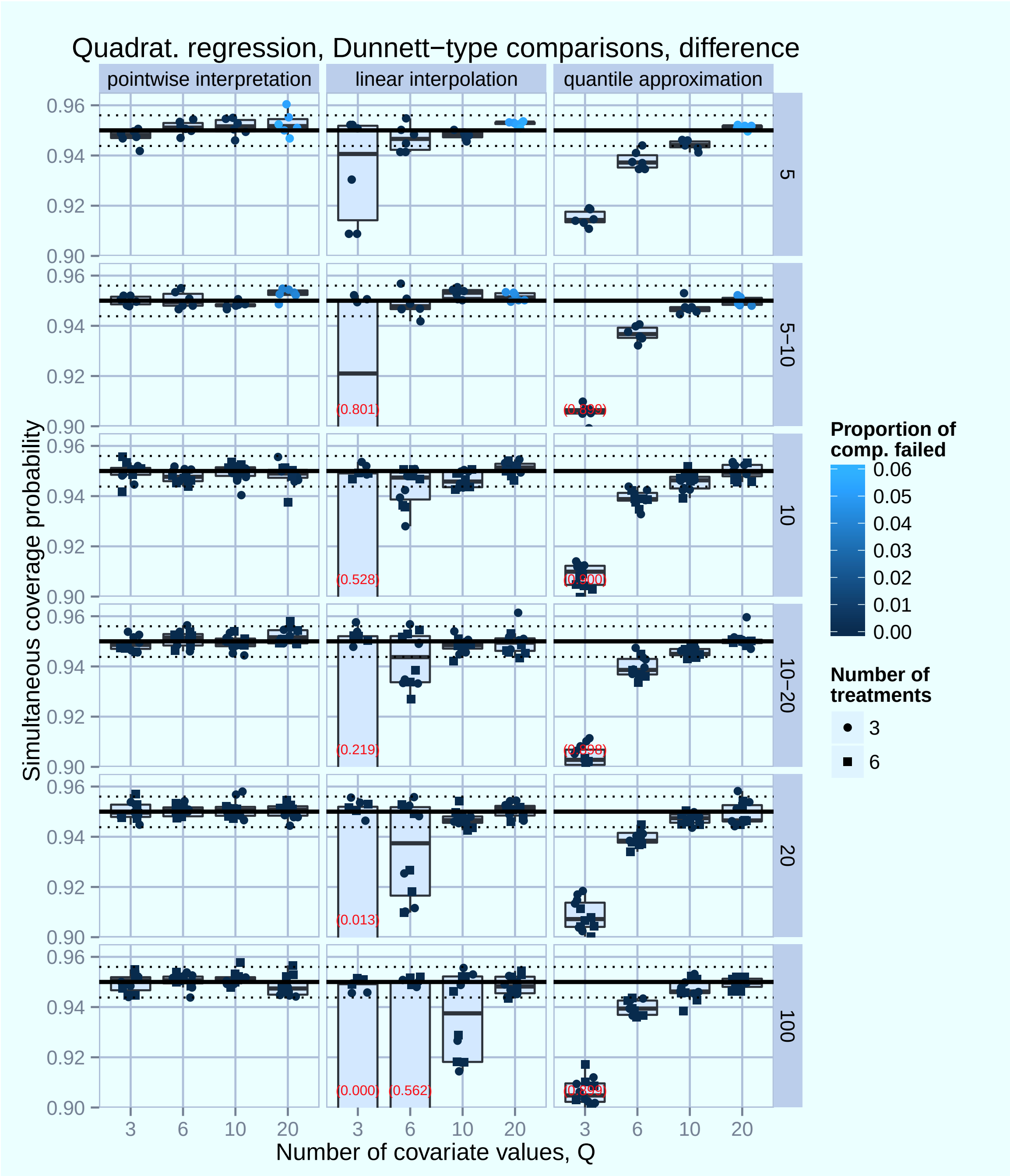

Figure 3: Estimated simultaneous coverage probabilities (5000 simulations) for Dunnett-type comparisons to a control treatment in terms of **differences** of treatment-specific model predictions in the **quadratic** model (**quad.reg**). Dotted horizontal lines show the range in which simulated simultaneous coverage probability should fall in 95% of cases, if the applied method has exactly 95% simultaneous confidence level. Settings with different sample size per group are shown in separate rows, columns distinguish simultaneous coverage probability of point wise interpretation from two options to approximate simultaneous confidence bands. Symbols distinguish simulations with $I = 3$ and $I = 6$ groups, the colour scale show the proportion of simulations in which the multivariate t-quantile could not be calculated. Where estimated coverage probabilities below 0.9 occured these are not shown in the plot, the minimal coverage probability is given as a number in parentheses, in red.