

Supplementary Material for  
Location-adjusted Wald statistic for scalar parameters

Claudia Di Caterina  
[dicaterina@stat.unipd.it](mailto:dicaterina@stat.unipd.it)

Department of Statistical Sciences, University of Padova  
Via Cesare Battisti 241-243, 35121 Padova, Italy

Ioannis Kosmidis  
[i.kosmidis@ucl.ac.uk](mailto:i.kosmidis@ucl.ac.uk)

Department of Statistical Science, University College London  
Gower Street, London, WC1E 6BT, UK

The Alan Turing Institute  
British Library, 96 Euston Road, London NW1 2DB

October 30, 2017

## 1 Further performance evaluation of the adjusted Wald test

### 1.1 Introduction

The following illustrative results allow to make a more thorough comparison between the inferential performance of the location-adjusted Wald statistic  $t^*$  and those of the most relevant competitors. As done throughout the paper, the one-sided Wald statistic and the signed root of the likelihood ratio statistic are denoted by  $t$  and  $r$ , respectively, whereas here  $s$  indicates the signed root of the score statistic. The second-order test based on the signed root of the modified likelihood ratio statistic  $r^*$  serves as a general benchmark of accuracy.

In the figures below, the comparison is in terms of how close the discrepancy between the empirical null  $p$ -values distribution of the two-sided tests and the uniform distribution is to zero. Apart from the studies under the meta-analysis setting where the R package `metaLik` (Guolo et al., 2012) is available, the values of the likelihood ratio statistic and its modification are obtained using the package `likelihoodAsy` (Bellio and Pierce, 2016), which calculates a simulation-based estimate of  $r^*$ . Due to the proximity of  $r$  to zero, computation of  $r^*$  can be subject to numerical instabilities that significantly misguide results. Thus, exploiting the approximately linear relationship between the two statistics (Fraser et al., 2008), whenever  $|r| < 0.15$  we derive the corresponding value of  $r^*$  by cubic spline interpolation. Furthermore, in the sequel, the tables contrast the various tests focusing on the empirical null rejection probabilities recorded for the most common nominal significance levels 1% and 5%.

In the experiments involving simulations for different sample sizes  $n$ , the Monte Carlo samples are generated under the maximum likelihood fit reported in the paper for each example, by replicating a suitable number of times the set of covariates in the original dataset.

## 1.2 Gamma regression for clotting times of blood

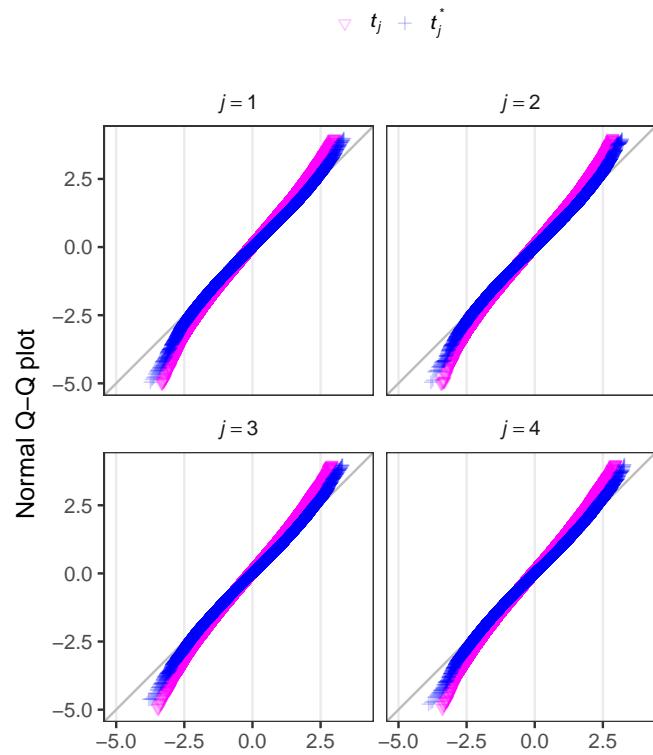


Figure 1: Normal Q-Q plots based on values of  $t_j$  and  $t_j^*$  computed under the hypothesis  $H_0 : \beta_j = \beta_{j0}$  ( $j = 1, \dots, 4$ ) on 50 000 samples generated under the maximum likelihood fit of the gamma regression model for the blood clotting dataset in § 3.3.

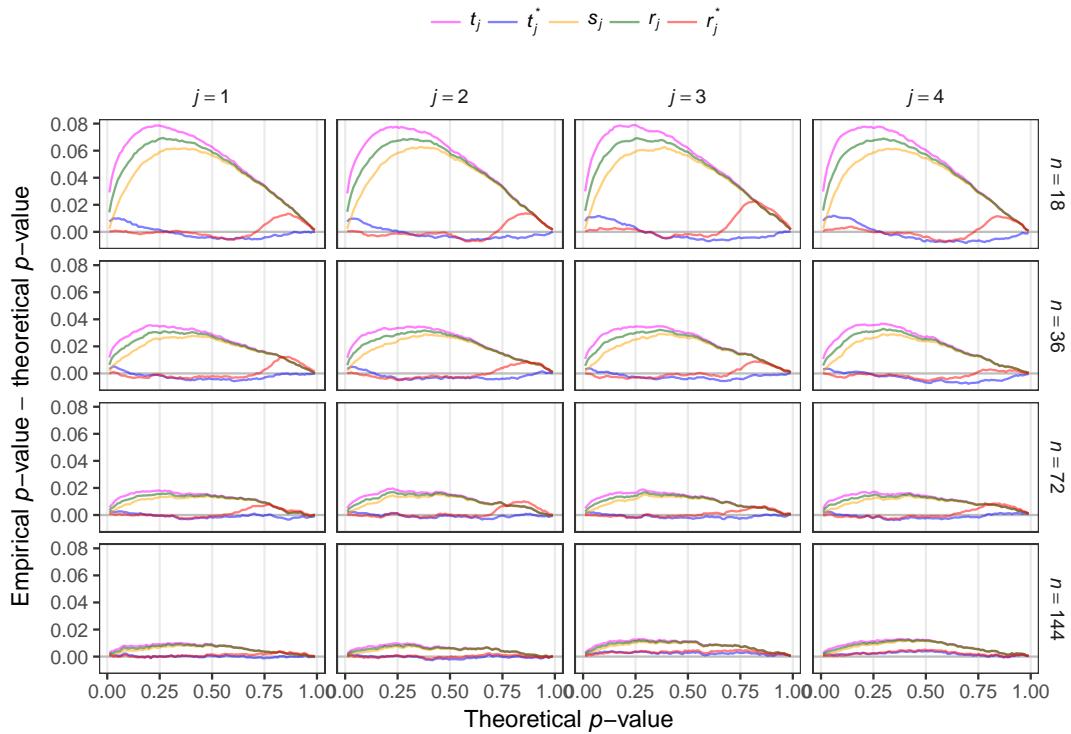


Figure 2: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta_j = \beta_{j0}$  against  $H_1 : \beta_j \neq \beta_{j0}$  based on the normal approximation to the distribution of  $t_j$ ,  $t_j^*$ ,  $s_j$ ,  $r_j$  and  $r_j^*$  ( $j = 1, \dots, 4$ ) in the gamma regression model for the blood clotting dataset of § 3.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 50 000 replications for each  $n \in \{18, 36, 72, 144\}$ .

Table 1: Empirical null rejection probabilities when testing  $H_0 : \beta_j = \beta_{j0}$  against  $H_1 : \beta_j \neq \beta_{j0}$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t_j$ ,  $t_j^*$ ,  $s_j$ ,  $r_j$  and  $r_j^*$  ( $j = 1, \dots, 4$ ) in the gamma regression model for the blood clotting dataset of § 3.3. Reported rates obtained from a simulation study with 50 000 replications for each  $n \in \{18, 36, 72, 144\}$ .

	$\alpha = 1\%$					$\alpha = 5\%$				
$n = 18$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$
$j = 1$	3.9	1.8	1.2	2.4	1.0	10.4	6.0	7.0	8.8	5.1
$j = 2$	3.8	1.8	1.2	2.5	1.1	10.3	6.0	6.9	8.6	5.0
$j = 3$	4.0	1.8	1.2	2.6	1.0	10.7	6.1	7.2	9.0	5.1
$j = 4$	4.0	1.8	1.2	2.6	1.0	10.7	6.1	7.2	9.0	5.2
$n = 36$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$
$j = 1$	2.2	1.3	1.2	1.6	1.0	7.3	5.4	5.9	6.5	5.0
$j = 2$	2.2	1.3	1.2	1.6	1.0	7.3	5.5	5.9	6.6	5.0
$j = 3$	2.1	1.3	1.1	1.6	1.0	7.1	5.2	5.7	6.4	4.8
$j = 4$	2.1	1.3	1.1	1.6	1.0	7.2	5.3	5.8	6.4	4.9
$n = 72$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$
$j = 1$	1.6	1.2	1.1	1.4	1.1	6.2	5.3	5.5	5.9	5.0
$j = 2$	1.5	1.2	1.1	1.3	1.0	6.1	5.2	5.4	5.8	5.0
$j = 3$	1.5	1.1	1.0	1.2	1.0	6.1	5.2	5.4	5.8	5.0
$j = 4$	1.4	1.1	1.0	1.2	1.0	6.1	5.2	5.4	5.8	5.0
$n = 144$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$
$j = 1$	1.3	1.1	1.1	1.2	1.1	5.5	5.1	5.2	5.4	5.0
$j = 2$	1.2	1.1	1.0	1.1	1.0	5.6	5.1	5.2	5.4	5.0
$j = 3$	1.3	1.1	1.1	1.2	1.0	5.7	5.2	5.3	5.5	5.1
$j = 4$	1.3	1.1	1.1	1.2	1.0	5.6	5.1	5.3	5.4	5.1

### 1.3 Logistic regression for crying babies

Notice that in this part  $t_c$ ,  $s_c$  and  $r_c$  indicate the versions of the one-sided Wald statistic and the signed roots of the score and likelihood ratio statistics, respectively, obtained via the conditional likelihood approach.

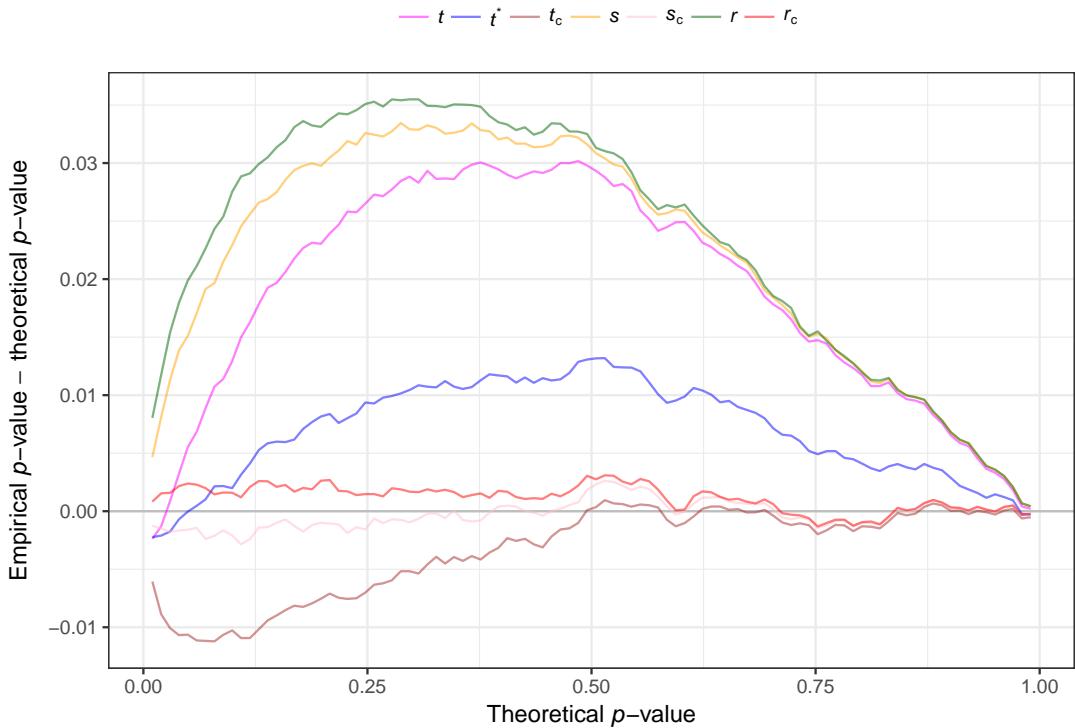


Figure 3: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \gamma = 0$  against  $H_1 : \gamma \neq 0$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $t_c$ ,  $s$ ,  $s_c$ ,  $r$  and  $r_c$  in the logistic regression model for the crying babies dataset of § 3.4. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 50 000 replications.

Table 2: Empirical null rejection probabilities when testing  $H_0 : \gamma = 0$  against  $H_1 : \gamma \neq 0$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $t_c$ ,  $s$ ,  $s_c$ ,  $r$  and  $r_c$  in the logistic regression model for the crying babies dataset of § 3.4. Reported rates obtained from a simulation study with 50 000 replications.

$\alpha = 1\%$							$\alpha = 5\%$						
$t$	$t^*$	$t_c$	$s$	$s_c$	$r$	$r_c$	$t$	$t^*$	$t_c$	$s$	$s_c$	$r$	$r_c$
0.8	0.8	0.4	1.5	0.9	1.8	1.1	5.5	5.0	3.9	6.5	4.8	7.0	5.2

## 1.4 Meta-analysis: simulation studies

Notice that in this part  $r_B$ , DL, KH and  $r_p$  denote the signed root of the Bartlett-corrected likelihood ratio statistic (see, e.g., Huizenga et al., 2011), the one-sided Wald statistic based on the DerSimonian & Laird estimator (1986), a modification of the latter proposed by Knapp and Hartung (2003) and the signed root of the penalized likelihood ratio statistic (Kosmidis et al., 2017), respectively.

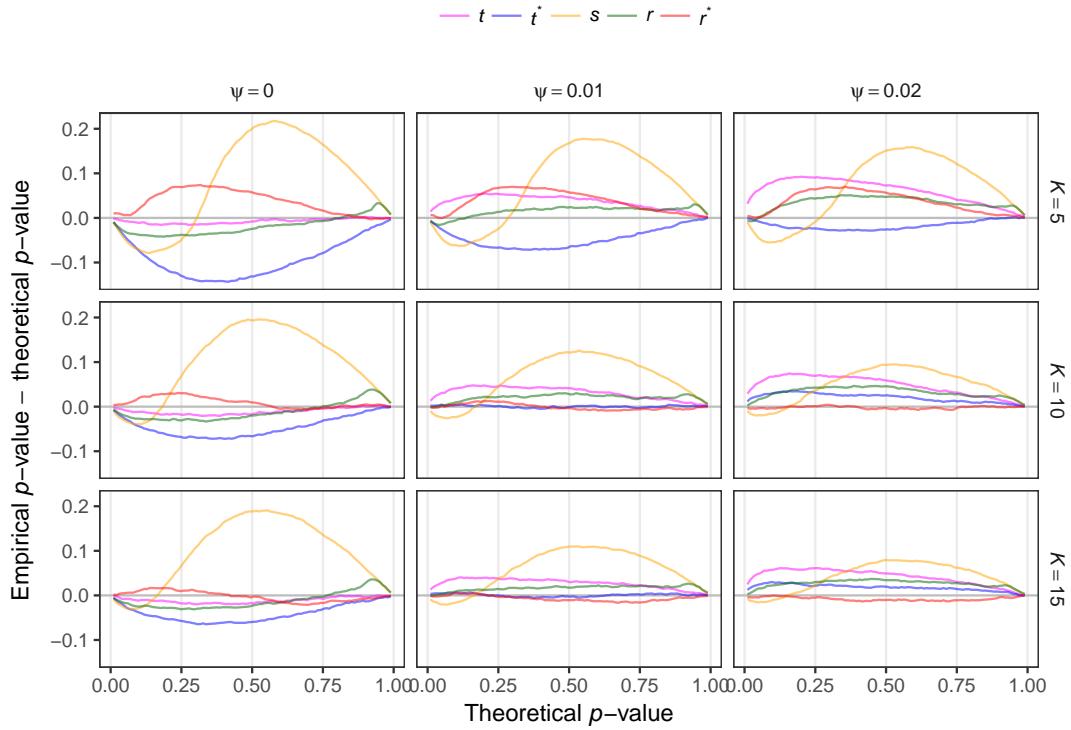


Figure 4: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

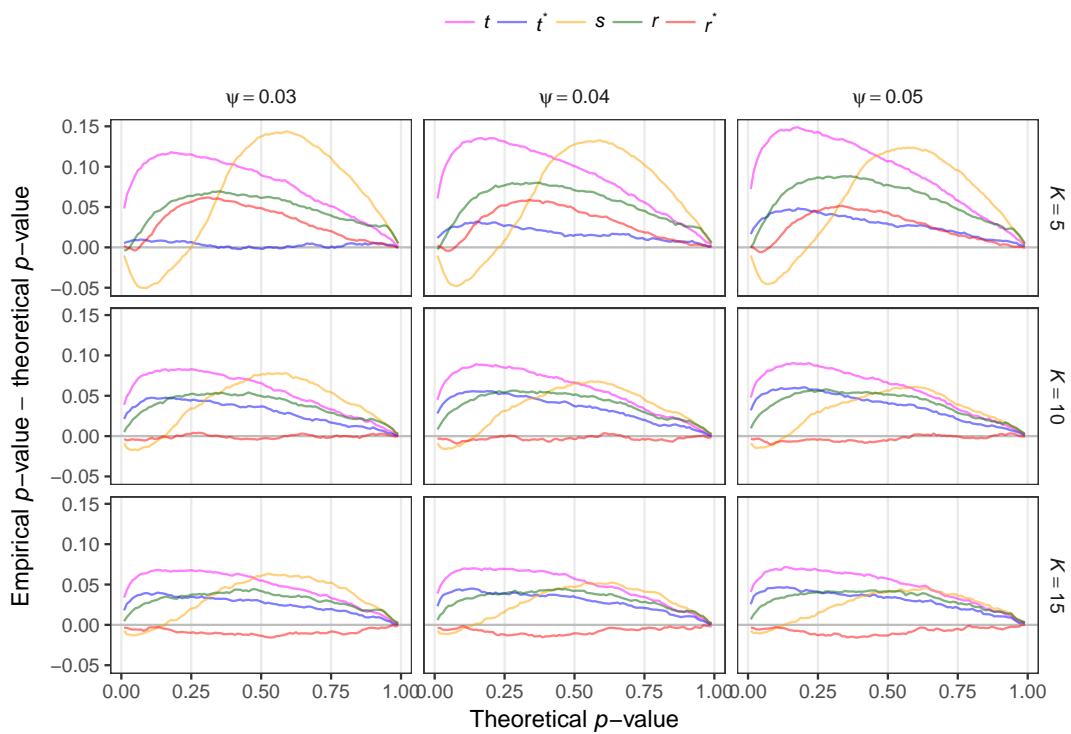


Figure 5: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

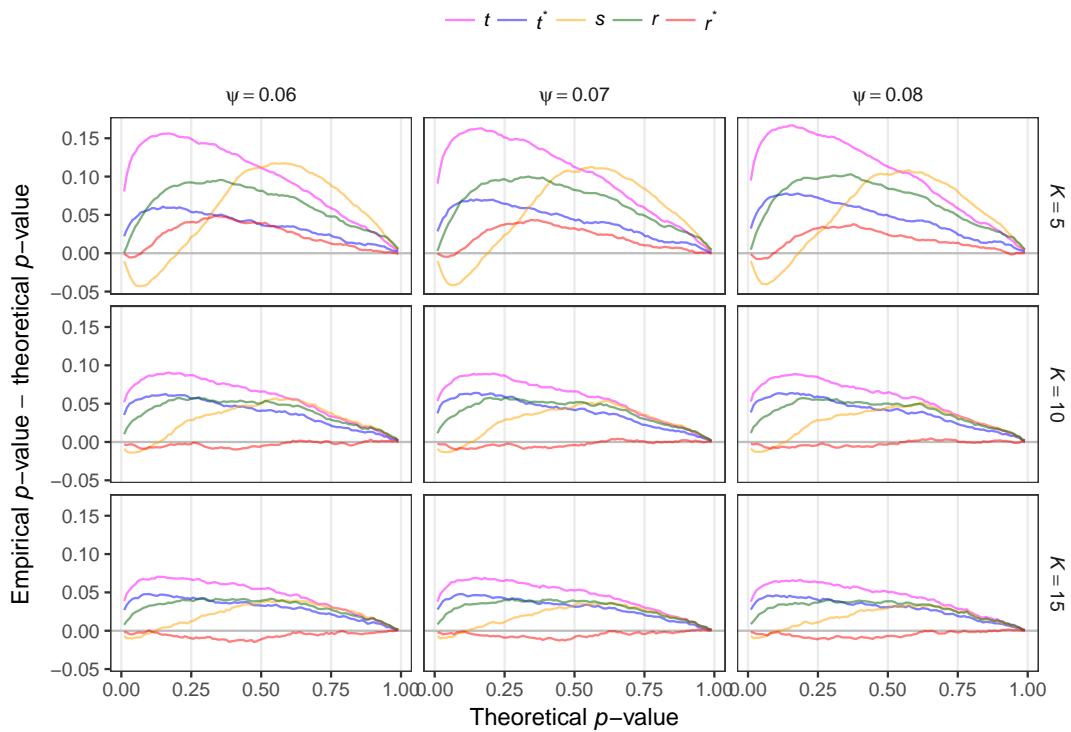


Figure 6: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

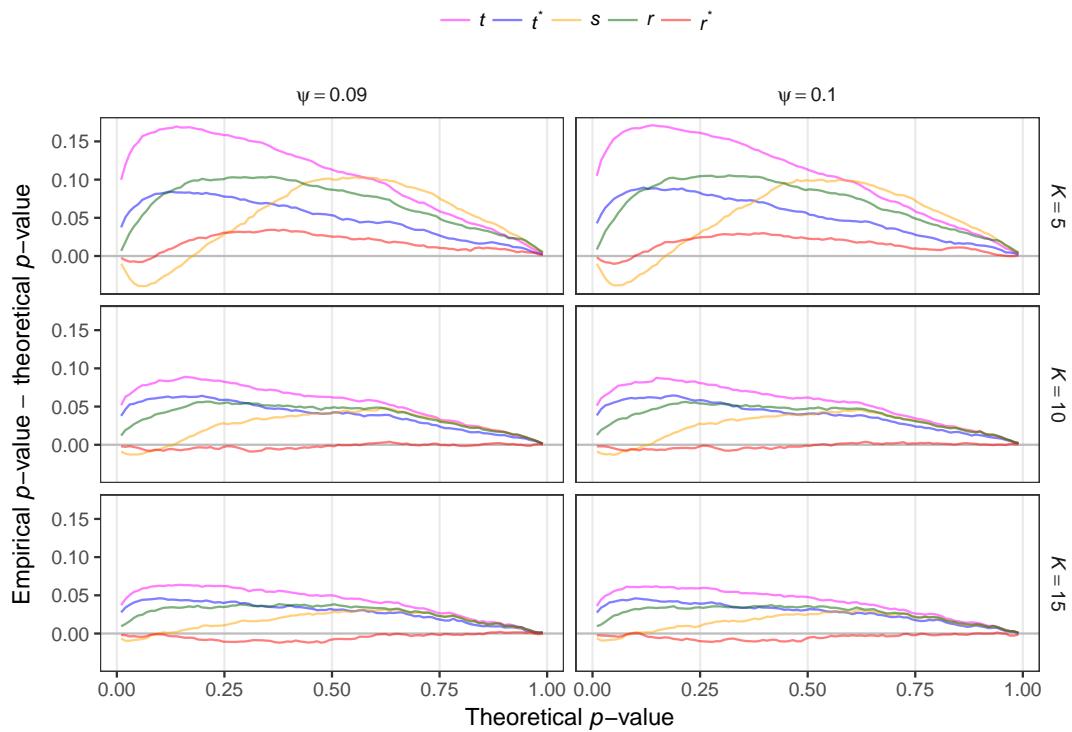


Figure 7: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

Table 3: Empirical null rejection probabilities when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. Reported rates obtained from a simulation study with 10 000 replications for each couple  $(K, \psi)$ .

	$K = 5$	$\alpha = 1\%$					$\alpha = 5\%$				
		$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
$\psi = 0$		0.9	0.1	0.0	0.1	2.0	4.2	0.8	0.2	1.9	5.6
$\psi = 0.01$		2.4	0.4	0.0	0.2	1.6	8.3	2.4	0.4	3.5	4.9
$\psi = 0.02$		4.2	0.9	0.0	0.4	1.4	11.3	4.2	0.4	5.0	5.2
$\psi = 0.03$		5.8	1.5	0.0	0.6	1.2	13.9	5.9	0.5	6.2	4.6
$\psi = 0.04$		7.0	2.2	0.0	0.8	1.2	15.8	7.3	0.6	7.1	4.6
$\psi = 0.05$		8.2	2.6	0.0	1.0	1.1	17.3	8.6	0.7	7.8	4.4
$\psi = 0.06$		9.1	3.2	0.0	1.1	0.9	18.2	9.6	0.9	8.5	4.5
$\psi = 0.07$		10.1	3.9	0.0	1.3	0.9	19.0	10.7	1.0	8.8	4.5
$\psi = 0.08$		10.5	4.2	0.0	1.5	0.9	19.6	11.2	1.0	9.3	4.3
$\psi = 0.09$		11.0	4.8	0.0	1.7	0.8	19.9	11.8	1.1	9.6	4.2
$\psi = 0.1$		11.5	5.3	0.0	1.8	0.8	20.4	12.6	1.1	9.8	4.0
$K = 10$		$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
		$\psi = 0$	0.8	0.2	0.0	0.3	1.4	4.2	2.3	1.8	3.0
$\psi = 0.01$		2.3	1.3	0.1	0.9	0.9	8.1	5.5	2.6	5.4	5.1
$\psi = 0.02$		3.9	2.1	0.1	1.2	0.7	10.3	7.3	3.1	6.7	4.7
$\psi = 0.03$		4.9	3.1	0.1	1.5	0.6	11.7	8.6	3.3	7.3	4.5
$\psi = 0.04$		5.5	3.8	0.1	1.8	0.7	12.2	9.7	3.5	7.8	4.6
$\psi = 0.05$		5.8	4.2	0.1	2.0	0.7	12.4	10.2	3.6	7.9	4.5
$\psi = 0.06$		6.2	4.5	0.1	2.0	0.7	12.5	10.4	3.6	8.0	4.5
$\psi = 0.07$		6.3	4.7	0.1	2.1	0.8	12.5	10.5	3.7	8.1	4.4
$\psi = 0.08$		6.3	4.9	0.1	2.2	0.8	12.6	10.5	3.7	8.1	4.5
$\psi = 0.09$		6.2	4.8	0.1	2.2	0.8	12.5	10.6	3.7	8.0	4.6
$\psi = 0.1$		6.1	4.8	0.1	2.3	0.8	12.3	10.5	3.6	8.1	4.5
$K = 15$		$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
		$\psi = 0$	0.9	0.5	0.1	0.4	1.2	4.2	2.6	2.2	3.3
$\psi = 0.01$		2.3	1.4	0.1	0.9	0.9	7.9	5.5	2.9	5.4	4.9
$\psi = 0.02$		3.5	2.2	0.2	1.2	0.6	9.9	7.4	3.5	6.6	4.5
$\psi = 0.03$		4.4	2.8	0.2	1.5	0.6	10.7	8.5	3.7	7.1	4.5
$\psi = 0.04$		4.9	3.3	0.3	1.6	0.6	11.1	9.0	3.9	7.3	4.4
$\psi = 0.05$		5.1	3.6	0.3	1.7	0.6	11.1	9.1	4.0	7.4	4.5
$\psi = 0.06$		4.9	3.7	0.3	1.8	0.8	11.1	9.3	4.2	7.4	4.7
$\psi = 0.07$		4.8	3.8	0.3	1.8	0.8	10.9	9.1	4.1	7.4	4.6
$\psi = 0.08$		4.8	3.8	0.3	1.9	0.8	10.8	9.2	4.2	7.4	4.5
$\psi = 0.09$		4.7	3.8	0.3	1.9	0.8	10.6	9.2	4.2	7.3	4.6
$\psi = 0.1$		4.6	3.7	0.4	1.9	0.8	10.6	9.2	4.2	7.4	4.7

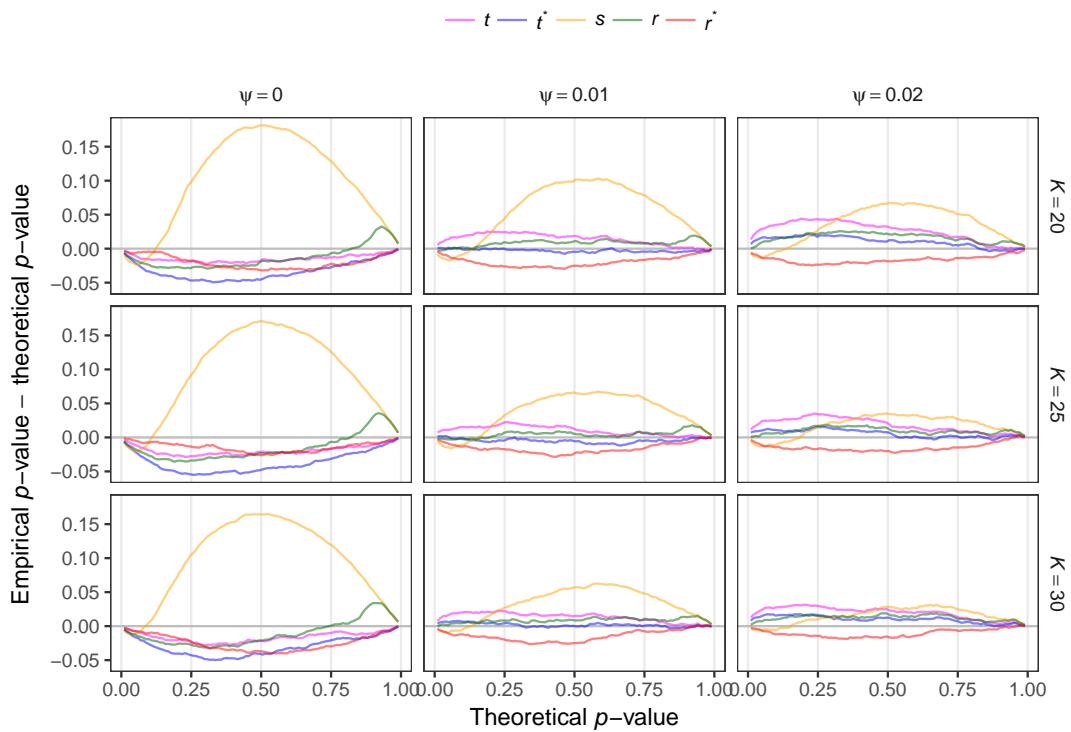


Figure 8: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

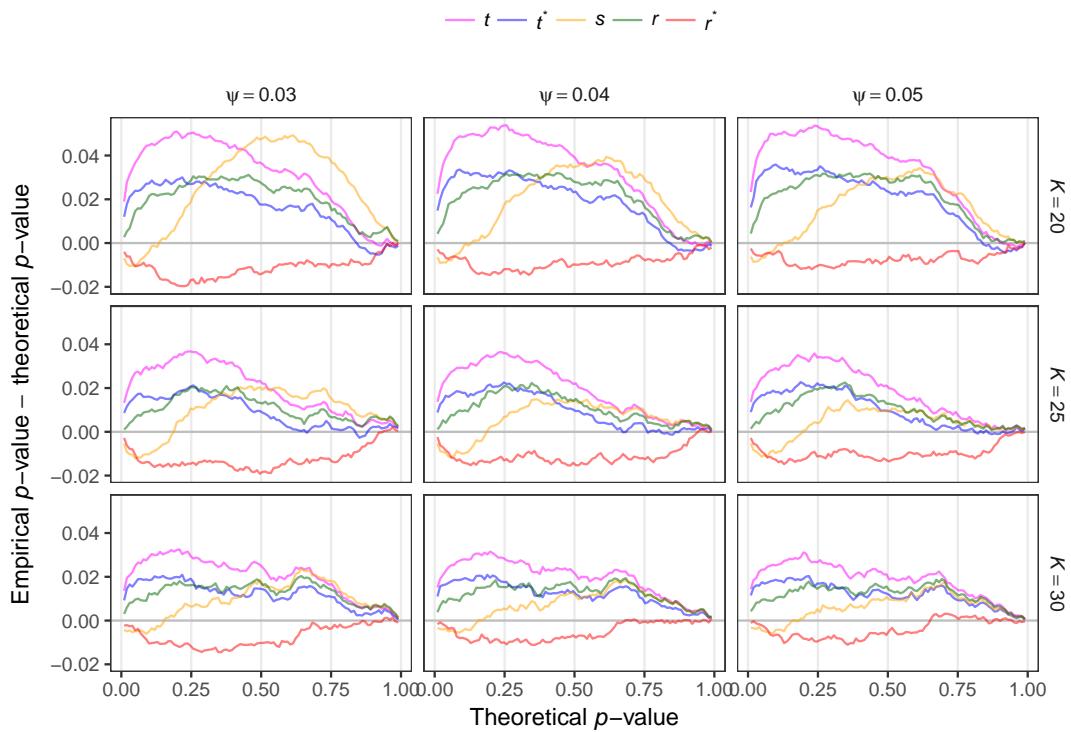


Figure 9: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

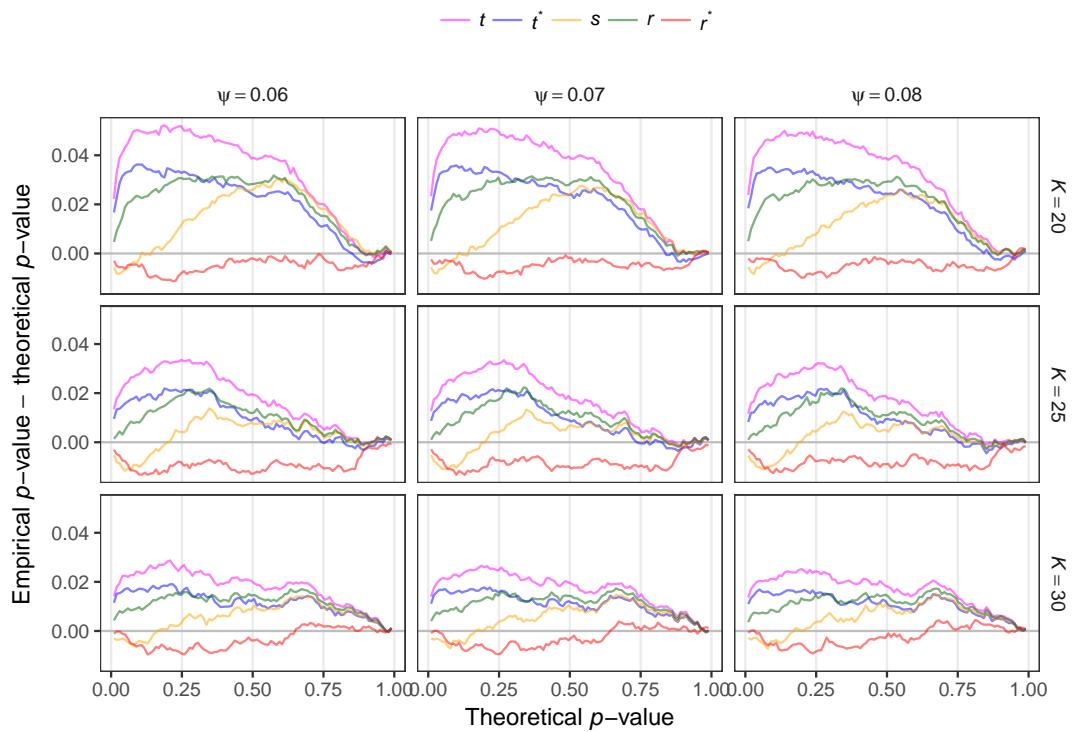


Figure 10: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

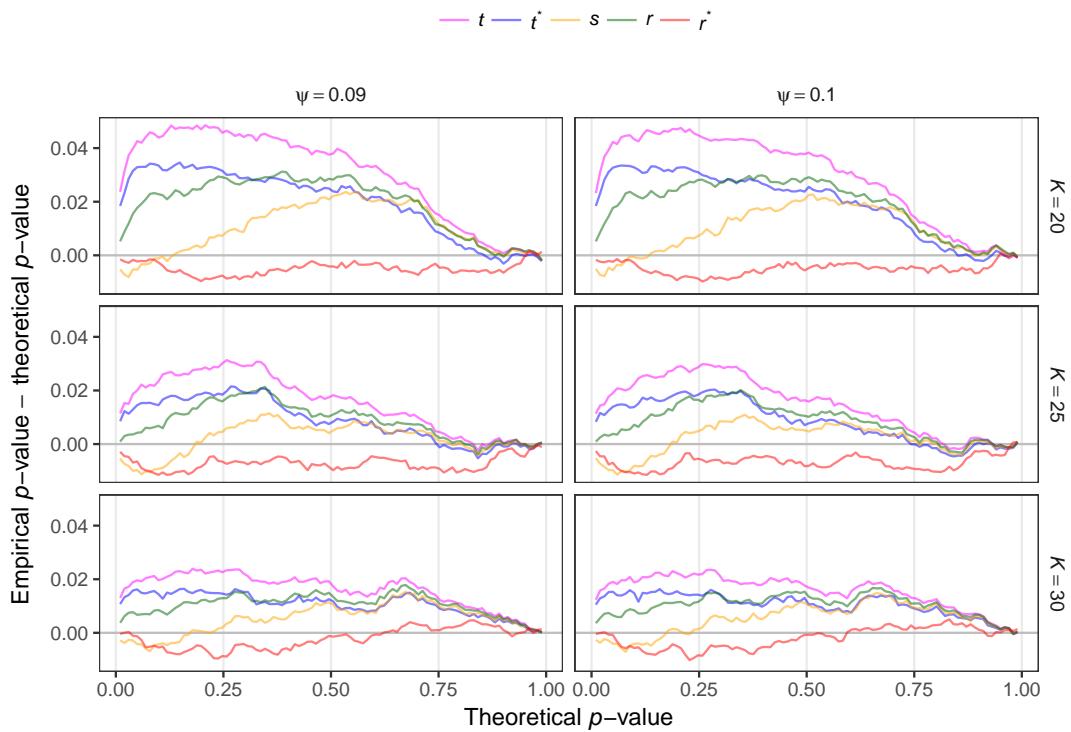


Figure 11: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

Table 4: Empirical null rejection probabilities when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. Reported rates obtained from a simulation study with 10 000 replications for each couple  $(K, \psi)$ .

$K = 20$	$\alpha = 1\%$					$\alpha = 5\%$				
	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
$\psi = 0$	0.7	0.3	0.1	0.4	0.7	4.0	2.9	2.7	3.3	4.3
$\psi = 0.01$	1.5	1.1	0.2	0.7	0.6	6.4	5.1	3.3	5.0	3.8
$\psi = 0.02$	2.4	1.7	0.3	1.0	0.4	7.9	6.7	3.7	6.0	3.6
$\psi = 0.03$	2.9	2.2	0.3	1.3	0.6	8.7	7.4	4.0	6.5	4.0
$\psi = 0.04$	3.2	2.5	0.4	1.4	0.7	9.2	7.9	4.2	6.8	4.3
$\psi = 0.05$	3.3	2.6	0.4	1.4	0.8	9.4	8.2	4.3	6.9	4.4
$\psi = 0.06$	3.2	2.7	0.4	1.5	0.7	9.3	8.3	4.3	6.9	4.4
$\psi = 0.07$	3.3	2.8	0.4	1.5	0.8	9.3	8.3	4.4	6.9	4.6
$\psi = 0.08$	3.4	2.8	0.4	1.5	0.8	9.3	8.3	4.4	6.8	4.8
$\psi = 0.09$	3.4	2.8	0.5	1.5	0.9	9.3	8.3	4.4	6.8	4.7
$\psi = 0.1$	3.3	2.8	0.5	1.5	0.8	9.2	8.2	4.5	6.8	4.8
$K = 25$	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
	0.6	0.4	0.2	0.4	0.9	3.8	2.9	3.0	3.3	4.5
$\psi = 0.01$	1.7	1.2	0.4	0.9	0.7	6.1	5.0	3.5	4.7	4.1
$\psi = 0.02$	2.1	1.7	0.4	1.1	0.7	7.4	6.0	3.7	5.5	3.9
$\psi = 0.03$	2.3	1.9	0.5	1.1	0.7	7.7	6.7	3.8	5.6	3.9
$\psi = 0.04$	2.4	1.9	0.5	1.1	0.8	7.6	6.6	3.9	5.6	4.1
$\psi = 0.05$	2.4	1.9	0.5	1.1	0.7	7.5	6.7	3.8	5.6	4.1
$\psi = 0.06$	2.3	1.9	0.5	1.1	0.7	7.4	6.6	3.8	5.5	4.2
$\psi = 0.07$	2.3	1.9	0.5	1.1	0.7	7.3	6.6	3.9	5.5	4.3
$\psi = 0.08$	2.2	1.8	0.5	1.1	0.7	7.2	6.5	4.0	5.4	4.3
$\psi = 0.09$	2.1	1.8	0.5	1.1	0.7	7.0	6.5	4.0	5.4	4.3
$\psi = 0.1$	2.1	1.8	0.5	1.1	0.7	6.9	6.4	4.0	5.4	4.3
$K = 30$	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
	0.7	0.5	0.3	0.5	0.8	4.3	3.4	3.9	3.8	4.3
$\psi = 0.01$	1.8	1.4	0.5	1.0	0.6	6.7	5.6	4.1	5.1	3.9
$\psi = 0.02$	2.3	1.8	0.5	1.2	0.8	7.4	6.5	4.4	5.7	4.3
$\psi = 0.03$	2.4	1.9	0.6	1.3	0.8	7.5	6.7	4.5	6.0	4.5
$\psi = 0.04$	2.5	2.1	0.7	1.4	0.8	7.4	6.7	4.5	6.0	4.6
$\psi = 0.05$	2.4	2.1	0.7	1.4	0.9	7.3	6.8	4.6	6.0	4.8
$\psi = 0.06$	2.4	2.1	0.7	1.4	0.9	7.3	6.7	4.6	5.8	4.9
$\psi = 0.07$	2.4	2.1	0.7	1.4	0.9	7.1	6.7	4.6	5.8	4.9
$\psi = 0.08$	2.4	2.1	0.7	1.4	0.9	7.1	6.7	4.6	5.7	4.9
$\psi = 0.09$	2.3	2.1	0.7	1.4	1.0	7.0	6.6	4.6	5.7	4.9
$\psi = 0.1$	2.2	2.1	0.7	1.4	1.0	6.9	6.5	4.6	5.6	4.9

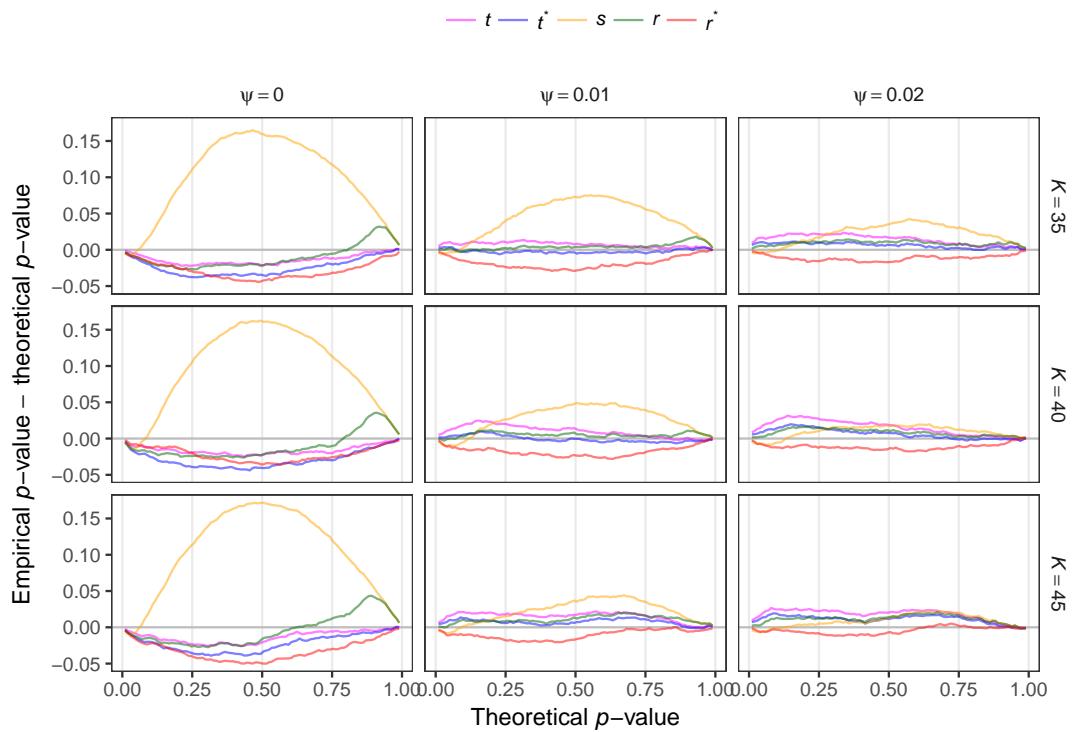


Figure 12: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

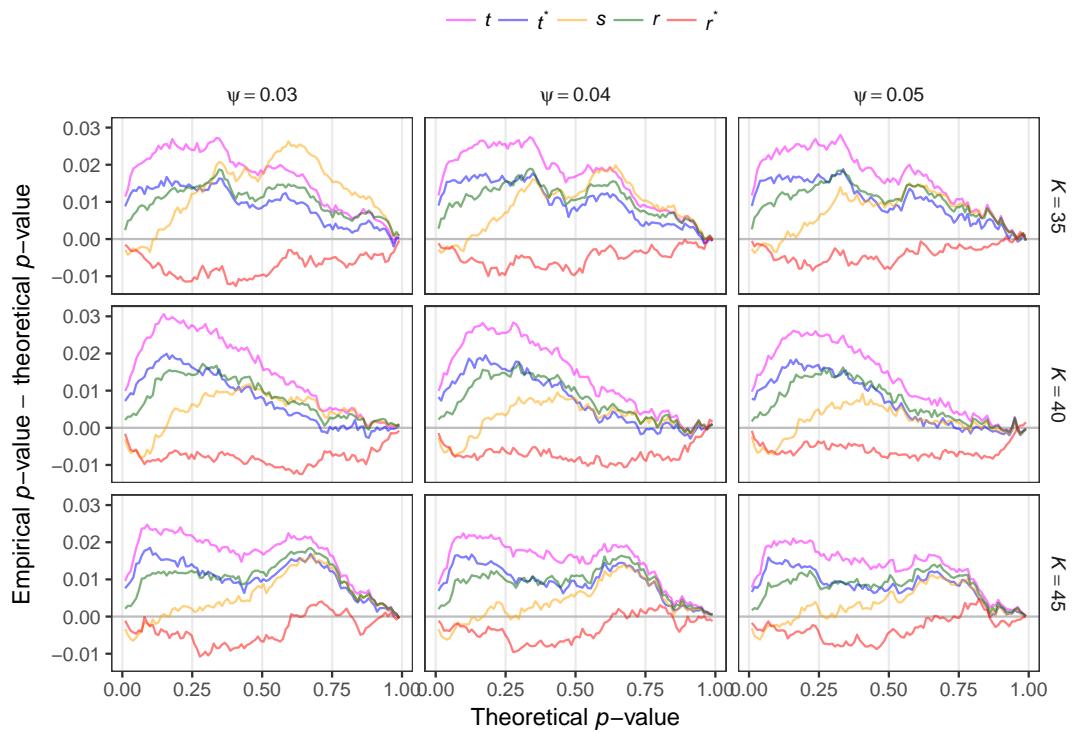


Figure 13: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

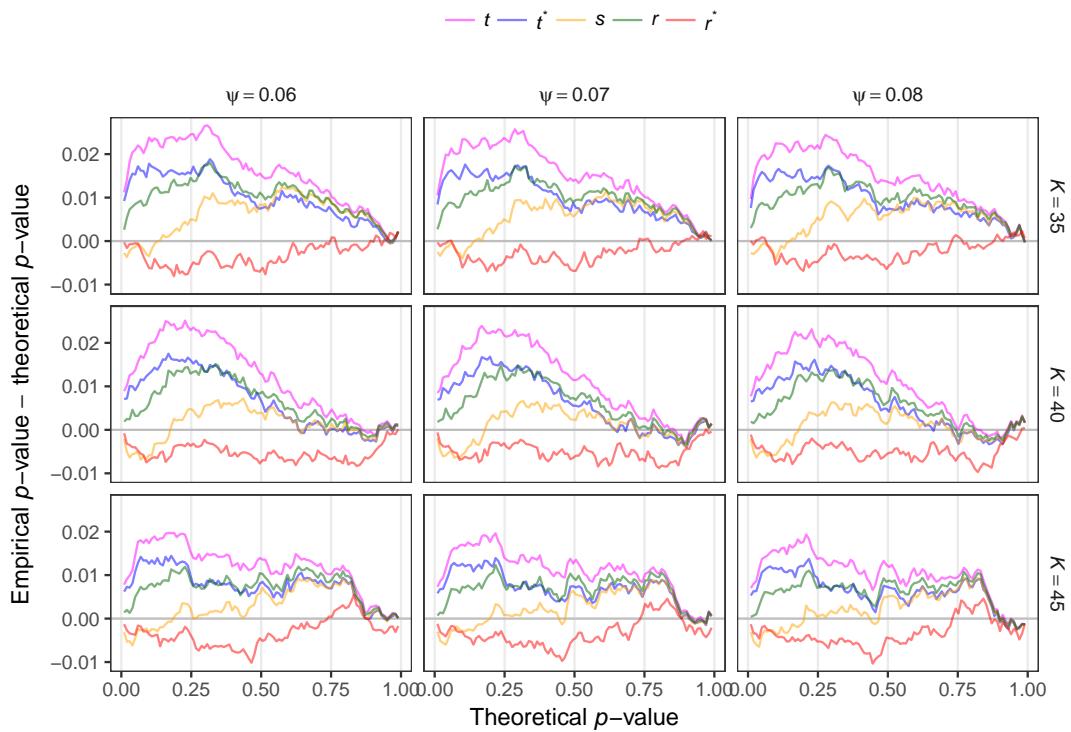


Figure 14: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

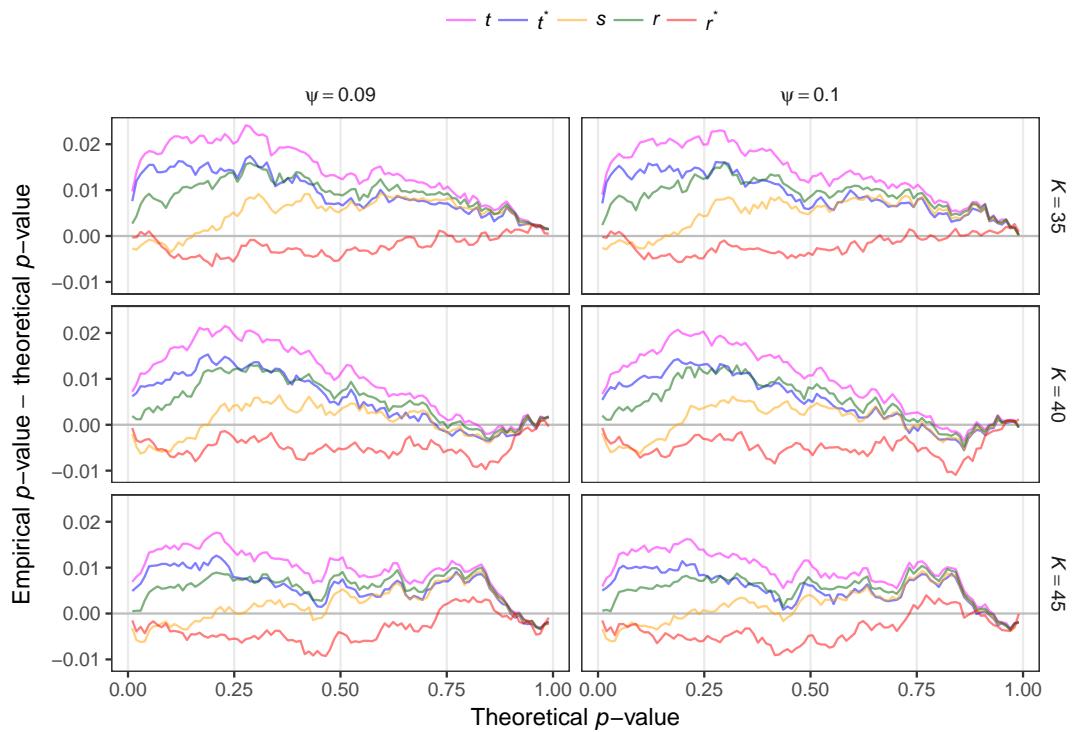


Figure 15: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

Table 5: Empirical null rejection probabilities when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. Reported rates obtained from a simulation study with 10 000 replications for each couple  $(K, \psi)$ .

	$\alpha = 1\%$					$\alpha = 5\%$				
	$K = 35$	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$
$\psi = 0$	0.9	0.7	0.5	0.6	0.6	4.5	3.8	4.8	4.0	4.0
$\psi = 0.01$	1.5	1.3	0.6	0.9	0.7	6.1	5.4	4.7	5.2	4.2
$\psi = 0.02$	1.9	1.7	0.7	1.1	0.8	6.7	6.1	4.7	5.5	4.5
$\psi = 0.03$	2.1	1.9	0.7	1.2	0.9	7.0	6.3	4.7	5.7	4.7
$\psi = 0.04$	2.2	1.9	0.7	1.3	0.9	7.1	6.6	4.8	5.9	4.8
$\psi = 0.05$	2.2	1.9	0.7	1.3	0.9	7.1	6.6	4.9	5.9	4.8
$\psi = 0.06$	2.1	1.9	0.7	1.3	1.0	7.1	6.5	4.9	5.9	5.0
$\psi = 0.07$	2.1	1.8	0.8	1.3	1.0	7.0	6.6	4.9	5.9	5.0
$\psi = 0.08$	2.0	1.8	0.7	1.3	1.0	7.0	6.6	4.9	5.9	5.1
$\psi = 0.09$	2.0	1.8	0.7	1.3	1.0	6.9	6.5	4.9	5.9	5.1
$\psi = 0.1$	1.9	1.7	0.7	1.2	1.0	6.8	6.5	4.9	5.9	5.1
$K = 40$	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
$\psi = 0$	0.7	0.4	0.4	0.6	0.8	3.8	3.0	3.8	3.4	4.3
$\psi = 0.01$	1.5	1.2	0.6	0.9	0.8	6.1	5.4	4.1	5.0	3.9
$\psi = 0.02$	1.8	1.6	0.7	1.2	0.8	6.8	5.9	4.3	5.2	4.1
$\psi = 0.03$	2.0	1.7	0.7	1.2	0.8	6.9	6.2	4.4	5.4	4.4
$\psi = 0.04$	2.0	1.8	0.7	1.2	0.9	6.7	6.1	4.5	5.4	4.5
$\psi = 0.05$	2.0	1.8	0.8	1.2	0.9	6.6	6.0	4.5	5.3	4.6
$\psi = 0.06$	1.9	1.7	0.8	1.2	0.9	6.3	5.9	4.4	5.3	4.6
$\psi = 0.07$	1.8	1.7	0.8	1.2	0.9	6.3	5.9	4.5	5.3	4.6
$\psi = 0.08$	1.8	1.7	0.8	1.2	0.9	6.2	5.8	4.5	5.4	4.6
$\psi = 0.09$	1.7	1.6	0.8	1.2	0.9	6.2	5.9	4.5	5.4	4.7
$\psi = 0.1$	1.7	1.5	0.8	1.2	0.9	6.1	5.9	4.5	5.4	4.7
$K = 45$	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
$\psi = 0$	0.8	0.6	0.4	0.6	0.6	4.0	3.5	4.3	3.8	3.2
$\psi = 0.01$	1.7	1.5	0.6	1.1	0.7	6.6	6.0	4.2	5.3	3.9
$\psi = 0.02$	1.9	1.7	0.6	1.2	0.9	7.0	6.4	4.4	5.5	4.4
$\psi = 0.03$	2.0	1.8	0.7	1.2	0.9	6.8	6.5	4.5	5.5	4.5
$\psi = 0.04$	1.9	1.7	0.7	1.2	0.9	6.7	6.3	4.4	5.5	4.6
$\psi = 0.05$	1.8	1.7	0.7	1.2	0.8	6.6	6.1	4.5	5.5	4.6
$\psi = 0.06$	1.8	1.6	0.7	1.1	0.9	6.5	6.0	4.6	5.5	4.7
$\psi = 0.07$	1.7	1.6	0.7	1.1	0.9	6.4	6.1	4.6	5.5	4.7
$\psi = 0.08$	1.7	1.5	0.7	1.0	0.9	6.4	6.1	4.6	5.5	4.8
$\psi = 0.09$	1.7	1.5	0.7	1.1	0.9	6.3	6.0	4.6	5.5	4.8
$\psi = 0.1$	1.6	1.5	0.7	1.1	0.9	6.3	6.0	4.6	5.5	4.8

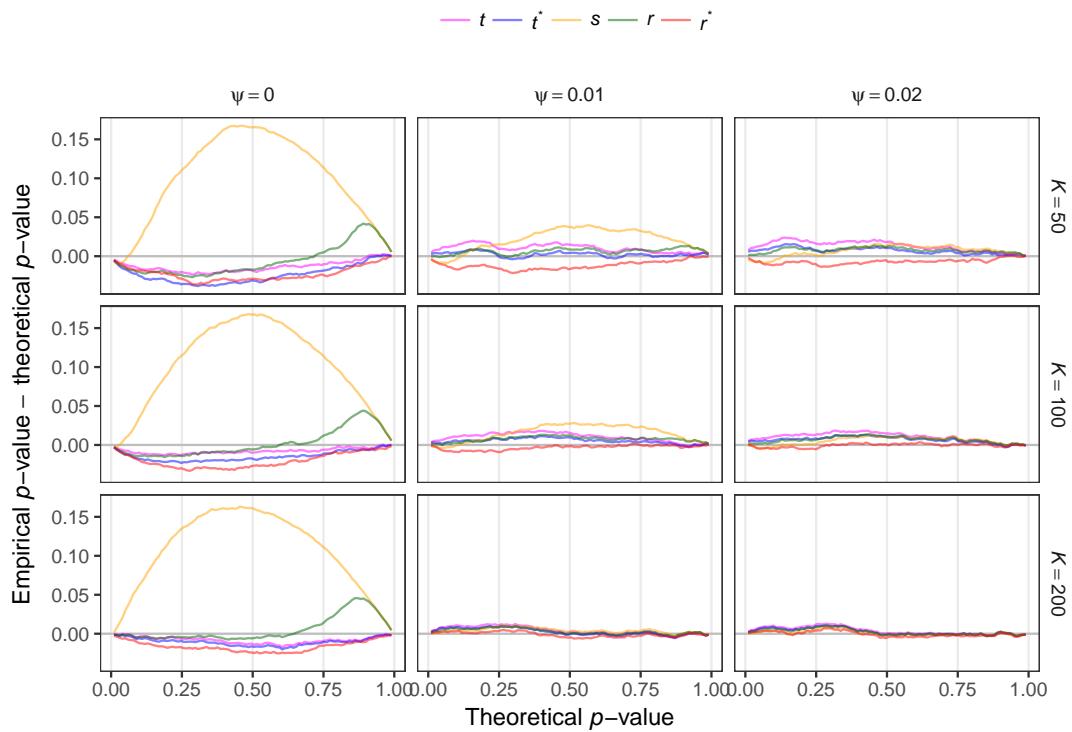


Figure 16: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

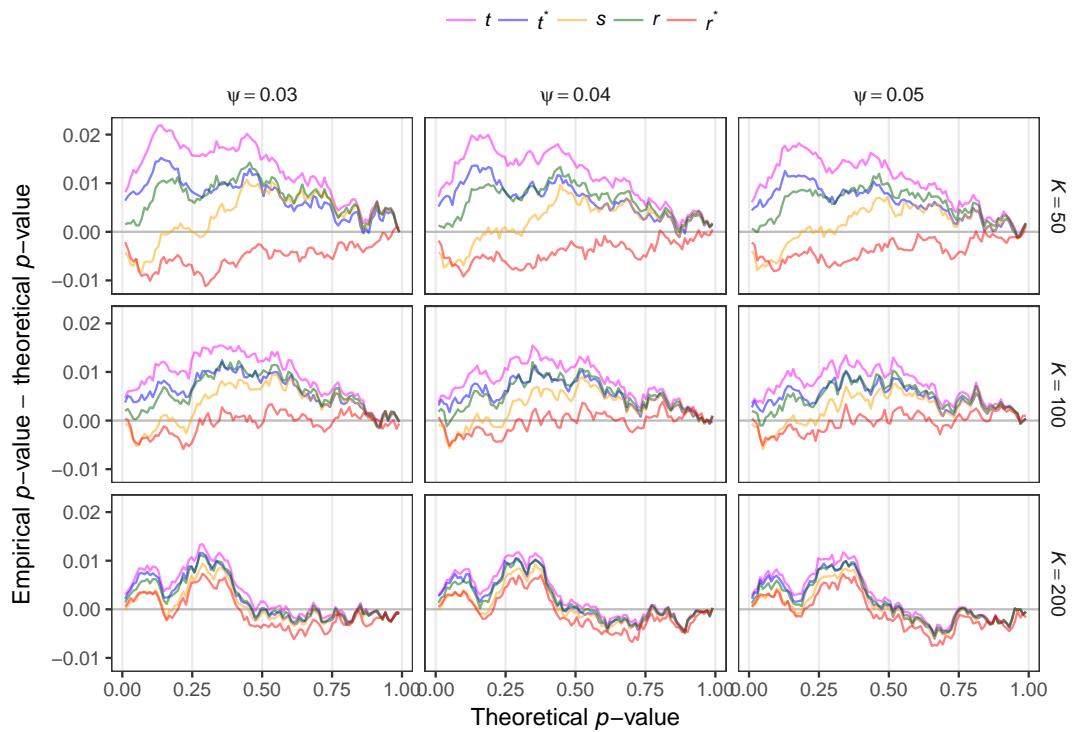


Figure 17: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

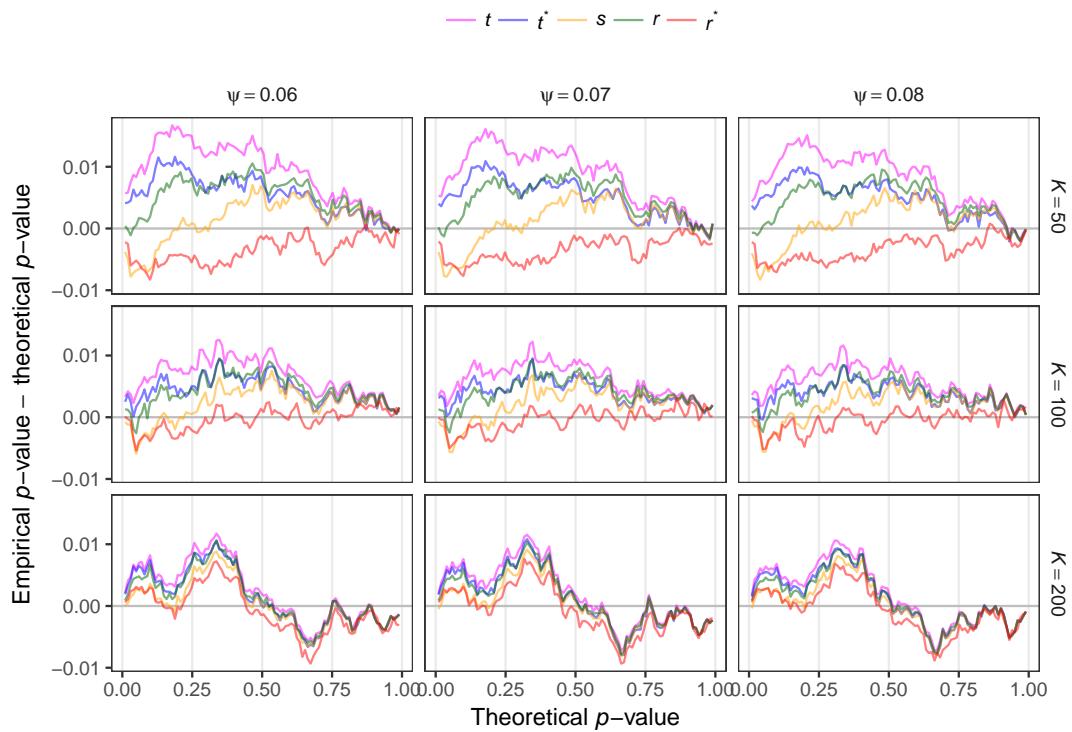


Figure 18: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

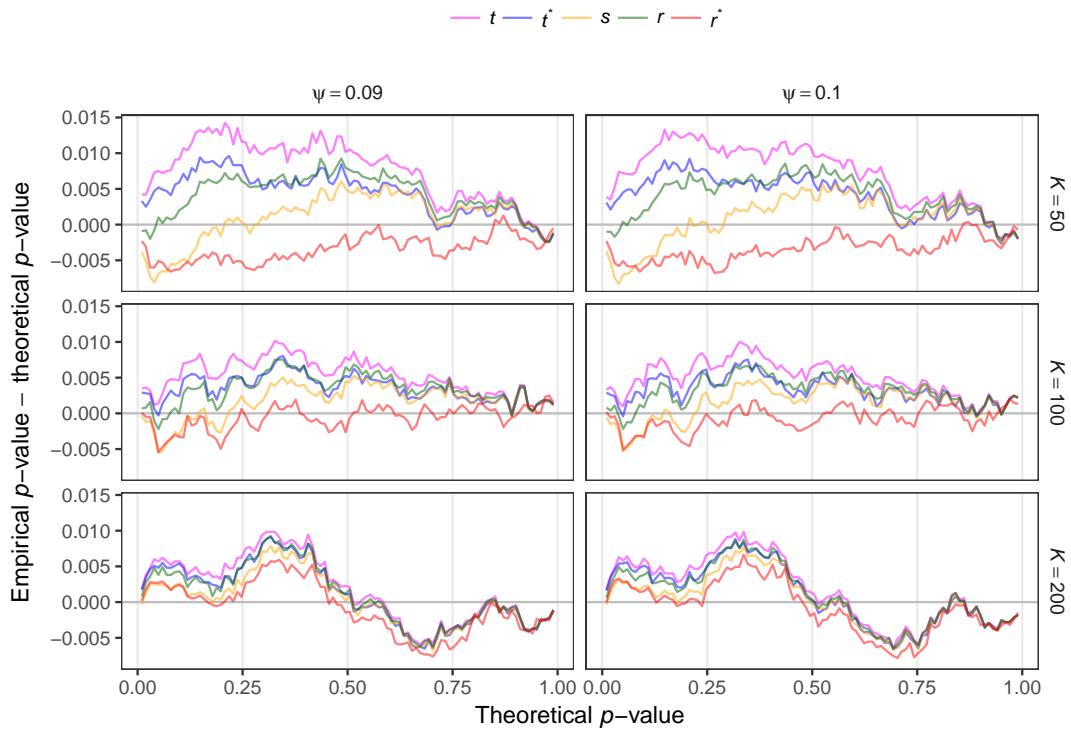


Figure 19: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

Table 6: Empirical null rejection probabilities when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t$ ,  $t^*$ ,  $s$ ,  $r$  and  $r^*$  in the meta-analysis model of § 4.3. Reported rates obtained from a simulation study with 10 000 replications for each couple  $(K, \psi)$ .

$K = 50$	$\alpha = 1\%$					$\alpha = 5\%$				
	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
$\psi = 0$	0.6	0.4	0.3	0.4	0.6	4.0	3.3	4.5	3.5	3.6
$\psi = 0.01$	1.5	1.3	0.6	1.0	0.7	6.1	5.4	4.2	4.9	3.9
$\psi = 0.02$	1.8	1.6	0.6	1.1	0.8	6.3	5.8	4.4	5.3	4.3
$\psi = 0.03$	1.8	1.6	0.6	1.1	0.8	6.2	5.8	4.5	5.1	4.4
$\psi = 0.04$	1.7	1.5	0.6	1.1	0.8	6.0	5.7	4.4	5.1	4.5
$\psi = 0.05$	1.6	1.4	0.6	1.1	0.8	5.9	5.5	4.4	5.1	4.5
$\psi = 0.06$	1.6	1.4	0.6	1.0	0.8	5.8	5.5	4.3	5.0	4.5
$\psi = 0.07$	1.5	1.4	0.6	1.0	0.8	5.9	5.6	4.3	5.1	4.4
$\psi = 0.08$	1.4	1.4	0.6	0.9	0.8	5.8	5.5	4.2	5.1	4.5
$\psi = 0.09$	1.4	1.3	0.6	0.9	0.8	5.8	5.5	4.3	5.1	4.5
$\psi = 0.1$	1.4	1.3	0.6	0.9	0.8	5.7	5.4	4.3	5.0	4.5
$K = 100$	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
	0.8	0.6	0.8	0.8	0.7	4.2	3.8	5.9	3.9	3.8
$\psi = 0$	1.4	1.3	0.8	1.1	0.9	5.7	5.1	4.6	5.0	4.3
$\psi = 0.01$	1.5	1.4	0.9	1.2	1.0	5.7	5.5	4.6	5.1	4.5
$\psi = 0.02$	1.5	1.4	0.9	1.2	1.0	5.6	5.4	4.5	5.0	4.5
$\psi = 0.03$	1.4	1.3	0.9	1.2	1.1	5.5	5.2	4.4	5.0	4.5
$\psi = 0.04$	1.4	1.3	0.9	1.2	1.1	5.4	5.2	4.4	4.9	4.5
$\psi = 0.05$	1.4	1.3	0.9	1.2	1.0	5.4	5.2	4.4	4.9	4.5
$\psi = 0.06$	1.3	1.2	0.9	1.1	1.0	5.4	5.2	4.4	4.8	4.4
$\psi = 0.07$	1.4	1.3	0.9	1.1	1.0	5.3	5.1	4.4	4.8	4.5
$\psi = 0.08$	1.4	1.3	0.9	1.1	1.0	5.2	5.0	4.4	4.8	4.5
$\psi = 0.09$	1.4	1.3	1.0	1.1	1.0	5.1	5.0	4.4	4.8	4.5
$\psi = 0.1$	1.3	1.3	0.9	1.1	1.0	5.1	4.9	4.5	4.8	4.5
$K = 200$	$t$	$t^*$	$s$	$r$	$r^*$	$t$	$t^*$	$s$	$r$	$r^*$
	0.9	0.9	1.1	0.9	0.7	4.9	4.7	8.0	4.8	4.4
$\psi = 0$	1.3	1.2	1.0	1.1	1.1	5.9	5.8	5.2	5.6	5.3
$\psi = 0.01$	1.3	1.2	1.0	1.1	1.1	5.8	5.7	5.2	5.5	5.3
$\psi = 0.02$	1.3	1.2	1.0	1.1	1.1	5.7	5.6	5.2	5.4	5.2
$\psi = 0.03$	1.3	1.2	1.0	1.1	1.1	5.6	5.5	5.3	5.4	5.3
$\psi = 0.04$	1.3	1.3	1.0	1.1	1.1	5.6	5.5	5.3	5.4	5.3
$\psi = 0.05$	1.3	1.2	1.0	1.1	1.1	5.6	5.5	5.3	5.4	5.3
$\psi = 0.06$	1.2	1.2	1.0	1.1	1.1	5.6	5.5	5.3	5.4	5.3
$\psi = 0.07$	1.2	1.2	0.9	1.1	1.1	5.5	5.5	5.2	5.5	5.3
$\psi = 0.08$	1.2	1.1	1.0	1.1	1.1	5.6	5.5	5.2	5.4	5.2
$\psi = 0.09$	1.2	1.2	1.0	1.1	1.0	5.6	5.5	5.2	5.4	5.2
$\psi = 0.1$	1.2	1.2	1.0	1.1	1.0	5.6	5.5	5.3	5.4	5.3

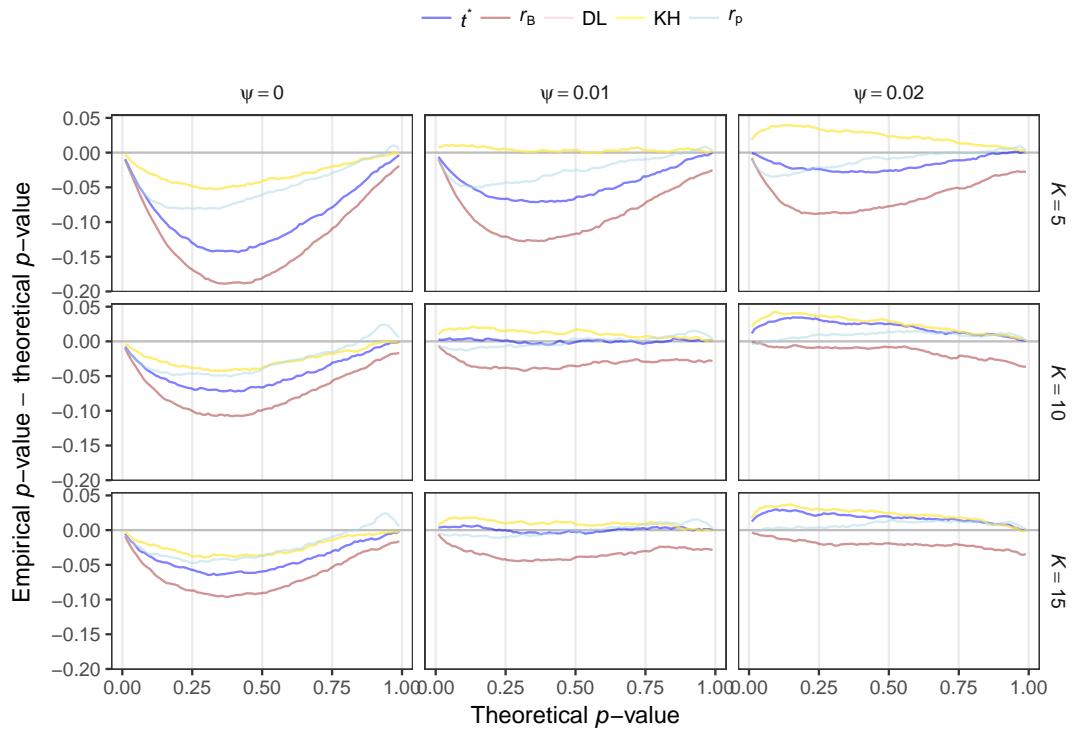


Figure 20: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

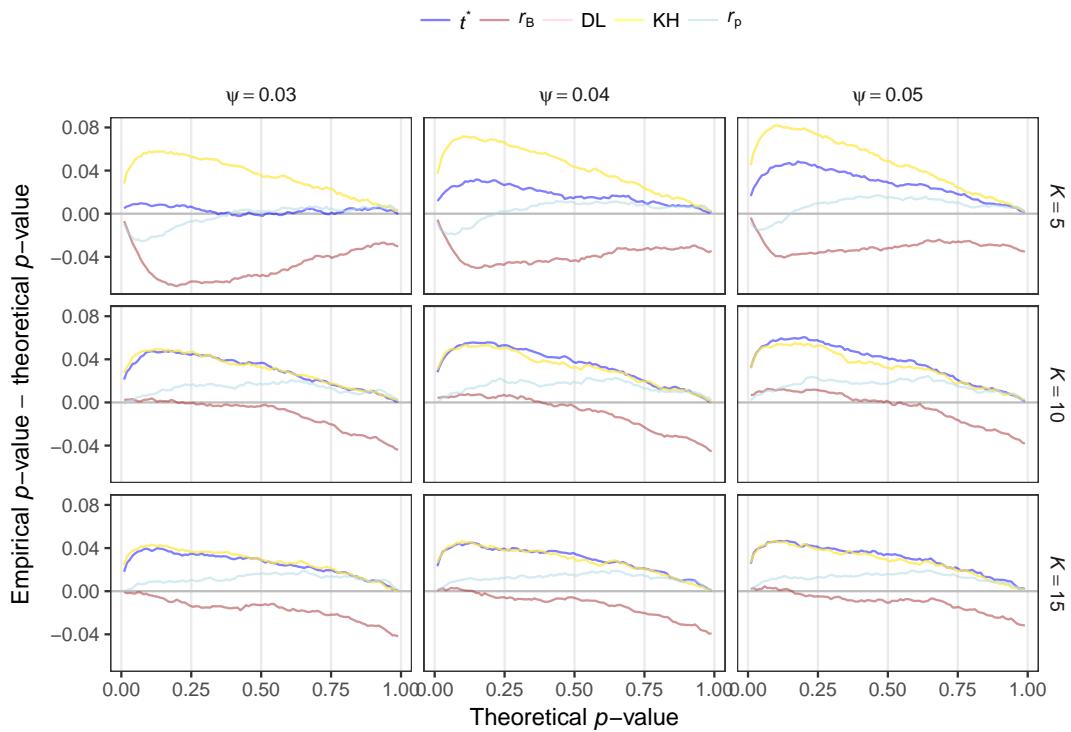


Figure 21: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

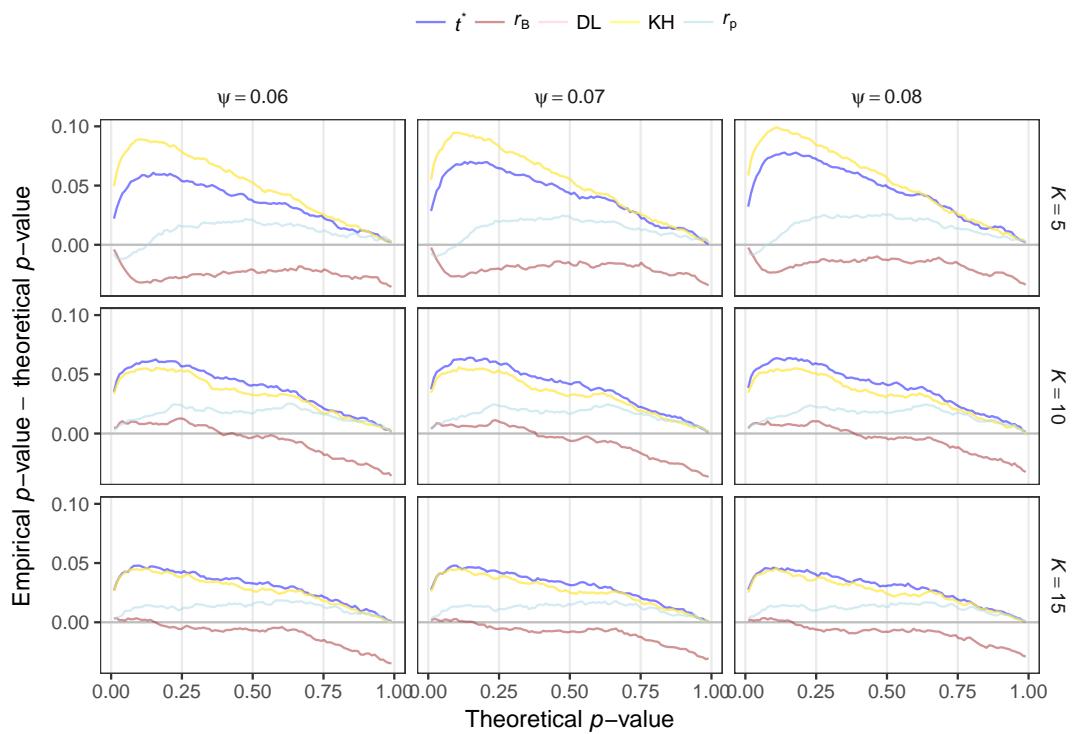


Figure 22: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

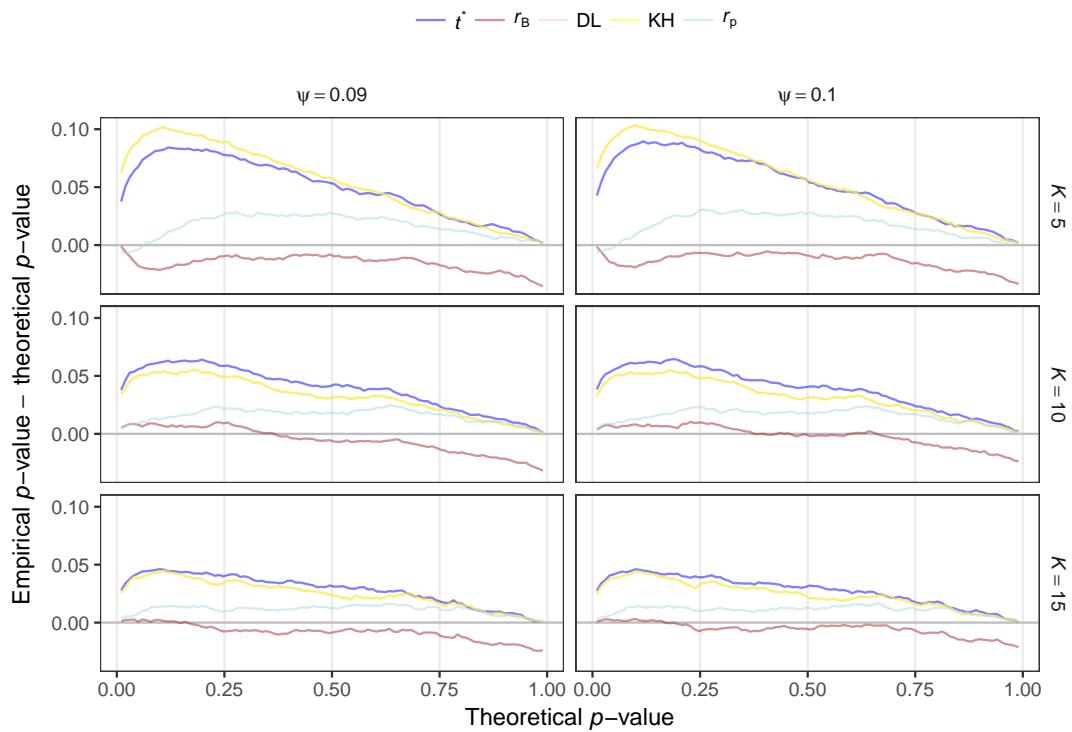


Figure 23: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

Table 7: Empirical null rejection probabilities when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. Reported rates obtained from a simulation study with 10 000 replications for each couple  $(K, \psi)$ .

	$K = 5$	$\alpha = 1\%$					$\alpha = 5\%$				
		$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
$\psi = 0$		0.1	0.0	0.7	0.7	0.0	0.8	0.2	3.1	3.1	0.7
$\psi = 0.01$		0.4	0.1	1.8	1.8	0.1	2.4	0.6	6.0	6.0	1.4
$\psi = 0.02$		0.9	0.2	2.9	2.9	0.1	4.2	1.2	8.2	8.2	2.1
$\psi = 0.03$		1.5	0.3	3.8	3.8	0.1	5.9	1.8	10.0	10.0	2.5
$\psi = 0.04$		2.2	0.4	4.7	4.7	0.2	7.3	2.2	11.3	11.3	3.1
$\psi = 0.05$		2.6	0.6	5.5	5.5	0.2	8.6	2.6	12.3	12.3	3.5
$\psi = 0.06$		3.2	0.6	5.9	5.9	0.2	9.6	2.8	12.8	12.8	3.9
$\psi = 0.07$		3.9	0.8	6.5	6.5	0.3	10.7	3.0	13.1	13.1	4.2
$\psi = 0.08$		4.2	0.8	6.9	6.9	0.4	11.2	3.2	13.7	13.7	4.4
$\psi = 0.09$		4.8	0.9	7.2	7.2	0.4	11.8	3.3	13.9	13.9	4.6
$\psi = 0.1$		5.3	0.9	7.6	7.6	0.4	12.6	3.4	14.3	14.3	4.9
$K = 10$		$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
		$\psi = 0$	0.2	0.0	0.6	0.6	0.1	2.3	1.2	3.6	3.6
$\psi = 0.01$		1.3	0.4	2.0	2.0	0.4	5.5	3.1	6.9	6.9	3.9
$\psi = 0.02$		2.1	0.9	3.2	3.2	0.6	7.3	4.6	8.4	8.4	4.8
$\psi = 0.03$		3.1	1.2	3.8	3.8	0.9	8.6	5.2	9.4	9.4	5.4
$\psi = 0.04$		3.8	1.4	4.1	4.1	1.0	9.7	5.4	9.8	9.8	5.7
$\psi = 0.05$		4.2	1.7	4.2	4.2	1.2	10.2	6.0	10.1	10.1	5.8
$\psi = 0.06$		4.5	1.5	4.4	4.4	1.2	10.4	5.9	10.1	10.1	5.9
$\psi = 0.07$		4.7	1.4	4.5	4.5	1.3	10.5	5.7	10.1	10.1	5.9
$\psi = 0.08$		4.9	1.5	4.5	4.5	1.4	10.5	5.8	10.1	10.1	5.9
$\psi = 0.09$		4.8	1.5	4.4	4.4	1.4	10.6	5.8	10.0	10.0	5.9
$\psi = 0.1$		4.8	1.4	4.3	4.3	1.4	10.5	5.6	10.0	10.0	5.8
$K = 15$		$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
		$\psi = 0$	0.5	0.1	0.7	0.7	0.3	2.6	1.5	3.5	3.5
$\psi = 0.01$		1.4	0.4	1.7	1.7	0.6	5.5	3.0	6.7	6.7	4.2
$\psi = 0.02$		2.2	0.7	2.7	2.7	0.8	7.4	4.2	8.2	8.2	5.1
$\psi = 0.03$		2.8	0.9	3.4	3.4	1.0	8.5	4.8	8.8	8.8	5.5
$\psi = 0.04$		3.3	1.1	3.6	3.6	1.1	9.0	5.2	9.1	9.1	5.7
$\psi = 0.05$		3.6	1.2	3.7	3.7	1.2	9.1	5.2	9.2	9.2	5.8
$\psi = 0.06$		3.7	1.3	3.6	3.6	1.3	9.3	5.1	9.2	9.2	6.0
$\psi = 0.07$		3.8	1.1	3.6	3.6	1.3	9.1	5.1	9.2	9.2	6.0
$\psi = 0.08$		3.8	1.1	3.5	3.5	1.3	9.2	5.1	9.0	9.0	5.8
$\psi = 0.09$		3.8	1.1	3.6	3.6	1.3	9.2	5.1	8.8	8.8	5.8
$\psi = 0.1$		3.7	1.1	3.5	3.5	1.3	9.2	5.2	8.9	8.9	5.7

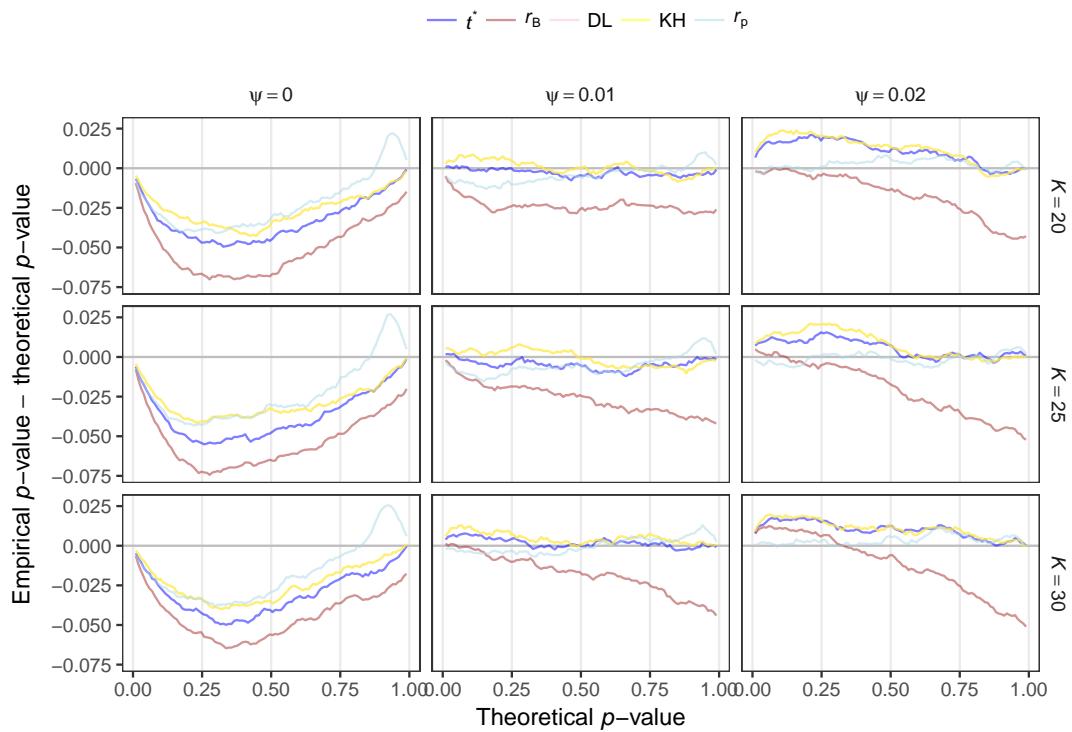


Figure 24: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

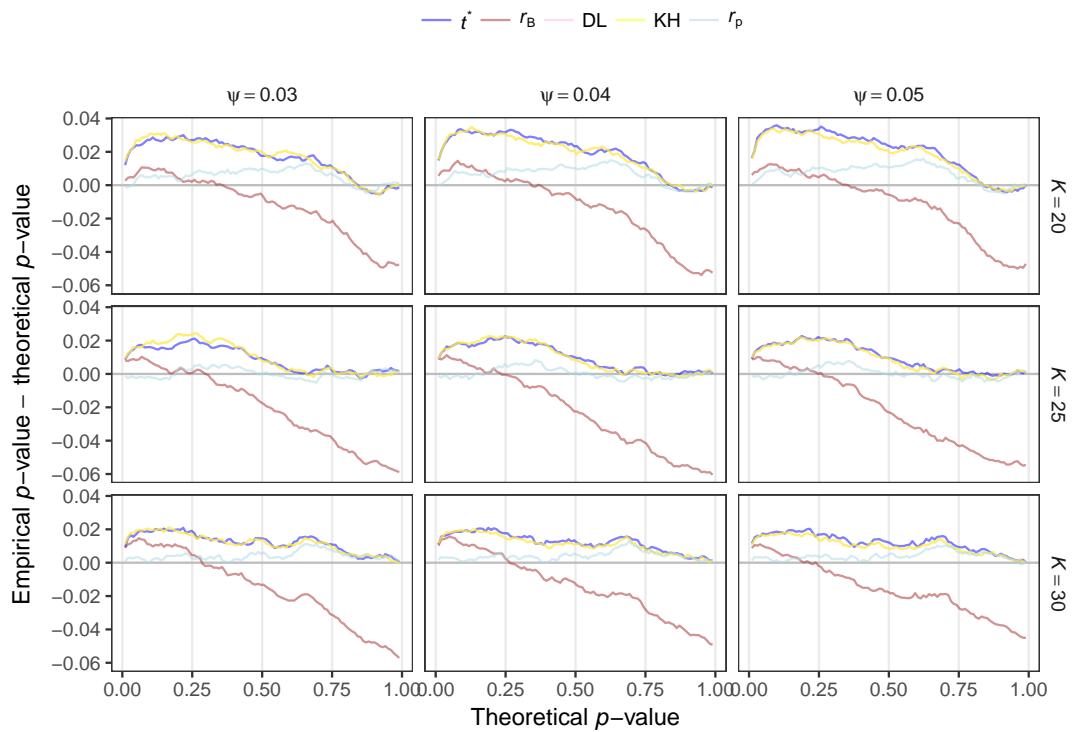


Figure 25: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

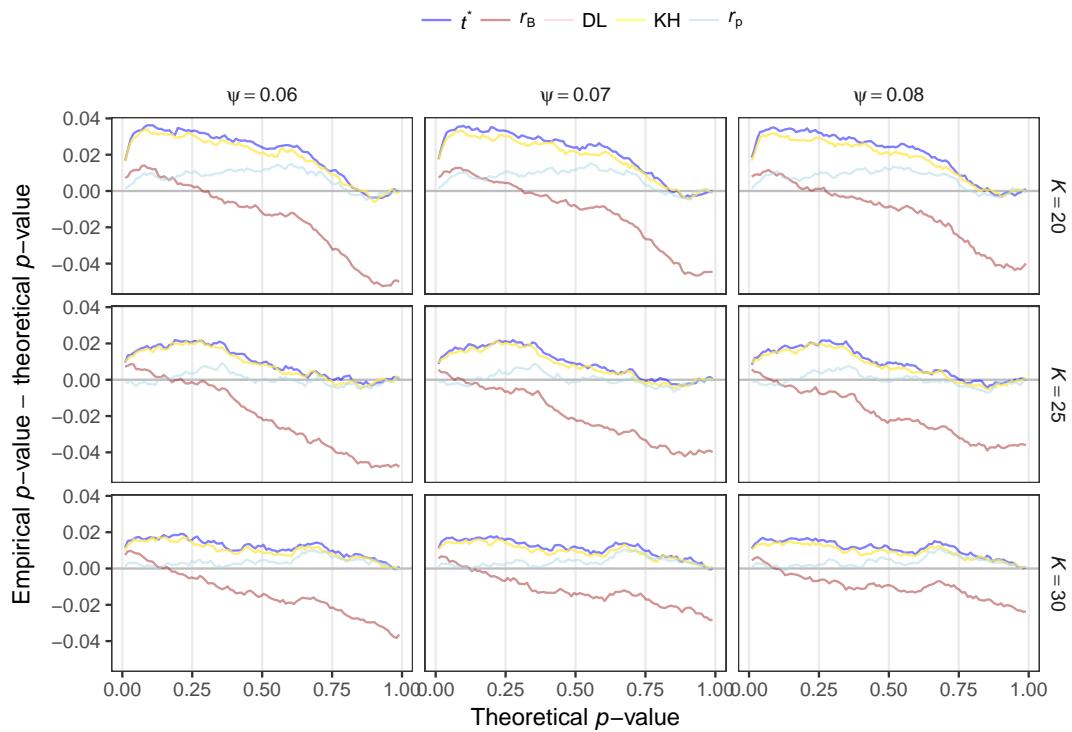


Figure 26: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

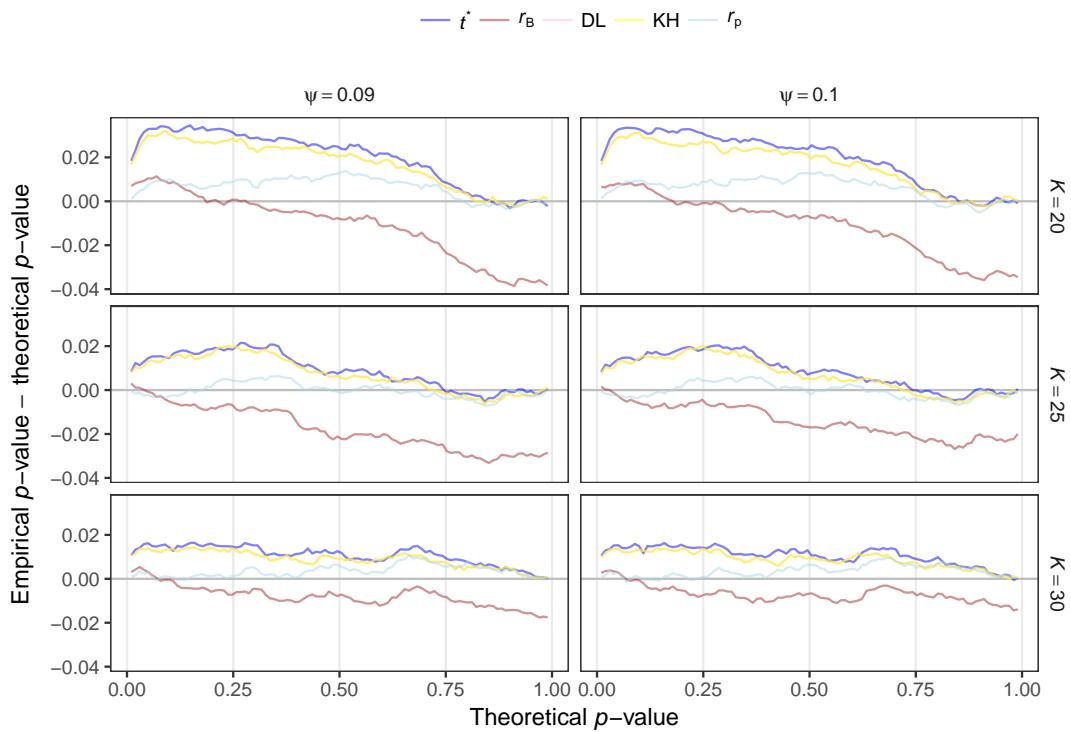


Figure 27: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

Table 8: Empirical null rejection probabilities when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. Reported rates obtained from a simulation study with 10 000 replications for each couple  $(K, \psi)$ .

$K = 20$	$\alpha = 1\%$					$\alpha = 5\%$				
	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
$\psi = 0$	0.3	0.1	0.5	0.5	0.3	2.9	2.0	3.4	3.4	2.9
$\psi = 0.01$	1.1	0.5	1.3	1.3	0.5	5.1	3.6	5.6	5.6	4.1
$\psi = 0.02$	1.7	0.8	2.0	2.0	0.8	6.7	4.8	6.9	6.9	4.9
$\psi = 0.03$	2.2	1.3	2.4	2.4	0.9	7.4	5.7	7.7	7.7	5.4
$\psi = 0.04$	2.5	1.6	2.5	2.5	1.0	7.9	5.9	8.0	8.0	5.5
$\psi = 0.05$	2.6	1.6	2.6	2.6	1.1	8.2	6.0	8.1	8.1	5.6
$\psi = 0.06$	2.7	1.7	2.6	2.6	1.1	8.3	6.2	8.0	8.0	5.7
$\psi = 0.07$	2.8	1.8	2.7	2.7	1.1	8.3	6.0	8.0	8.0	5.7
$\psi = 0.08$	2.8	1.8	2.7	2.7	1.1	8.3	6.0	8.1	8.1	5.7
$\psi = 0.09$	2.8	1.7	2.7	2.7	1.1	8.3	6.0	8.0	8.0	5.7
$\psi = 0.1$	2.8	1.7	2.7	2.7	1.1	8.2	5.8	7.9	7.9	5.7
$K = 25$	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
	0.4	0.2	0.6	0.6	0.3	2.9	2.0	3.4	3.4	2.9
$\psi = 0.01$	1.2	0.8	1.6	1.6	0.7	5.0	4.0	5.2	5.2	4.1
$\psi = 0.02$	1.7	1.5	1.8	1.8	0.8	6.0	5.1	6.4	6.4	4.6
$\psi = 0.03$	1.9	1.8	1.9	1.9	0.9	6.7	5.9	6.6	6.6	4.7
$\psi = 0.04$	1.9	1.9	1.9	1.9	0.9	6.6	6.0	6.5	6.5	4.7
$\psi = 0.05$	1.9	1.9	2.0	2.0	0.9	6.7	5.8	6.6	6.6	4.7
$\psi = 0.06$	1.9	1.7	2.0	2.0	0.9	6.6	5.6	6.5	6.5	4.7
$\psi = 0.07$	1.9	1.6	1.9	1.9	0.9	6.6	5.2	6.5	6.5	4.7
$\psi = 0.08$	1.8	1.6	1.9	1.9	0.9	6.5	5.2	6.3	6.3	4.8
$\psi = 0.09$	1.8	1.3	1.8	1.8	0.9	6.5	4.9	6.3	6.3	4.8
$\psi = 0.1$	1.8	1.2	1.8	1.8	0.9	6.4	4.7	6.3	6.3	4.7
$K = 30$	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
	0.5	0.4	0.7	0.7	0.4	3.4	2.5	3.8	3.8	3.5
$\psi = 0.01$	1.4	1.1	1.6	1.6	0.8	5.6	5.0	6.1	6.1	4.7
$\psi = 0.02$	1.8	1.8	1.9	1.9	1.0	6.5	6.1	6.9	6.9	5.2
$\psi = 0.03$	1.9	1.9	2.1	2.1	1.1	6.7	6.3	6.9	6.9	5.3
$\psi = 0.04$	2.1	2.0	2.2	2.2	1.2	6.7	6.6	6.9	6.9	5.3
$\psi = 0.05$	2.1	1.9	2.1	2.1	1.1	6.8	6.0	6.6	6.6	5.3
$\psi = 0.06$	2.1	1.7	2.0	2.0	1.1	6.7	5.9	6.6	6.6	5.3
$\psi = 0.07$	2.1	1.6	2.1	2.1	1.1	6.7	5.4	6.5	6.5	5.2
$\psi = 0.08$	2.1	1.4	2.1	2.1	1.1	6.7	5.4	6.4	6.4	5.2
$\psi = 0.09$	2.1	1.3	2.1	2.1	1.1	6.6	5.3	6.4	6.4	5.3
$\psi = 0.1$	2.1	1.3	2.0	2.0	1.1	6.5	5.3	6.3	6.3	5.3

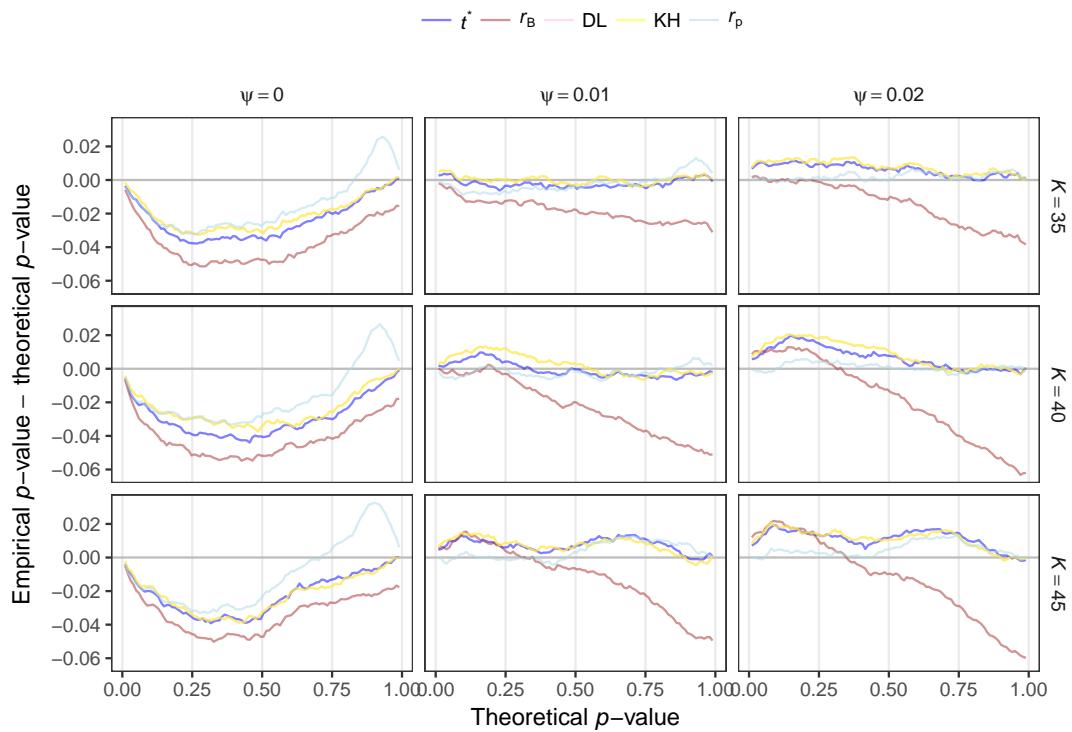


Figure 28: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

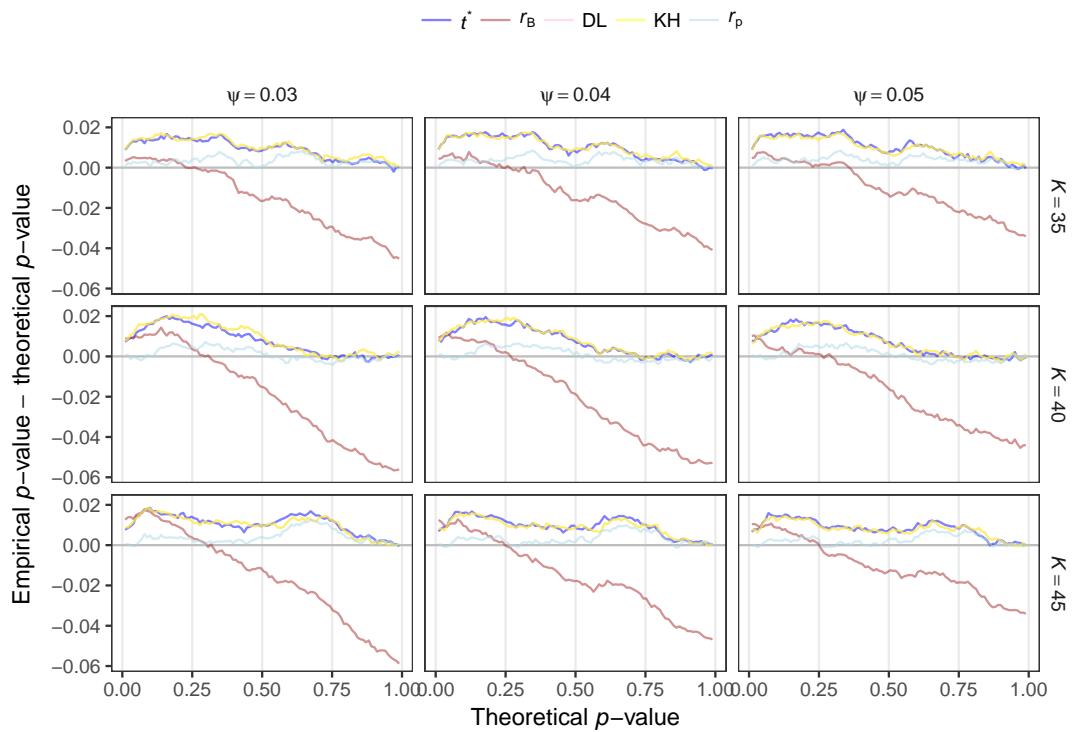


Figure 29: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

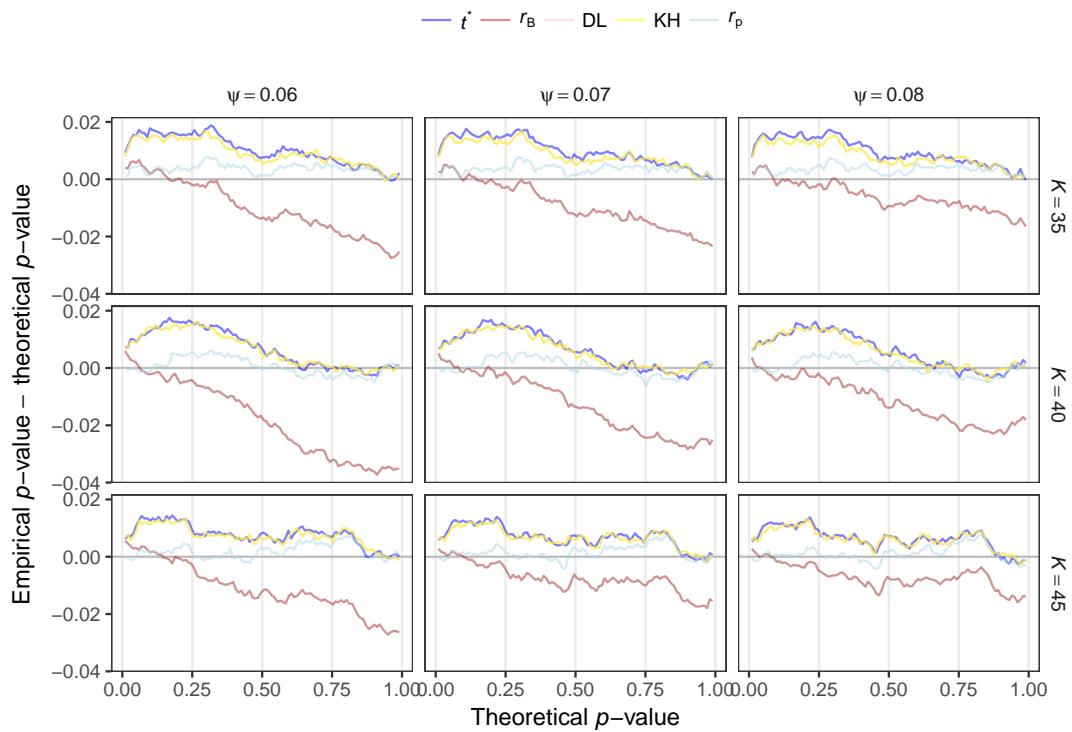


Figure 30: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

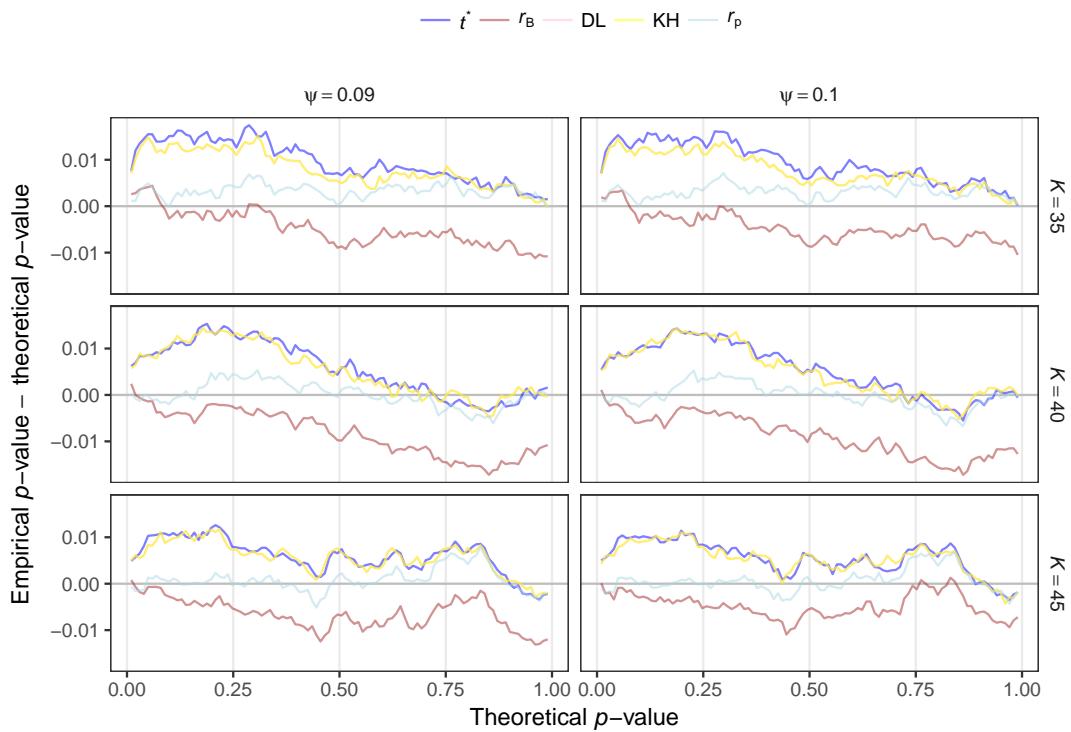


Figure 31: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

Table 9: Empirical null rejection probabilities when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. Reported rates obtained from a simulation study with 10 000 replications for each couple  $(K, \psi)$ .

$K = 35$	$\alpha = 1\%$					$\alpha = 5\%$				
	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
$\psi = 0$	0.7	0.4	0.8	0.8	0.6	3.8	3.1	4.1	4.1	3.8
$\psi = 0.01$	1.3	0.8	1.4	1.4	0.9	5.4	4.6	5.5	5.5	4.8
$\psi = 0.02$	1.7	1.2	1.8	1.8	1.0	6.1	5.1	6.1	6.1	5.1
$\psi = 0.03$	1.9	1.3	1.9	1.9	1.1	6.3	5.5	6.4	6.4	5.3
$\psi = 0.04$	1.9	1.4	1.9	1.9	1.1	6.6	5.6	6.6	6.6	5.4
$\psi = 0.05$	1.9	1.5	1.9	1.9	1.1	6.6	5.8	6.6	6.6	5.5
$\psi = 0.06$	1.9	1.4	1.8	1.8	1.1	6.5	5.6	6.5	6.5	5.4
$\psi = 0.07$	1.8	1.3	1.8	1.8	1.1	6.6	5.4	6.5	6.5	5.4
$\psi = 0.08$	1.8	1.3	1.8	1.8	1.1	6.6	5.4	6.6	6.6	5.5
$\psi = 0.09$	1.8	1.3	1.7	1.7	1.1	6.5	5.4	6.5	6.5	5.5
$\psi = 0.1$	1.7	1.2	1.7	1.7	1.1	6.5	5.3	6.4	6.4	5.4
$K = 40$	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
	$\psi = 0$	0.4	0.3	0.6	0.6	0.5	3.0	2.4	3.4	3.4
$\psi = 0.01$	1.2	1.0	1.3	1.3	0.8	5.4	4.8	5.8	5.8	4.5
$\psi = 0.02$	1.6	1.9	1.7	1.7	0.9	5.9	6.0	6.3	6.3	4.8
$\psi = 0.03$	1.7	1.9	1.8	1.8	1.0	6.2	6.0	6.4	6.4	4.9
$\psi = 0.04$	1.8	1.9	1.7	1.7	1.1	6.1	6.0	6.2	6.2	5.0
$\psi = 0.05$	1.8	2.0	1.7	1.7	1.1	6.0	5.7	6.1	6.1	5.0
$\psi = 0.06$	1.7	1.6	1.7	1.7	1.1	5.9	5.1	5.9	5.9	4.9
$\psi = 0.07$	1.7	1.5	1.6	1.6	1.1	5.9	5.2	5.9	5.9	5.0
$\psi = 0.08$	1.7	1.4	1.6	1.6	1.0	5.8	4.9	5.9	5.9	5.0
$\psi = 0.09$	1.6	1.2	1.6	1.6	1.0	5.9	4.8	5.8	5.8	5.0
$\psi = 0.1$	1.5	1.1	1.5	1.5	1.0	5.9	4.8	5.8	5.8	5.0
$K = 45$	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
	$\psi = 0$	0.6	0.5	0.8	0.8	0.5	3.5	3.0	3.6	3.6
$\psi = 0.01$	1.5	1.5	1.7	1.7	1.0	6.0	5.6	6.1	6.1	4.8
$\psi = 0.02$	1.7	2.2	1.8	1.8	1.0	6.4	6.5	6.5	6.5	4.9
$\psi = 0.03$	1.8	2.3	1.9	1.9	1.1	6.5	6.5	6.3	6.3	4.9
$\psi = 0.04$	1.7	2.2	1.8	1.8	1.1	6.3	6.1	6.2	6.2	4.9
$\psi = 0.05$	1.7	2.0	1.7	1.7	1.0	6.1	5.9	6.1	6.1	5.1
$\psi = 0.06$	1.6	1.6	1.7	1.7	1.0	6.0	5.3	5.9	5.9	5.0
$\psi = 0.07$	1.6	1.3	1.6	1.6	1.0	6.1	5.1	5.8	5.8	5.0
$\psi = 0.08$	1.5	1.3	1.6	1.6	1.0	6.1	5.0	5.9	5.9	5.0
$\psi = 0.09$	1.5	1.1	1.5	1.5	0.9	6.0	5.0	5.9	5.9	5.1
$\psi = 0.1$	1.5	1.0	1.4	1.4	0.9	6.0	4.9	5.8	5.8	5.1

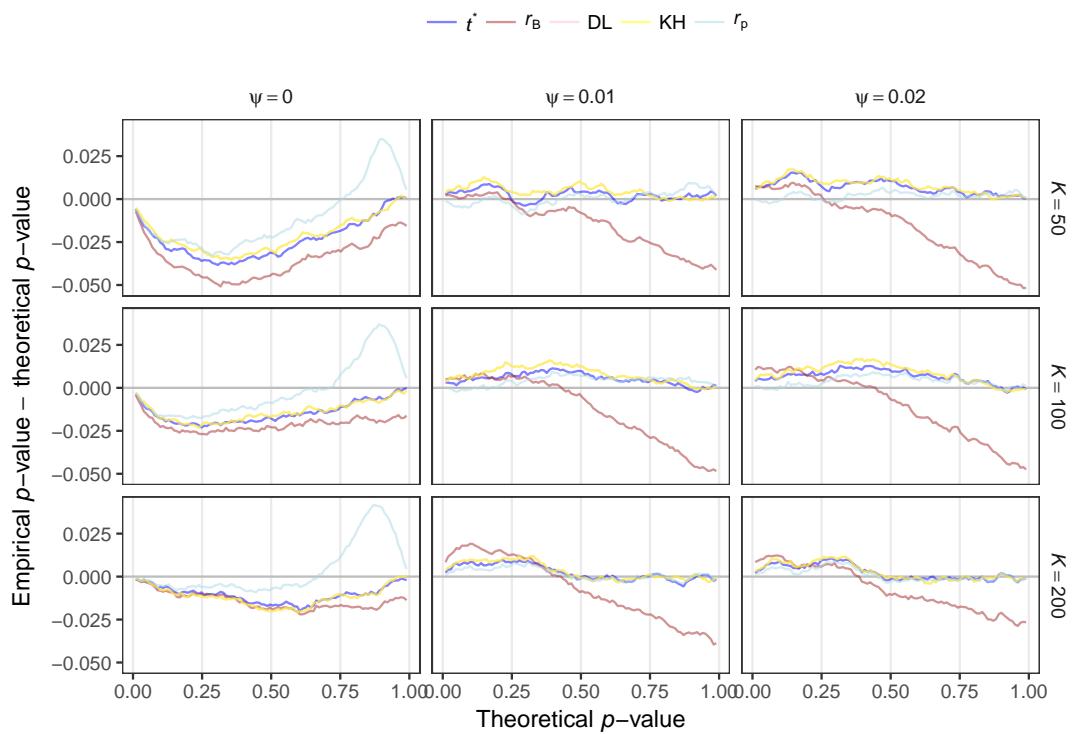


Figure 32: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

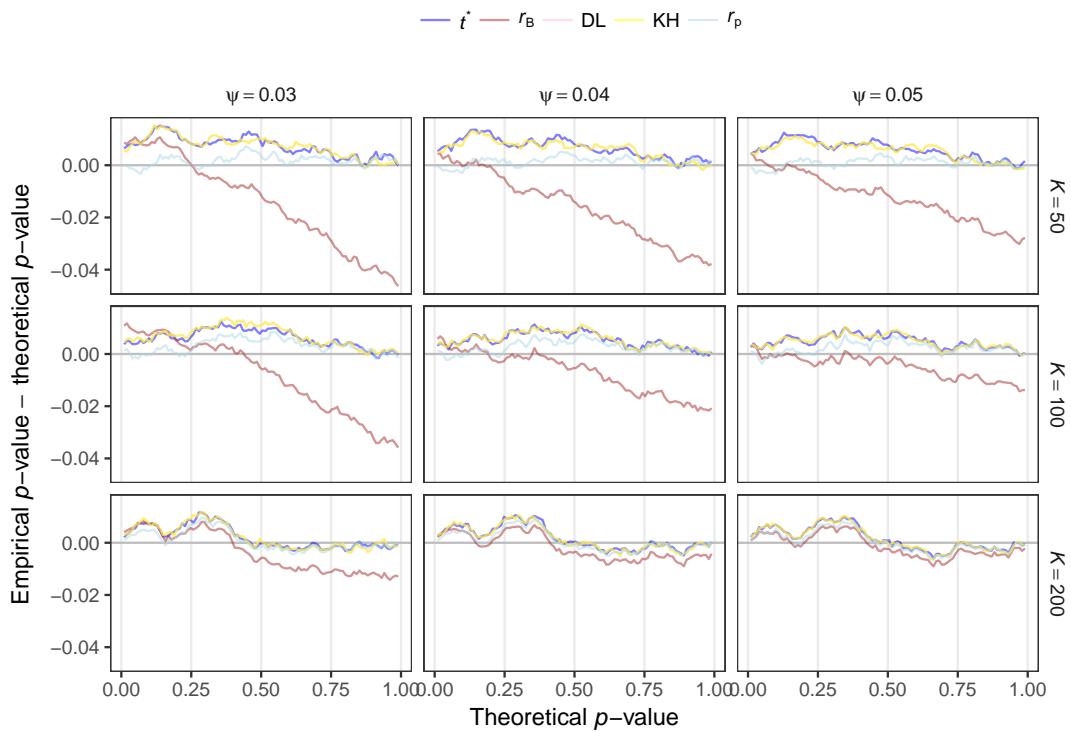


Figure 33: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

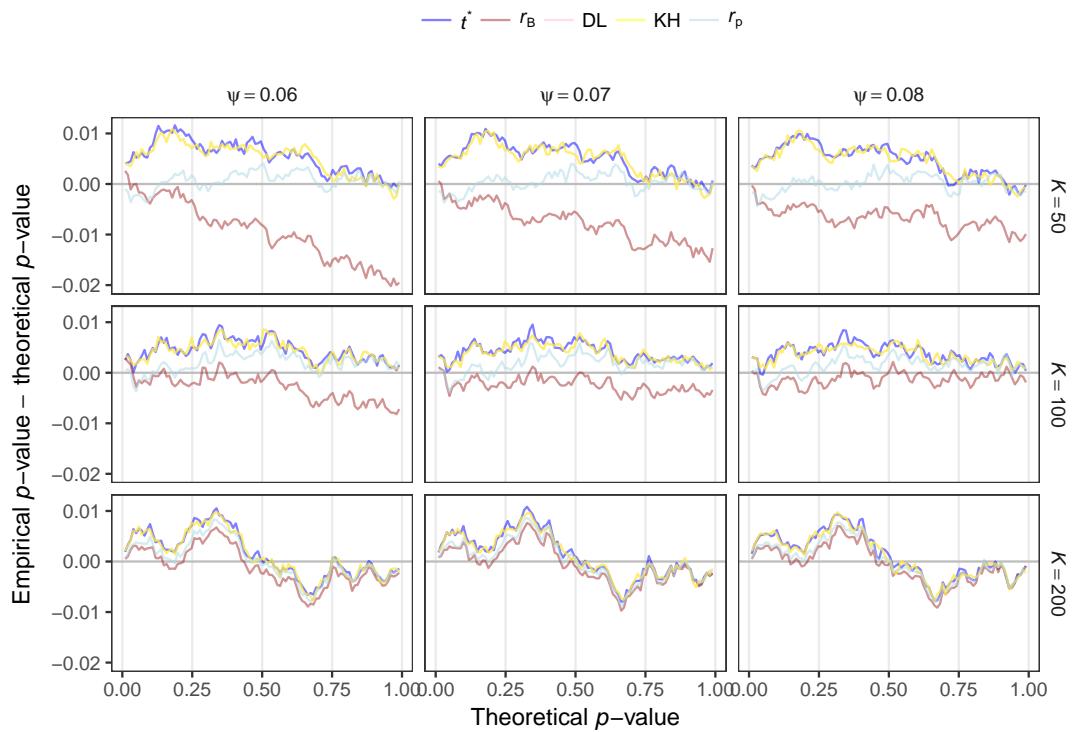


Figure 34: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

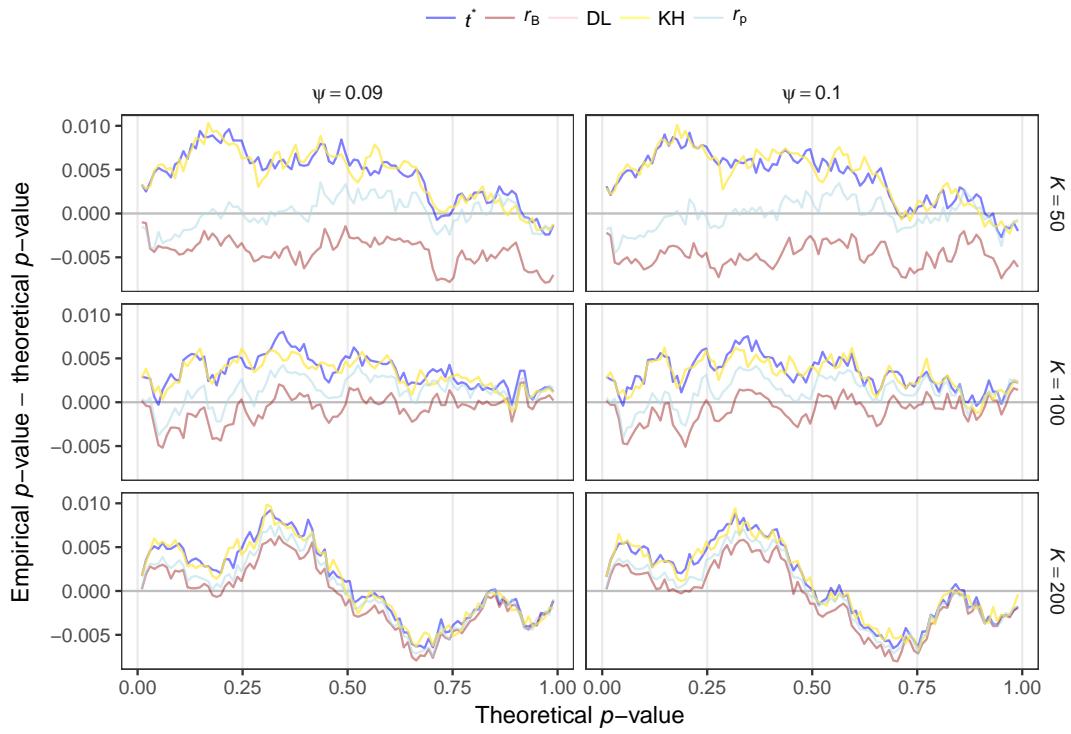


Figure 35: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 10 000 replications for each couple  $(\psi, K)$ .

Table 10: Empirical null rejection probabilities when testing  $H_0 : \beta = 0.5$  against  $H_1 : \beta \neq 0.5$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t^*$ ,  $r_B$ , DL, KH and  $r_p$  in the meta-analysis model of § 4.3. Reported rates obtained from a simulation study with 10 000 replications for each couple  $(K, \psi)$ .

$K = 50$	$\alpha = 1\%$					$\alpha = 5\%$				
	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
$\psi = 0$	0.4	0.3	0.5	0.5	0.4	3.3	2.8	3.6	3.6	3.3
$\psi = 0.01$	1.3	1.2	1.4	1.4	0.9	5.4	5.1	5.7	5.7	4.6
$\psi = 0.02$	1.6	1.8	1.6	1.6	1.0	5.8	5.7	6.0	6.0	4.9
$\psi = 0.03$	1.6	1.9	1.6	1.6	1.0	5.8	6.1	5.9	5.9	4.8
$\psi = 0.04$	1.5	1.4	1.5	1.5	0.9	5.7	5.4	5.8	5.8	4.8
$\psi = 0.05$	1.4	1.4	1.5	1.5	0.9	5.5	5.2	5.7	5.7	4.8
$\psi = 0.06$	1.4	1.3	1.4	1.4	0.9	5.5	5.0	5.6	5.6	4.8
$\psi = 0.07$	1.4	1.1	1.3	1.3	0.9	5.6	4.8	5.5	5.5	4.8
$\psi = 0.08$	1.4	1.0	1.4	1.4	0.8	5.5	4.7	5.5	5.5	4.8
$\psi = 0.09$	1.3	0.9	1.3	1.3	0.9	5.5	4.7	5.5	5.5	4.8
$\psi = 0.1$	1.3	0.8	1.3	1.3	0.8	5.4	4.5	5.5	5.5	4.7
$K = 100$	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
	$\psi = 0$	0.6	0.6	0.7	0.7	0.7	3.8	3.5	4.0	4.0
$\psi = 0.01$	1.3	1.5	1.4	1.4	1.1	5.1	5.6	5.7	5.7	4.8
$\psi = 0.02$	1.4	2.1	1.5	1.5	1.1	5.5	6.1	5.8	5.8	4.9
$\psi = 0.03$	1.4	2.1	1.5	1.5	1.1	5.4	5.9	5.5	5.5	4.9
$\psi = 0.04$	1.3	1.6	1.4	1.4	1.1	5.2	5.2	5.3	5.3	4.8
$\psi = 0.05$	1.3	1.3	1.3	1.3	1.1	5.2	4.8	5.2	5.2	4.6
$\psi = 0.06$	1.2	1.3	1.3	1.3	1.0	5.2	4.7	5.2	5.2	4.6
$\psi = 0.07$	1.3	1.1	1.3	1.3	1.0	5.1	4.5	5.1	5.1	4.6
$\psi = 0.08$	1.3	1.0	1.3	1.3	1.0	5.0	4.6	5.2	5.2	4.6
$\psi = 0.09$	1.3	1.0	1.3	1.3	1.0	5.0	4.5	5.1	5.1	4.6
$\psi = 0.1$	1.3	1.0	1.3	1.3	1.1	4.9	4.5	5.2	5.2	4.6
$K = 200$	$t^*$	$r_B$	DL	KH	$r_p$	$t^*$	$r_B$	DL	KH	$r_p$
	$\psi = 0$	0.9	0.8	0.9	0.9	0.9	4.7	4.6	4.7	4.7
$\psi = 0.01$	1.2	1.8	1.4	1.4	1.1	5.8	6.6	5.8	5.8	5.5
$\psi = 0.02$	1.2	1.8	1.3	1.3	1.1	5.7	6.2	5.7	5.7	5.4
$\psi = 0.03$	1.2	1.4	1.3	1.3	1.1	5.6	5.7	5.7	5.7	5.3
$\psi = 0.04$	1.3	1.2	1.3	1.3	1.1	5.5	5.5	5.7	5.7	5.3
$\psi = 0.05$	1.2	1.1	1.2	1.2	1.1	5.5	5.4	5.6	5.6	5.3
$\psi = 0.06$	1.2	1.1	1.2	1.2	1.1	5.5	5.3	5.6	5.6	5.4
$\psi = 0.07$	1.2	1.1	1.2	1.2	1.1	5.5	5.3	5.5	5.5	5.4
$\psi = 0.08$	1.1	1.1	1.2	1.2	1.1	5.5	5.3	5.5	5.5	5.3
$\psi = 0.09$	1.2	1.0	1.2	1.2	1.0	5.5	5.3	5.5	5.5	5.3
$\psi = 0.1$	1.2	1.0	1.2	1.2	1.0	5.5	5.3	5.5	5.5	5.3

## 1.5 Beta regression for reading skills

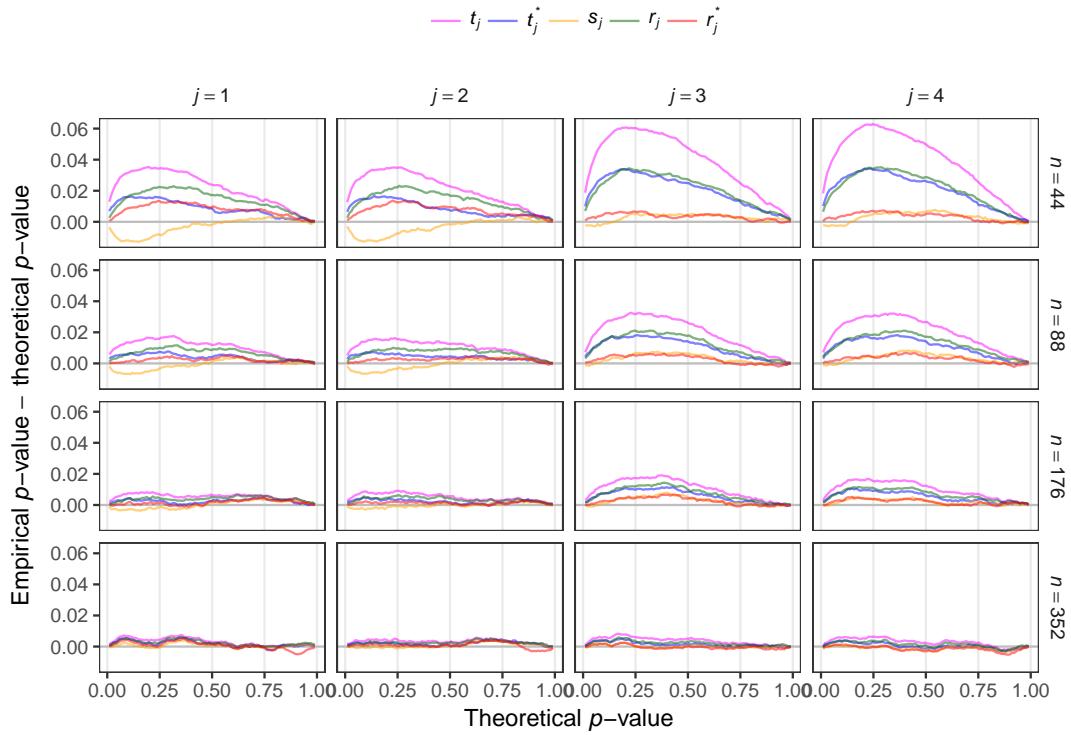


Figure 36: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \beta_j = \beta_{j0}$  against  $H_1 : \beta_j \neq \beta_{j0}$  based on the normal approximation to the distribution of  $t_j$ ,  $t_j^*$ ,  $s_j$ ,  $r_j$  and  $r_j^*$  ( $j = 1, \dots, 4$ ) in the beta regression model for the reading skills dataset of Example 1.1. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 50 000 replications for each  $n \in \{44, 88, 176, 352\}$ .

Table 11: Empirical null rejection probabilities when testing  $H_0 : \beta_j = \beta_{j0}$  against  $H_1 : \beta_j \neq \beta_{j0}$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t_j$ ,  $t_j^*$ ,  $s_j$ ,  $r_j$  and  $r_j^*$  ( $j = 1, \dots, 4$ ) in the beta regression model for the reading skills dataset of Example 1.1. Reported rates obtained from a simulation study with 50 000 replications for each  $n \in \{44, 88, 176, 352\}$ .

$n = 44$	$\alpha = 1\%$					$\alpha = 5\%$				
	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$
$j = 1$	2.3	1.7	0.7	1.3	1.1	7.6	6.3	3.9	6.0	5.4
$j = 2$	2.3	1.7	0.7	1.3	1.1	7.6	6.3	3.9	6.0	5.5
$j = 3$	2.9	2.0	0.8	1.7	1.1	9.0	7.2	4.7	6.8	5.3
$j = 4$	2.8	2.0	0.8	1.6	1.0	8.8	7.2	4.8	6.8	5.2
$n = 88$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$
$j = 1$	1.6	1.3	0.8	1.1	1.0	6.1	5.5	4.4	5.4	5.0
$j = 2$	1.5	1.3	0.8	1.1	1.0	6.1	5.6	4.4	5.4	5.1
$j = 3$	1.7	1.4	0.9	1.3	1.1	6.8	6.0	4.9	6.0	5.2
$j = 4$	1.7	1.4	0.9	1.3	1.0	6.7	6.0	4.9	5.8	5.1
$n = 176$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$
$j = 1$	1.3	1.2	0.9	1.1	1.0	5.5	5.3	4.7	5.2	5.0
$j = 2$	1.3	1.2	0.9	1.1	1.1	5.6	5.4	4.8	5.3	5.1
$j = 3$	1.3	1.2	1.0	1.2	1.0	5.8	5.5	4.9	5.5	5.1
$j = 4$	1.3	1.2	1.0	1.2	1.0	6.0	5.6	5.1	5.6	5.2
$n = 352$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$
$j = 1$	1.1	1.1	1.0	1.1	1.0	5.5	5.4	5.1	5.3	5.2
$j = 2$	1.2	1.1	1.0	1.1	1.0	5.4	5.3	4.9	5.2	5.1
$j = 3$	1.1	1.0	0.9	1.0	1.0	5.5	5.4	5.1	5.3	5.2
$j = 4$	1.0	1.0	0.9	1.0	0.9	5.5	5.2	5.0	5.3	5.0

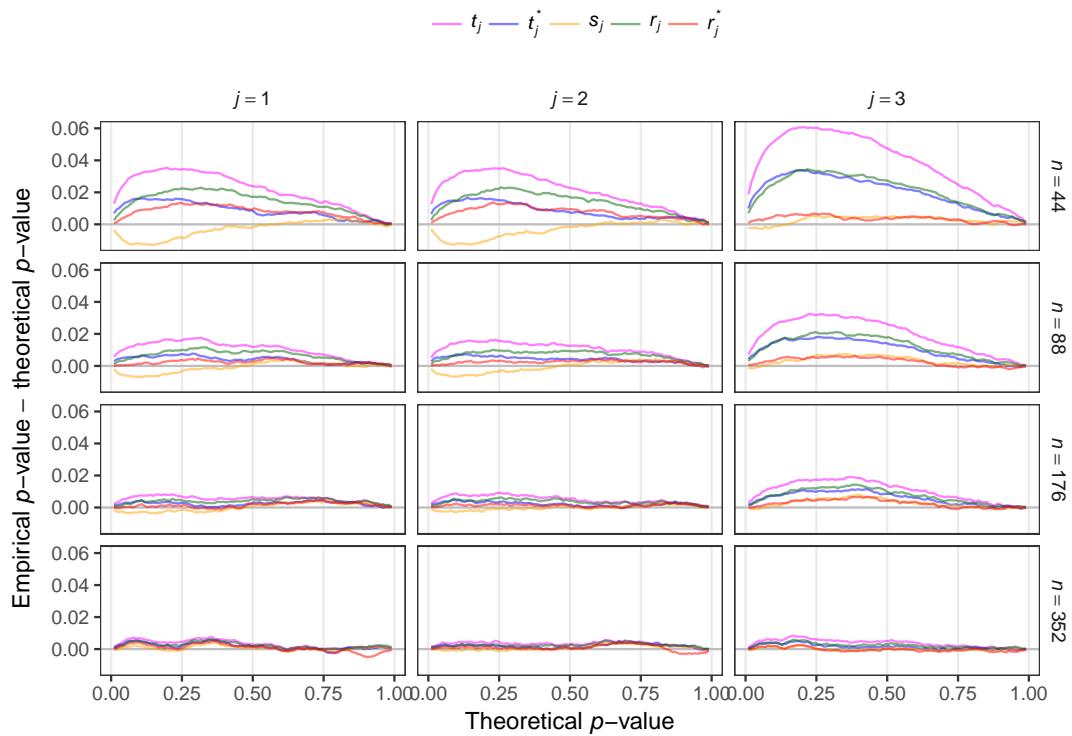


Figure 37: Difference between empirical and theoretical null  $p$ -values when testing  $H_0 : \gamma_j = \gamma_{j0}$  against  $H_1 : \gamma_j \neq \gamma_{j0}$  based on the normal approximation to the distribution of  $t_j$ ,  $t_j^*$ ,  $s_j$ ,  $r_j$  and  $r_j^*$  ( $j = 1, 2, 3$ ) in the beta regression model for the reading skills dataset of Example 1.1. The reference zero line (grey) indicates equality between empirical and uniform  $p$ -values. Results obtained from a simulation study with 50 000 replications for each  $n \in \{44, 88, 176, 352\}$ .

Table 12: Empirical null rejection probabilities when testing  $H_0 : \gamma_j = \gamma_{j0}$  against  $H_1 : \gamma_j \neq \gamma_{j0}$  at nominal level  $\alpha \in \{1\%, 5\%\}$ , based on the normal approximation to the distribution of  $t_j$ ,  $t_j^*$ ,  $s_j$ ,  $r_j$  and  $r_j^*$  ( $j = 1, 2, 3$ ) in the beta regression model for the reading skills dataset of Example 1.1. Reported rates obtained from a simulation study with 50 000 replications for each  $n \in \{44, 88, 176, 352\}$ .

$n = 44$	$\alpha = 1\%$					$\alpha = 5\%$					
	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	
$j = 1$	7.1	2.4	0.7	4.4	1.0	17.0	7.9	6.9	13.6	5.0	
$j = 2$	3.9	3.4	0.9	1.4	0.9	10.9	10.0	4.9	6.1	4.8	
$j = 3$	5.6	4.3	0.5	1.4	0.9	14.0	11.4	3.6	6.4	4.8	
$n = 88$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	
	$j = 1$	3.5	1.4	0.8	2.4	0.9	10.4	6.1	5.9	9.0	4.9
	$j = 2$	2.1	1.9	0.9	1.2	0.9	7.6	7.2	4.9	5.5	4.8
	$j = 3$	2.6	2.2	0.8	1.3	1.0	8.9	7.9	4.4	6.0	5.0
$n = 176$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	
	$j = 1$	2.2	1.3	1.0	1.7	1.1	7.8	5.7	5.7	7.1	5.0
	$j = 2$	1.5	1.4	1.0	1.1	1.0	6.2	6.0	4.9	5.1	4.8
	$j = 3$	1.6	1.5	0.9	1.1	1.0	6.7	6.3	4.8	5.4	4.8
$n = 352$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	$t_j$	$t_j^*$	$s_j$	$r_j$	$r_j^*$	
	$j = 1$	1.5	1.0	0.9	1.2	0.9	6.1	5.2	5.2	5.8	4.9
	$j = 2$	1.2	1.1	1.0	1.0	1.0	5.4	5.3	4.9	5.0	4.8
	$j = 3$	1.2	1.2	0.9	1.0	0.9	5.7	5.5	4.7	5.1	4.8

## References

- Bellio, R. and D. Pierce (2016). *likelihoodAsy: Functions for Likelihood Asymptotics*. R package version 0.45.
- DerSimonian, R. and N. Laird (1986). Meta-analysis in clinical trials. *Controlled clinical trials* 7, 177–188.
- Fraser, D. A. S., L. Faye, Y. Cai, K. Ji, M. Mallo, and R. Thinniyam (2008). Is  $r^*$  linear in  $r$ ? *Journal of Statistical Research* 42, 1–13.
- Guolo, A., C. Varin, et al. (2012). The R package metaLik for likelihood inference in meta-analysis. *Journal of Statistical Software* 50, 1–14.
- Huizenga, H. M., I. Visser, and C. V. Dolan (2011). Testing overall and moderator effects in random effects meta-regression. *British Journal of Mathematical and Statistical Psychology* 64, 1–19.
- Knapp, G. and J. Hartung (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in medicine* 22, 2693–2710.
- Kosmidis, I., A. Guolo, and C. Varin (2017). Improving the accuracy of likelihood-based inference in meta-analysis and meta-regression. *Biometrika* 104, 489–496.