Simulation results in the main text of the Double calibration

September 2, 2024

1 Case I: Sparse PS & Dense OR

First, we consider the case where the PS is a sparse logistic-linear model and the OR is a dense nonlinear model:

$$T|X = x \sim \text{Bernoulli}\left(\pi(x) = \psi(x^\top\gamma)\right),$$
$$Y|X = x, T = t \sim (r(x,t), 1),$$

where $\psi(z) \equiv 1/(1+\exp(-z))$ in this section. In particular, we choose γ to satisfy the following: $\|\gamma\|_2 = 1$, $\gamma_j \sim \mathrm{Uniform}([1,2])$ for $j = 1, \cdots, s$ and $\gamma_{s+1} = \cdots = \gamma_p = 0$ so only the first s coordinates of γ are active. For the outcome model $r(x,t)$, we first transform the covariates X nonlinearly as follows:

$$\tilde{X}_1 \equiv \mathtt{bs}(X_1, 100)^\top \cdot (1, 1/2, 1/3, \cdots, 1/100), \tilde{X}_2 \equiv 2/(1+\exp(-X_2)),$$
$$\tilde{X}_3 \equiv \exp(X_3/2), \tilde{X}_4 \equiv X_4/(1+\exp(X_3)), \tilde{X}_5 = X_4X_5/10,$$

and then scale $\tilde{X}_2, \cdots, \tilde{X}_5$ to have zero mean and unit variance. Here $\mathtt{bs}(x, 100)$ denotes the B-spline transformation of x with 100 degrees-of-freedom. We then define the outcome model $r(x,t)$ as

$$r(x,t) \equiv (|\tilde{x}_1 + \tilde{x}_2 - \tilde{x}_3/2 + \tilde{x}_4/3 - \tilde{x}_5/4| + 0.05)^{-1} + x^\top\beta - t,$$

where $\|\beta\|_2 \equiv 1$ and $\beta_j \propto j^{-1}$ for $j = 1, \cdots, p$.

The simulation results are displayed in Figure 1 and Table 1.

Tables 1 below summarize the coverage probabilities and the lengths of 95% confidence intervals, Mean/Median Absolute Biases, standard errors, and Root-Mean-Squared-Errors (RMSE) of different estimators for the Sparse PS & Dense OR simulation setting from the main text.

Table 1: Coverage probability of 95% confidence intervals (Coverage), length of 95% confidence intervals (CI length), Mean Absolute Bias (MAB†), Median Absolute Bias (MAB‡), standard error (std), root-mean-squared error (RMSE) for estimating τ under the *sparse PS & dense OR* setting from the main text.

Setting	**measure**	**g-formula**	**IPW**	**AIPW**	**TMLE**	**ARB**	**hdCBPS**	**RCAL**	**RCAL***	**DCal**
$s = 10$, $p = 400$, $n = 200$	MAB†	1.34	1.06	0.98	0.70	0.53	0.78	0.78	0.80	0.47
	MAB‡	1.39	1.08	1.00	0.71	0.50	0.80	0.80	0.82	0.41
	std	0.02	0.58	0.31	0.67	0.49	0.50	0.32	0.33	0.53
	RMSE	0.93	0.85	0.72	0.71	0.53	0.69	0.60	0.61	0.52

Table 1 continued from previous page

	CP	0.00%	57.00%	15.00%	90.50%	92.00%	53.50%	32.00%	32.00%	96.00%
	Length	0.06	2.28	1.21	2.64	1.92	1.98	1.27	1.29	2.09
$s = 10$, $p = 800$, $n = 400$	MAB†	1.35	1.02	0.92	0.56	0.48	0.60	0.77	0.79	0.40
	MAB‡	1.41	1.02	0.93	0.54	0.44	0.56	0.80	0.80	0.33
	std	0.01	0.45	0.25	0.53	0.37	0.46	0.24	0.25	0.43
	RMSE	0.91	0.76	0.65	0.53	0.42	0.57	0.55	0.56	0.42
	CP	0.50%	28.50%	8.50%	93.50%	82.00%	60.50%	15.50%	16.00%	91.50%
	Length	0.05	1.77	0.97	2.06	1.45	1.81	0.94	0.97	1.69
$s = 10$, $p = 1000$, $n = 800$	MAB†	1.11	0.91	0.70	0.32	0.40	0.32	0.61	0.64	0.33
	MAB‡	1.11	0.93	0.71	0.28	0.39	0.23	0.59	0.58	0.30
	std	0.02	0.35	0.19	0.41	0.27	0.44	0.20	0.19	0.33
	RMSE	0.76	0.67	0.50	0.37	0.34	0.40	0.44	0.46	0.34
	CP	0.00%	18.00%	11.50%	97.50%	72.00%	88.00%	22.00%	15.00%	89.00%
	Length	0.09	1.38	0.75	1.62	1.05	1.71	0.77	0.73	1.28
$s = 10$, $p = 1000$, $n = 1600$	MAB†	0.87	0.81	0.51	0.20	0.37	0.20	0.45	0.45	0.34
	MAB‡	0.86	0.82	0.51	0.18	0.37	0.14	0.43	0.45	0.34
	std	0.03	0.29	0.15	0.30	0.19	0.34	0.15	0.15	0.24
	RMSE	0.59	0.59	0.37	0.26	0.29	0.28	0.33	0.33	0.29
	CP	0.00%	15.50%	18.00%	99.50%	49.00%	98.50%	24.12%	25.00%	75.50%
	Length	0.10	1.12	0.59	1.19	0.73	1.33	0.58	0.59	0.93
$s = 20$, $p = 400$, $n = 200$	MAB†	1.44	1.15	1.08	0.66	0.51	0.84	0.84	0.88	0.47
	MAB‡	1.48	1.15	1.11	0.62	0.48	0.87	0.82	0.87	0.36
	std	0.01	0.56	0.29	0.96	0.52	0.49	0.32	0.32	0.56
	RMSE	1.00	0.90	0.78	0.84	0.53	0.72	0.63	0.66	0.53
	CP	0.00%	38.00%	11.50%	98.50%	95.00%	47.00%	32.50%	30.50%	96.00%
	Length	0.05	2.21	1.15	3.75	2.04	1.92	1.27	1.25	2.18
$s = 20$, $p = 800$, $n = 400$	MAB†	1.36	1.06	0.94	0.40	0.40	0.59	0.70	0.72	0.34
	MAB‡	1.38	1.05	0.94	0.35	0.36	0.53	0.69	0.74	0.29
	std	0.02	0.41	0.22	0.71	0.38	0.43	0.23	0.24	0.42
	RMSE	0.92	0.77	0.66	0.58	0.39	0.53	0.50	0.52	0.38
	CP	0.00%	15.00%	6.50%	97.50%	92.00%	64.00%	22.00%	23.00%	93.00%
	Length	0.06	1.61	0.87	2.79	1.49	1.68	0.92	0.93	1.65
$s = 20$, $p = 1000$, $n = 800$	MAB†	1.26	0.98	0.80	0.29	0.33	0.36	0.62	0.64	0.28
	MAB‡	1.27	0.98	0.81	0.25	0.29	0.27	0.59	0.60	0.23
	std	0.02	0.33	0.18	0.58	0.28	0.41	0.17	0.18	0.34
	RMSE	0.86	0.71	0.56	0.46	0.31	0.41	0.44	0.45	0.32
	CP	0.00%	10.00%	7.50%	100.00%	85.50%	82.50%	12.00%	13.50%	95.50%
	Length	0.08	1.29	0.72	2.28	1.11	1.61	0.68	0.69	1.34
$s = 20$, $p = 1000$, $n = 1600$	MAB†	1.03	0.89	0.61	0.32	0.26	0.25	0.46	0.48	0.22
	MAB‡	1.00	0.90	0.60	0.29	0.24	0.21	0.46	0.46	0.19
	std	0.02	0.26	0.14	0.46	0.20	0.34	0.14	0.14	0.25
	RMSE	0.70	0.64	0.43	0.40	0.23	0.30	0.34	0.34	0.24
	CP	0.00%	7.50%	6.00%	99.50%	82.50%	92.50%	20.92%	19.50%	93.50%

Table 1 continued from previous page

	Length	0.09	1.03	0.56	1.79	0.77	1.32	0.57	0.56	0.97
$s = 50$, $p = 400$, $n = 200$	MAB†	1.24	1.09	1.00	0.60	0.51	0.71	0.77	0.79	0.47
	MAB‡	1.25	1.09	1.04	0.56	0.45	0.67	0.78	0.81	0.40
	std	0.02	0.55	0.27	1.26	0.52	0.45	0.32	0.32	0.54
	RMSE	0.86	0.86	0.73	1.01	0.53	0.62	0.59	0.60	0.52
	CP	0.50%	43.00%	10.00%	99.00%	94.00%	62.00%	31.00%	29.50%	96.50%
	Length	0.09	2.15	1.08	4.96	2.04	1.77	1.26	1.26	2.11
$s = 50$, $p = 800$, $n = 400$	MAB†	1.14	0.98	0.85	0.42	0.37	0.44	0.58	0.61	0.34
	MAB‡	1.17	0.98	0.87	0.36	0.31	0.36	0.58	0.63	0.27
	std	0.02	0.41	0.21	0.97	0.38	0.38	0.24	0.24	0.41
	RMSE	0.77	0.72	0.59	0.74	0.37	0.42	0.43	0.45	0.38
	CP	0.00%	17.00%	7.00%	99.50%	94.00%	79.50%	37.00%	31.50%	95.50%
	Length	0.10	1.60	0.82	3.82	1.50	1.48	0.93	0.93	1.61
$s = 50$, $p = 1000$, $n = 800$	MAB†	1.06	0.92	0.73	0.40	0.25	0.32	0.49	0.51	0.23
	MAB‡	1.07	0.93	0.73	0.32	0.20	0.30	0.49	0.50	0.18
	std	0.03	0.31	0.16	0.78	0.28	0.30	0.18	0.18	0.32
	RMSE	0.73	0.67	0.52	0.62	0.27	0.31	0.36	0.37	0.28
	CP	0.00%	8.50%	4.00%	100.00%	93.50%	89.50%	23.00%	20.00%	97.50%
	Length	0.11	1.22	0.63	3.05	1.11	1.16	0.69	0.70	1.25
$s = 50$, $p = 1000$, $n = 1600$	MAB†	0.94	0.84	0.60	0.48	0.19	0.26	0.45	0.46	0.18
	MAB‡	0.94	0.88	0.61	0.45	0.16	0.25	0.47	0.48	0.16
	std	0.03	0.25	0.13	0.62	0.20	0.23	0.13	0.13	0.24
	RMSE	0.64	0.60	0.42	0.56	0.20	0.25	0.32	0.33	0.21
	CP	0.00%	7.00%	5.50%	98.00%	94.50%	85.50%	13.50%	12.00%	98.50%
	Length	0.11	0.99	0.51	2.44	0.80	0.91	0.51	0.51	0.94

2 Case II: Dense PS & Sparse OR

Next, we consider the case where the PS is a dense nonlinear model bounded between 0 and 1 and the OR is a sparse linear model:

$$T|X = x \sim \text{Bernoulli}(\pi(x)),$$
$$Y|X = x, T = t \sim (r(x,t), 1).$$

Here we take the PS model to be

$$\pi(X) \equiv \min\{\max\{\psi(\check{X}_1 - \frac{1}{2}\check{X}_2 + \frac{1}{4}\check{X}_3 - \frac{1}{8}\check{X}_4 + \gamma^\top X), 0.05\}, 0.95\},$$

where $\check{X}_1 \coloneqq e^{0.5X_1}$, $\check{X}_2 \coloneqq 10 + X_2/(1+e^{X_1})$, $\check{X}_3 \coloneqq (0.05X_1X_3 + 0.6)^2$, $\check{X}_4 \coloneqq (X_2 + X_4 + 10)^2$, and $\gamma_j \propto j^{-1}$ with $\|\gamma\|_2 \equiv 2$. We take the OR model to be

$$r(X,t) \equiv \beta^\top X - 1 + 2t$$

where we first draw $\beta_j \sim \text{Uniform}([1,2])$ for $j = 1, \cdots, s$ and set $\beta_j \equiv 0$ for $j = s+1, \cdots, p$. We eventually normalize β such that $\|\beta\|_2 = 1$.

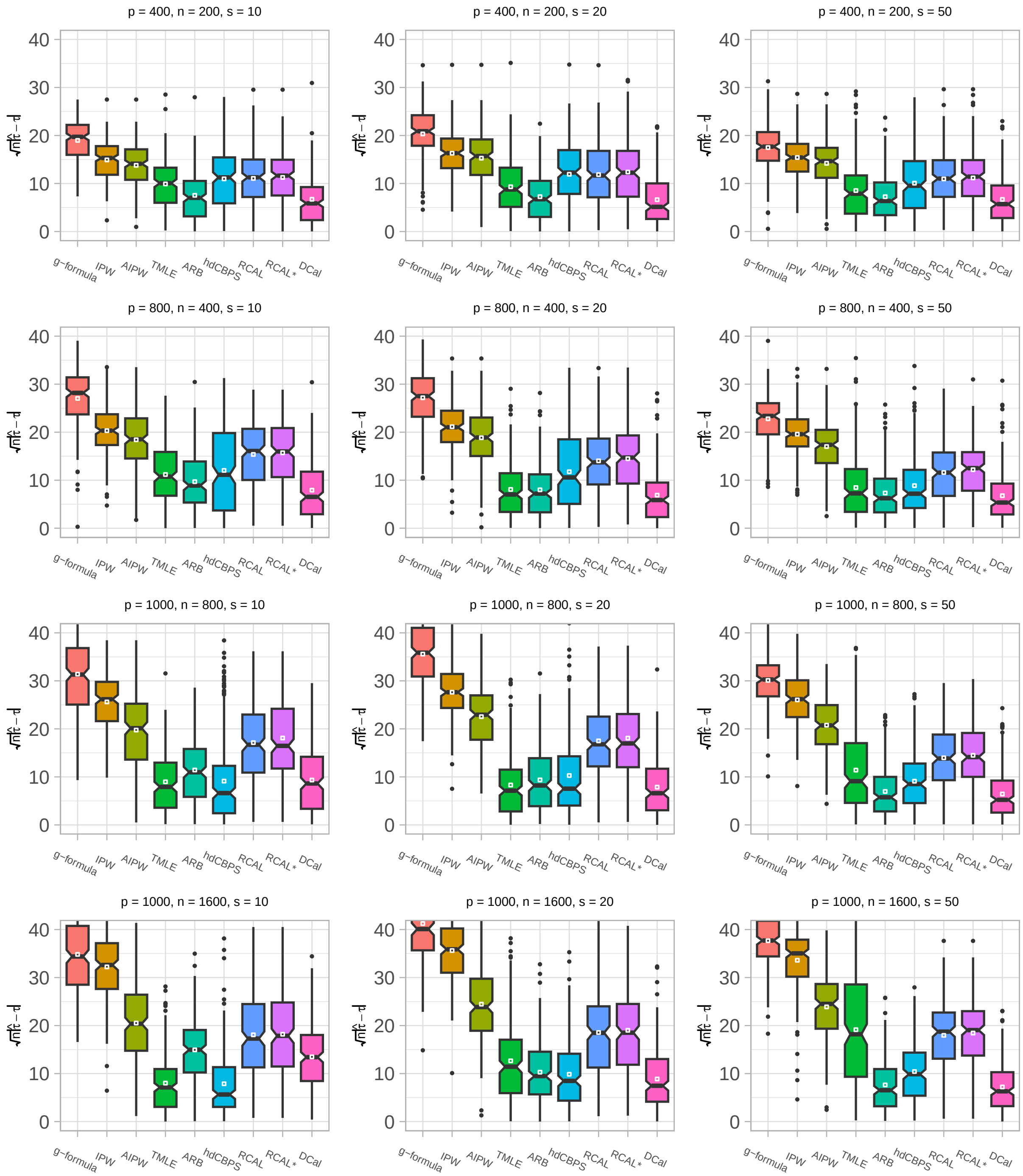

Figure 1: Boxplots of the $\sqrt{n}$-scaled estimation error $\sqrt{n}|\hat{\tau} - \tau|$ under different sample size n, dimension p and sparsity level s under the *Sparse PS & Dense OR* setting. The white dots correspond to mean.

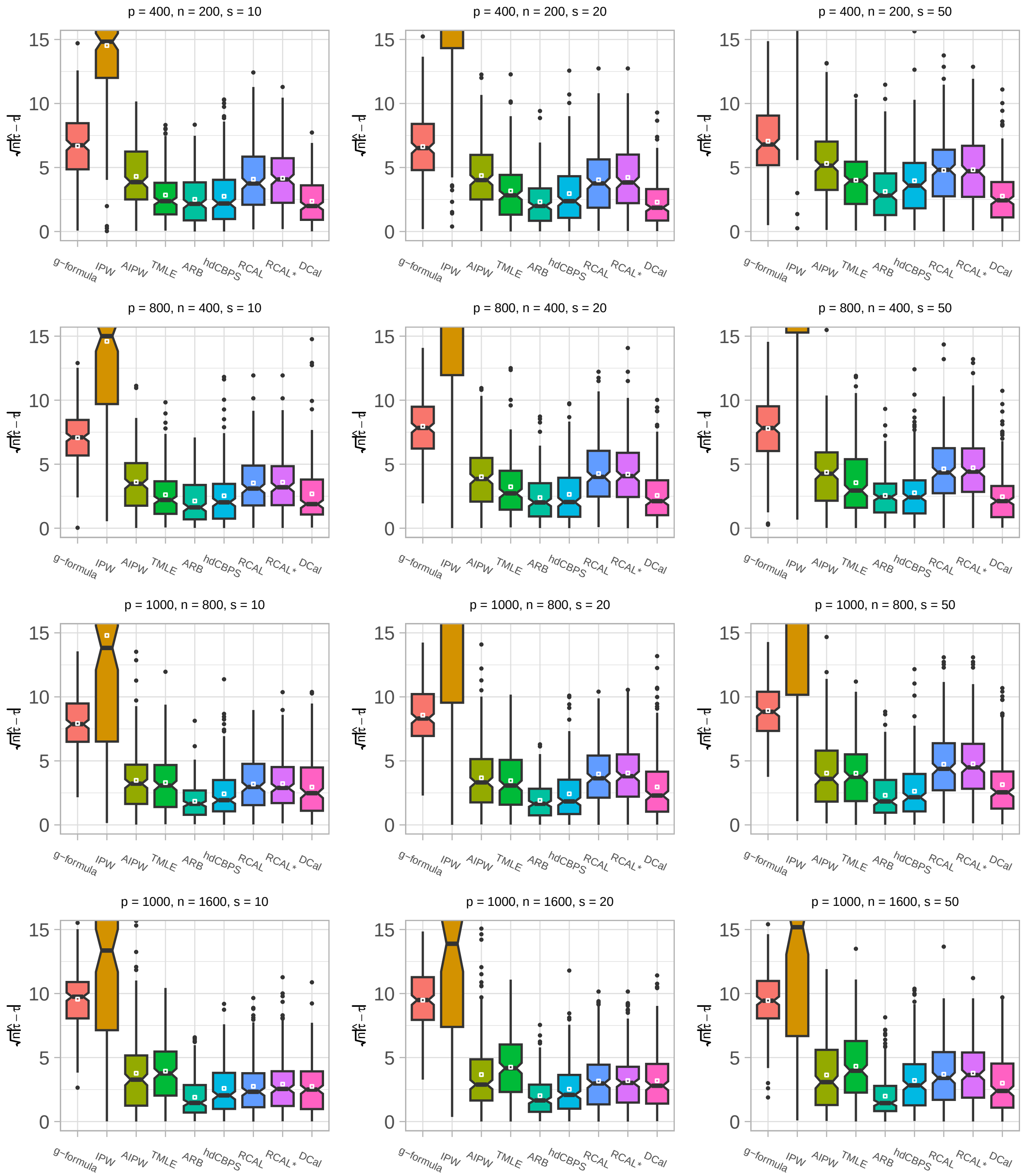

Figure 2: Boxplots of the $\sqrt{n}$-scaled estimation error $\sqrt{n}|\hat{\tau} - \tau|$ under different sample size n, dimension p and sparsity level s under the *Dense PS & Sparse OR* setting. The white dots correspond to mean.

The simulation results are displayed in Figure 2, and Table 2.

Table 2: Coverage probability of 95% confidence intervals (Coverage), length of 95% confidence intervals (CI length), Mean Absolute Bias (MAB†), Median Absolute Bias (MAB‡), standard error (std), root-mean-squared error (RMSE) for estimating τ under the *dense PS & sparse OR* setting from the main text.

Setting	measure	g-formula	IPW	AIPW	TMLE	ARB	hdCBPS	RCAL	RCAL*	DCal
$s = 10$, $p = 400$, $n = 200$	MAB†	1.34	1.06	0.98	0.70	0.53	0.78	0.78	0.80	0.47
	MAB‡	1.39	1.08	1.00	0.71	0.50	0.80	0.80	0.82	0.41
	std	0.02	0.58	0.31	0.67	0.49	0.50	0.32	0.33	0.53
	RMSE	0.93	0.85	0.72	0.71	0.53	0.69	0.60	0.61	0.52
	CP	0.00%	57.00%	15.00%	90.50%	92.00%	53.50%	32.00%	32.00%	96.00%
	Length	0.06	2.28	1.21	2.64	1.92	1.98	1.27	1.29	2.09
$s = 10$, $p = 800$, $n = 400$	MAB†	1.35	1.02	0.92	0.56	0.48	0.60	0.77	0.79	0.40
	MAB‡	1.41	1.02	0.93	0.54	0.44	0.56	0.80	0.80	0.33
	std	0.01	0.45	0.25	0.53	0.37	0.46	0.24	0.25	0.43
	RMSE	0.91	0.76	0.65	0.53	0.42	0.57	0.55	0.56	0.42
	CP	0.50%	28.50%	8.50%	93.50%	82.00%	60.50%	15.50%	16.00%	91.50%
	Length	0.05	1.77	0.97	2.06	1.45	1.81	0.94	0.97	1.69
$s = 10$, $p = 1000$, $n = 800$	MAB†	1.11	0.91	0.70	0.32	0.40	0.32	0.61	0.64	0.33
	MAB‡	1.11	0.93	0.71	0.28	0.39	0.23	0.59	0.58	0.30
	std	0.02	0.35	0.19	0.41	0.27	0.44	0.20	0.19	0.33
	RMSE	0.76	0.67	0.50	0.37	0.34	0.40	0.44	0.46	0.34
	CP	0.00%	18.00%	11.50%	97.50%	72.00%	88.00%	22.00%	15.00%	89.00%
	Length	0.09	1.38	0.75	1.62	1.05	1.71	0.77	0.73	1.28
$s = 10$, $p = 1000$, $n = 1600$	MAB†	0.87	0.81	0.51	0.20	0.37	0.20	0.45	0.45	0.34
	MAB‡	0.86	0.82	0.51	0.18	0.37	0.14	0.43	0.45	0.34
	std	0.03	0.29	0.15	0.30	0.19	0.34	0.15	0.15	0.24
	RMSE	0.59	0.59	0.37	0.26	0.29	0.28	0.33	0.33	0.29
	CP	0.00%	15.50%	18.00%	99.50%	49.00%	98.50%	24.12%	25.00%	75.50%
	Length	0.10	1.12	0.59	1.19	0.73	1.33	0.58	0.59	0.93
$s = 20$, $p = 400$, $n = 200$	MAB†	1.44	1.15	1.08	0.66	0.51	0.84	0.84	0.88	0.47
	MAB‡	1.48	1.15	1.11	0.62	0.48	0.87	0.82	0.87	0.36
	std	0.01	0.56	0.29	0.96	0.52	0.49	0.32	0.32	0.56
	RMSE	1.00	0.90	0.78	0.84	0.53	0.72	0.63	0.66	0.53
	CP	0.00%	38.00%	11.50%	98.50%	95.00%	47.00%	32.50%	30.50%	96.00%
	Length	0.05	2.21	1.15	3.75	2.04	1.92	1.27	1.25	2.18
$s = 20$, $p = 800$, $n = 400$	MAB†	1.36	1.06	0.94	0.40	0.40	0.59	0.70	0.72	0.34
	MAB‡	1.38	1.05	0.94	0.35	0.36	0.53	0.69	0.74	0.29
	std	0.02	0.41	0.22	0.71	0.38	0.43	0.23	0.24	0.42
	RMSE	0.92	0.77	0.66	0.58	0.39	0.53	0.50	0.52	0.38
	CP	0.00%	15.00%	6.50%	97.50%	92.00%	64.00%	22.00%	23.00%	93.00%
	Length	0.06	1.61	0.87	2.79	1.49	1.68	0.92	0.93	1.65
$s = 20$, $p = 1000$, $n = 800$	MAB†	1.26	0.98	0.80	0.29	0.33	0.36	0.62	0.64	0.28
	MAB‡	1.27	0.98	0.81	0.25	0.29	0.27	0.59	0.60	0.23

Table 2 continued from previous page

	std	0.02	0.33	0.18	0.58	0.28	0.41	0.17	0.18	0.34
	RMSE	0.86	0.71	0.56	0.46	0.31	0.41	0.44	0.45	0.32
	CP	0.00%	10.00%	7.50%	100.00%	85.50%	82.50%	12.00%	13.50%	95.50%
	Length	0.08	1.29	0.72	2.28	1.11	1.61	0.68	0.69	1.34
$s = 20$, $p = 1000$, $n = 1600$	MAB†	1.03	0.89	0.61	0.32	0.26	0.25	0.46	0.48	0.22
	MAB‡	1.00	0.90	0.60	0.29	0.24	0.21	0.46	0.46	0.19
	std	0.02	0.26	0.14	0.46	0.20	0.34	0.14	0.14	0.25
	RMSE	0.70	0.64	0.43	0.40	0.23	0.30	0.34	0.34	0.24
	CP	0.00%	7.50%	6.00%	99.50%	82.50%	92.50%	20.92%	19.50%	93.50%
	Length	0.09	1.03	0.56	1.79	0.77	1.32	0.57	0.56	0.97
$s = 50$, $p = 400$, $n = 200$	MAB†	1.24	1.09	1.00	0.60	0.51	0.71	0.77	0.79	0.47
	MAB‡	1.25	1.09	1.04	0.56	0.45	0.67	0.78	0.81	0.40
	std	0.02	0.55	0.27	1.26	0.52	0.45	0.32	0.32	0.54
	RMSE	0.86	0.86	0.73	1.01	0.53	0.62	0.59	0.60	0.52
	CP	0.50%	43.00%	10.00%	99.00%	94.00%	62.00%	31.00%	29.50%	96.50%
	Length	0.09	2.15	1.08	4.96	2.04	1.77	1.26	1.26	2.11
$s = 50$, $p = 800$, $n = 400$	MAB†	1.14	0.98	0.85	0.42	0.37	0.44	0.58	0.61	0.34
	MAB‡	1.17	0.98	0.87	0.36	0.31	0.36	0.58	0.63	0.27
	std	0.02	0.41	0.21	0.97	0.38	0.38	0.24	0.24	0.41
	RMSE	0.77	0.72	0.59	0.74	0.37	0.42	0.43	0.45	0.38
	CP	0.00%	17.00%	7.00%	99.50%	94.00%	79.50%	37.00%	31.50%	95.50%
	Length	0.10	1.60	0.82	3.82	1.50	1.48	0.93	0.93	1.61
$s = 50$, $p = 1000$, $n = 800$	MAB†	1.06	0.92	0.73	0.40	0.25	0.32	0.49	0.51	0.23
	MAB‡	1.07	0.93	0.73	0.32	0.20	0.30	0.49	0.50	0.18
	std	0.03	0.31	0.16	0.78	0.28	0.30	0.18	0.18	0.32
	RMSE	0.73	0.67	0.52	0.62	0.27	0.31	0.36	0.37	0.28
	CP	0.00%	8.50%	4.00%	100.00%	93.50%	89.50%	23.00%	20.00%	97.50%
	Length	0.11	1.22	0.63	3.05	1.11	1.16	0.69	0.70	1.25
$s = 50$, $p = 1000$, $n = 1600$	MAB†	0.94	0.84	0.60	0.48	0.19	0.26	0.45	0.46	0.18
	MAB‡	0.94	0.88	0.61	0.45	0.16	0.25	0.47	0.48	0.16
	std	0.03	0.25	0.13	0.62	0.20	0.23	0.13	0.13	0.24
	RMSE	0.64	0.60	0.42	0.56	0.20	0.25	0.32	0.33	0.21
	CP	0.00%	7.00%	5.50%	98.00%	94.50%	85.50%	13.50%	12.00%	98.50%
	Length	0.11	0.99	0.51	2.44	0.80	0.91	0.51	0.51	0.94