Supplementary material for "Statistical inference based on accelerated failure time models under model misspecification and small samples"

1 Simulation Results

In this section, we show results for all the settings discussed in Section 4.

Table 1: Test size (%) in simulation 1. The true model is the Weibull regression model while the fitting model is the Weibull regression model.

		Wald			LR		
κ	n	Naive	Robust	Boot	w	w_K	w^*
$q = 0.2$							
0.5	20	6.09	6.35	6.10	5.86	6.15	4.82
0.5	50	5.36	5.43	5.52	5.30	5.39	4.96
0.5	100	5.16	5.26	5.33	5.13	5.23	5.01
0.5	200	5.11	5.13	5.26	5.09	5.12	5.03
0.5	500	5.06	5.08	5.17	5.05	5.07	5.03
1.0	20	6.52	7.30	6.45	5.94	6.67	4.89
1.0	50	5.44	5.80	5.61	5.22	5.57	4.91
1.0	100	5.29	5.49	5.49	5.17	5.39	5.04
1.0	200	5.18	5.30	5.31	5.13	5.24	5.06
1.0	500	5.15	5.18	5.23	5.10	5.15	5.12
2.0	20	6.98	8.05	6.73	6.05	7.18	4.87
2.0	50	5.66	6.12	5.76	5.32	5.79	4.86
2.0	100	5.40	5.67	5.54	5.21	5.49	5.10
2.0	200	5.18	5.34	5.35	5.11	5.26	5.06
2.0	500	5.15	5.18	5.29	5.11	5.14	5.10
$q = 0.4$							
0.5	20	5.11	5.16	5.82	5.73	5.76	4.85
0.5	50	5.00	5.01	5.33	5.28	5.27	4.95
0.5	100	5.17	5.19	5.43	5.32	5.31	5.19
0.5	200	5.03	5.02	5.19	5.08	5.09	5.04
0.5	500	4.96	4.95	5.11	4.98	4.98	5.00
1.0	20	6.07	6.52	6.09	5.95	6.40	4.80
1.0	50	5.36	5.60	5.58	5.32	5.55	4.98
1.0	100	5.30	5.35	5.45	5.29	5.34	5.10
1.0	200	5.10	5.14	5.25	5.08	5.13	5.04
1.0	500	5.12	5.13	5.27	5.11	5.13	5.11
2.0	20	6.95	8.07	6.51	6.11	7.22	4.75
2.0	50	5.69	6.20	5.75	5.37	5.90	4.97
2.0	100	5.37	5.62	5.49	5.24	5.47	5.04
2.0	200	5.19	5.32	5.30	5.11	5.24	5.04
2.0	500	5.09	5.18	5.23	5.06	5.15	5.12

Note: κ is the shape parameter in the true model. q is the parameter of the censoring mechanism. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 2: Test size (%) in simulation 1. The true model is the log-logistic regression model while the fitting model is the Weibull regression model.

		Wald			LR		
κ	n	Naive	Robust	Boot	w	w_K	w^*
$q = 0.2$							
0.5	20	13.67	10.20	7.56	11.91	8.48	4.81
0.5	50	12.96	7.69	6.69	12.19	6.94	5.05
0.5	100	13.03	6.72	6.24	12.61	6.31	5.09
0.5	200	13.06	6.11	5.98	12.80	5.90	5.16
0.5	500	13.07	5.51	5.51	12.97	5.43	5.13
1.0	20	10.14	7.81	6.68	9.27	6.96	4.89
1.0	50	9.05	6.17	5.90	8.74	5.87	5.01
1.0	100	8.84	5.63	5.55	8.68	5.46	5.08
1.0	200	8.63	5.32	5.38	8.55	5.26	5.08
1.0	500	8.49	5.12	5.17	8.47	5.10	5.02
2.0	20	7.93	6.51	6.14	7.74	6.28	4.87
2.0	50	7.17	5.48	5.47	7.11	5.41	4.94
2.0	100	7.16	5.34	5.40	7.13	5.32	5.09
2.0	200	7.00	5.14	5.23	6.98	5.13	5.05
2.0	500	6.85	5.04	5.13	6.85	5.03	5.01
$q = 0.4$							
0.5	20	10.86	8.73	6.78	9.70	7.60	4.82
0.5	50	10.10	6.85	6.24	9.57	6.35	5.04
0.5	100	10.07	6.10	5.86	9.76	5.82	5.05
0.5	200	9.95	5.74	5.69	9.82	5.58	5.16
0.5	500	9.83	5.32	5.40	9.77	5.28	5.11
1.0	20	7.51	6.47	6.14	7.54	6.53	4.96
1.0	50	7.04	5.62	5.67	7.05	5.64	5.05
1.0	100	6.81	5.32	5.34	6.84	5.34	5.04
1.0	200	6.79	5.20	5.32	6.81	5.20	5.06
1.0	500	6.64	5.02	5.12	6.64	5.03	5.01
2.0	20	5.70	5.28	5.74	6.46	6.03	4.89
2.0	50	5.75	5.12	5.36	6.09	5.44	5.02
2.0	100	5.76	5.06	5.25	5.92	5.22	5.02
2.0	200	5.72	5.01	5.17	5.79	5.08	5.03
2.0	500	5.76	5.04	5.14	5.80	5.08	5.06

Note: κ is the shape parameter in the true model. q is the parameter of the censoring mechanism. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 3: Test size (%) in simulation 1. The true model is the log-normal regression model while the fitting model is the Weibull regression model.

		Wald			LR		
κ	n	Naive	Robust	Boot	w	w_K	w^*
$q = 0.2$							
0.5	20	13.52	10.65	7.59	11.68	8.75	4.82
0.5	50	12.59	7.94	6.76	11.76	7.12	5.03
0.5	100	12.38	6.74	6.24	11.95	6.33	5.15
0.5	200	12.42	6.13	5.99	12.19	5.90	5.15
0.5	500	12.48	5.61	5.59	12.40	5.52	5.19
1.0	20	11.72	9.14	7.15	10.40	7.79	4.89
1.0	50	10.73	7.00	6.34	10.22	6.44	5.04
1.0	100	10.37	6.09	5.87	10.08	5.81	5.12
1.0	200	10.36	5.72	5.67	10.24	5.57	5.20
1.0	500	10.25	5.34	5.35	10.19	5.29	5.16
2.0	20	9.36	7.36	6.52	8.77	6.78	4.88
2.0	50	8.49	5.98	5.79	8.24	5.76	5.05
2.0	100	8.18	5.48	5.46	8.05	5.37	5.01
2.0	200	7.97	5.27	5.35	7.91	5.21	5.07
2.0	500	7.98	5.14	5.24	7.96	5.12	5.07
$q = 0.4$							
0.5	20	12.81	10.73	7.33	10.90	8.82	4.81
0.5	50	11.75	7.78	6.55	10.92	6.97	4.97
0.5	100	11.60	6.74	6.23	11.16	6.30	5.19
0.5	200	11.52	6.14	5.91	11.28	5.91	5.21
0.5	500	11.73	5.63	5.61	11.61	5.54	5.21
1.0	20	9.99	8.22	6.67	9.14	7.38	4.93
1.0	50	9.02	6.38	5.92	8.64	6.04	5.01
1.0	100	8.83	5.81	5.69	8.63	5.63	5.08
1.0	200	8.63	5.44	5.45	8.56	5.33	5.08
1.0	500	8.71	5.24	5.32	8.68	5.21	5.12
2.0	20	7.21	6.21	5.98	7.45	6.44	4.91
2.0	50	6.86	5.49	5.50	6.97	5.59	5.00
2.0	100	6.67	5.16	5.25	6.73	5.20	4.93
2.0	200	6.60	5.03	5.13	6.61	5.04	4.94
2.0	500	6.77	5.17	5.26	6.78	5.18	5.14

Note: κ is the shape parameter in the true model. q is the parameter of the censoring mechanism. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 4: Coverage probability (%) and mean of $\hat{\beta}_g$ in simulation 1 when the true model is the Weibull regression model. The parameter q of the censoring mechanism is 0.2.

				Wald			LR		
κ	n	α_g	$\hat{\beta}_g$	Naive	Robust	Boot	w	w_K	w^*
0.5	20	0.0	0.00	93.9	93.7	93.9	94.1	93.9	95.2
0.5	20	0.5	0.50	94.2	94.0	93.9	94.3	94.0	95.1
0.5	20	1.0	1.01	94.1	94.1	94.1	94.1	94.1	95.1
0.5	50	0.0	0.00	94.6	94.6	94.5	94.7	94.6	95.0
0.5	50	0.5	0.51	94.6	94.6	94.6	94.6	94.6	95.0
0.5	50	1.0	1.01	94.9	94.8	94.6	94.7	94.6	95.2
0.5	100	0.0	0.00	94.8	94.7	94.7	94.9	94.8	95.0
0.5	100	0.5	0.50	94.6	94.7	94.4	94.6	94.7	94.9
0.5	100	1.0	1.00	94.7	94.7	94.7	94.6	94.7	94.9
0.5	200	0.0	0.00	94.9	94.9	94.7	94.9	94.9	95.0
0.5	200	0.5	0.50	94.8	94.7	94.7	94.8	94.8	95.0
0.5	200	1.0	1.00	94.7	94.6	94.6	94.7	94.6	94.9
0.5	500	0.0	0.00	94.9	94.9	94.8	95.0	94.9	95.0
0.5	500	0.5	0.50	94.3	94.2	94.2	94.3	94.2	94.3
0.5	500	1.0	1.00	94.4	94.4	94.2	94.5	94.4	94.5
1.0	20	0.0	0.00	93.5	92.7	93.6	94.1	93.3	95.1
1.0	20	0.5	0.50	93.4	92.9	93.5	93.8	93.3	95.1
1.0	20	1.0	1.00	93.6	93.1	93.9	94.0	93.6	95.3
1.0	50	0.0	0.00	94.6	94.2	94.4	94.8	94.4	95.1
1.0	50	0.5	0.51	94.6	94.2	94.3	94.8	94.3	95.2
1.0	50	1.0	1.01	94.6	94.2	94.2	94.6	94.2	95.4
1.0	100	0.0	0.00	94.7	94.5	94.5	94.8	94.6	95.0
1.0	100	0.5	0.50	94.7	94.6	94.7	94.8	94.7	95.1
1.0	100	1.0	1.00	94.6	94.4	94.4	94.7	94.5	95.5
1.0	200	0.0	0.00	94.8	94.7	94.7	94.9	94.8	94.9
1.0	200	0.5	0.50	94.6	94.5	94.4	94.7	94.6	95.0
1.0	200	1.0	1.00	94.9	94.8	94.7	94.9	94.9	95.7
1.0	500	0.0	0.00	94.9	94.8	94.8	94.9	94.9	94.9
1.0	500	0.5	0.50	94.3	94.3	94.3	94.3	94.3	94.7
1.0	500	1.0	1.00	94.5	94.5	94.2	94.5	94.6	95.2
2.0	20	0.0	0.00	93.0	92.0	93.3	94.0	92.8	95.1
2.0	20	0.5	0.50	93.0	92.0	93.4	93.9	92.8	95.5
2.0	20	1.0	1.00	93.1	91.9	93.4	94.0	92.8	95.8
2.0	50	0.0	0.00	94.3	93.9	94.2	94.7	94.2	95.1
2.0	50	0.5	0.50	94.4	93.9	94.3	94.7	94.2	95.7
2.0	50	1.0	1.00	94.3	93.9	94.4	94.5	94.1	96.1
2.0	100	0.0	0.00	94.6	94.3	94.5	94.8	94.5	94.9
2.0	100	0.5	0.50	94.8	94.5	94.6	94.9	94.7	95.9
2.0	100	1.0	1.00	94.5	94.1	94.3	94.7	94.4	96.1
2.0	200	0.0	0.00	94.8	94.7	94.6	94.9	94.7	94.9
2.0	200	0.5	0.50	94.8	94.6	94.7	94.8	94.7	95.8
2.0	200	1.0	1.00	94.9	94.8	94.6	95.1	94.9	96.2
2.0	500	0.0	0.00	94.9	94.8	94.7	94.9	94.9	94.9
2.0	500	0.5	0.50	94.3	94.3	94.2	94.4	94.4	95.4
2.0	500	1.0	1.00	94.6	94.5	94.5	94.5	94.4	95.9

Note: α_g and β_g are group effects in the true model and fitting model, respectively. κ is the shape parameter in the true model. $\hat{\beta}_g$ is the maximum likelihood estimator of β_g. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 5: Coverage probability (%) and mean of $\hat{\beta}_g$ in simulation 1 when the true model is the Weibull regression model. The parameter q of the censoring mechanism is 0.4.

				Wald			LR		
κ	n	α_g	$\hat{\beta}_g$	Naive	Robust	Boot	w	w_K	w^*
0.5	20	0.0	0.00	94.9	94.8	94.2	94.3	94.2	95.1
0.5	20	0.5	0.50	94.9	94.8	94.2	94.2	94.1	95.0
0.5	20	1.0	1.02	95.0	94.9	94.2	94.2	94.3	95.2
0.5	50	0.0	0.00	95.0	95.0	94.7	94.7	94.7	95.0
0.5	50	0.5	0.51	95.0	95.2	94.7	94.6	94.8	95.1
0.5	50	1.0	1.02	95.0	95.1	94.6	94.6	94.7	95.0
0.5	100	0.0	0.00	94.8	94.8	94.6	94.7	94.7	94.8
0.5	100	0.5	0.50	94.6	94.5	94.4	94.5	94.4	94.6
0.5	100	1.0	1.00	94.9	95.0	94.7	94.7	94.8	94.9
0.5	200	0.0	0.00	95.0	95.0	94.8	94.9	94.9	95.0
0.5	200	0.5	0.50	95.1	95.0	94.8	94.9	94.9	94.9
0.5	200	1.0	1.00	95.1	95.1	95.0	94.9	94.9	95.0
0.5	500	0.0	0.00	95.0	95.1	94.9	95.0	95.0	95.0
0.5	500	0.5	0.50	94.4	94.4	94.2	94.3	94.3	94.4
0.5	500	1.0	1.00	94.6	94.6	94.5	94.6	94.6	94.6
1.0	20	0.0	0.00	93.9	93.5	93.9	94.0	93.6	95.2
1.0	20	0.5	0.50	94.2	93.8	94.1	94.1	93.9	95.2
1.0	20	1.0	1.02	94.3	93.8	94.0	94.0	93.7	95.3
1.0	50	0.0	0.00	94.6	94.4	94.4	94.7	94.5	95.0
1.0	50	0.5	0.51	94.8	94.6	94.6	94.7	94.5	95.1
1.0	50	1.0	1.01	95.0	94.7	94.3	94.6	94.5	95.3
1.0	100	0.0	0.00	94.7	94.6	94.5	94.7	94.7	94.9
1.0	100	0.5	0.50	94.7	94.5	94.5	94.6	94.5	94.8
1.0	100	1.0	1.00	95.2	95.1	95.1	94.9	94.9	95.4
1.0	200	0.0	0.00	94.9	94.9	94.8	94.9	94.9	95.0
1.0	200	0.5	0.50	94.9	94.9	94.8	94.9	94.8	95.1
1.0	200	1.0	1.00	95.0	95.1	94.9	95.0	95.0	95.5
1.0	500	0.0	0.00	94.9	94.9	94.7	94.9	94.9	94.9
1.0	500	0.5	0.50	94.8	94.8	94.6	94.8	94.8	95.0
1.0	500	1.0	1.00	94.9	94.9	94.9	94.8	94.8	95.3
2.0	20	0.0	0.00	93.1	91.9	93.5	93.9	92.8	95.2
2.0	20	0.5	0.50	93.1	91.9	93.6	93.9	92.6	95.4
2.0	20	1.0	0.98	93.3	91.9	93.8	93.8	92.5	95.8
2.0	50	0.0	0.00	94.3	93.8	94.2	94.6	94.1	95.0
2.0	50	0.5	0.50	94.4	93.9	94.4	94.6	94.0	95.7
2.0	50	1.0	1.00	94.1	93.3	94.0	94.4	93.6	96.0
2.0	100	0.0	0.00	94.6	94.4	94.5	94.8	94.5	95.0
2.0	100	0.5	0.50	94.8	94.4	94.6	94.9	94.5	95.6
2.0	100	1.0	1.00	94.9	94.6	94.9	95.0	94.7	96.5
2.0	200	0.0	0.00	94.8	94.7	94.7	94.9	94.8	95.0
2.0	200	0.5	0.50	94.7	94.7	94.8	94.7	94.7	95.6
2.0	200	1.0	1.00	94.8	94.7	94.6	94.8	94.7	96.3
2.0	500	0.0	0.00	94.9	94.8	94.8	94.9	94.9	94.9
2.0	500	0.5	0.50	94.4	94.4	94.3	94.5	94.4	95.4
2.0	500	1.0	1.00	94.7	94.6	94.4	94.7	94.5	96.2

Note: α_g and β_g are group effects in the true model and fitting model, respectively. κ is the shape parameter in the true model. $\hat{\beta}_g$ is the maximum likelihood estimator of β_g. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 6: Coverage probability (%) and mean of $\hat{\beta}_g$ in simulation 1 when the true model is the log-logistic regression model. The parameter q of the censoring mechanism is 0.2.

				Wald			LR		
κ	n	α_g	$\hat{\beta}_g$	Naive	Robust	Boot	w	w_K	w^*
0.5	20	0.0	0.00	86.3	89.0	92.4	88.1	91.5	95.2
0.5	20	0.5	0.48	86.6	89.9	92.3	88.2	91.5	94.5
0.5	20	1.0	0.96	88.0	90.1	92.2	89.6	91.5	94.2
0.5	50	0.0	0.00	87.0	92.3	93.3	87.8	93.1	94.9
0.5	50	0.5	0.47	88.0	92.4	93.2	88.8	93.0	94.5
0.5	50	1.0	0.95	88.9	92.1	92.8	89.8	93.0	94.0
0.5	100	0.0	0.00	87.0	93.3	93.8	87.4	93.7	94.9
0.5	100	0.5	0.47	87.7	93.3	93.8	88.4	93.8	94.6
0.5	100	1.0	0.95	88.2	92.4	92.7	88.8	93.2	93.8
0.5	200	0.0	0.00	86.9	93.9	94.0	87.2	94.1	94.8
0.5	200	0.5	0.47	86.9	93.1	93.2	87.2	93.4	93.8
0.5	200	1.0	0.95	86.5	91.7	91.7	86.9	92.3	92.4
0.5	500	0.0	0.00	86.9	94.5	94.5	87.0	94.6	94.9
0.5	500	0.5	0.47	86.2	93.3	93.4	86.4	93.4	93.5
0.5	500	1.0	0.95	82.2	88.8	88.7	82.8	89.2	89.0
1.0	20	0.0	0.00	89.9	92.2	93.3	90.7	93.0	95.1
1.0	20	0.5	0.46	90.1	92.2	93.1	90.9	93.1	94.9
1.0	20	1.0	0.92	90.5	92.3	93.3	91.4	93.2	94.8
1.0	50	0.0	0.00	91.0	93.8	94.1	91.3	94.1	95.0
1.0	50	0.5	0.45	91.1	93.4	93.7	91.3	93.8	94.6
1.0	50	1.0	0.90	91.1	93.2	93.3	91.5	93.7	94.5
1.0	100	0.0	0.00	91.2	94.4	94.4	91.3	94.5	94.9
1.0	100	0.5	0.45	91.0	94.2	94.1	91.2	94.4	94.8
1.0	100	1.0	0.91	90.0	92.6	92.8	90.6	93.1	93.6
1.0	200	0.0	0.00	91.4	94.7	94.6	91.4	94.7	94.9
1.0	200	0.5	0.45	90.4	93.6	93.5	90.6	93.8	94.0
1.0	200	1.0	0.91	88.6	91.5	91.6	89.1	92.0	92.2
1.0	500	0.0	0.00	91.5	94.9	94.8	91.5	94.9	95.0
1.0	500	0.5	0.45	89.7	93.4	93.2	89.9	93.5	93.5
1.0	500	1.0	0.91	83.4	87.5	87.4	83.9	87.9	88.0
2.0	20	0.0	0.00	92.1	93.5	93.9	92.3	93.7	95.1
2.0	20	0.5	0.45	92.1	93.4	93.6	92.2	93.5	94.7
2.0	20	1.0	0.89	92.0	93.1	93.3	92.1	93.4	94.6
2.0	50	0.0	0.00	92.8	94.5	94.5	92.9	94.6	95.1
2.0	50	0.5	0.42	93.0	94.4	94.4	93.0	94.5	95.0
2.0	50	1.0	0.86	92.6	94.2	94.1	92.8	94.3	94.7
2.0	100	0.0	0.00	92.8	94.7	94.6	92.9	94.7	94.9
2.0	100	0.5	0.43	92.6	94.4	94.3	92.6	94.5	94.7
2.0	100	1.0	0.87	91.8	93.6	93.5	91.9	93.8	94.0
2.0	200	0.0	0.00	93.0	94.9	94.8	93.0	94.9	94.9
2.0	200	0.5	0.43	92.5	94.3	94.3	92.6	94.3	94.3
2.0	200	1.0	0.87	91.4	93.1	93.0	91.5	93.2	93.4
2.0	500	0.0	0.00	93.2	95.0	94.9	93.2	95.0	95.0
2.0	500	0.5	0.43	91.9	94.0	93.9	92.0	94.0	94.1
2.0	500	1.0	0.87	88.5	90.8	90.5	88.7	91.0	91.2

Note: α_g and β_g are group effects in the true model and fitting model, respectively. κ is the shape parameter in the true model. $\hat{\beta}_g$ is the maximum likelihood estimator of β_g. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 7: Coverage probability (%) and mean of $\hat{\beta}_g$ in simulation 1 when the true model is the log-logistic regression model. The parameter q of the censoring mechanism is 0.4.

				Wald			LR		
κ	n	α_g	$\hat{\beta}_g$	Naive	Robust	Boot	w	w_K	w^*
0.5	20	0.0	0.00	89.1	91.3	93.2	90.3	92.4	95.2
0.5	20	0.5	0.48	90.2	91.5	93.3	91.1	92.4	94.9
0.5	20	1.0	0.97	91.9	91.9	93.2	92.2	92.6	95.1
0.5	50	0.0	0.00	89.9	93.2	93.8	90.4	93.7	95.0
0.5	50	0.5	0.47	90.9	93.3	93.8	91.6	94.0	95.1
0.5	50	1.0	0.95	91.9	93.1	93.8	92.9	94.0	95.2
0.5	100	0.0	0.00	89.9	93.9	94.1	90.2	94.2	95.0
0.5	100	0.5	0.47	90.8	93.7	93.7	91.3	94.1	94.6
0.5	100	1.0	0.96	91.5	92.8	93.1	92.0	93.5	94.8
0.5	200	0.0	0.00	90.1	94.3	94.3	90.2	94.4	94.8
0.5	200	0.5	0.47	89.9	93.4	93.2	90.4	93.7	94.1
0.5	200	1.0	0.96	90.7	92.2	92.5	91.3	93.0	94.1
0.5	500	0.0	0.00	90.2	94.7	94.6	90.2	94.7	94.9
0.5	500	0.5	0.47	89.3	92.9	92.8	89.6	93.0	93.3
0.5	500	1.0	0.95	88.0	89.9	90.1	88.7	90.6	91.9
1.0	20	0.0	0.00	92.5	93.5	93.9	92.5	93.5	95.0
1.0	20	0.5	0.47	92.6	93.4	93.7	92.3	93.4	94.8
1.0	20	1.0	0.95	93.2	93.6	93.8	93.0	93.6	95.0
1.0	50	0.0	0.00	93.0	94.4	94.3	93.0	94.4	95.0
1.0	50	0.5	0.46	93.2	94.4	94.3	93.3	94.4	95.1
1.0	50	1.0	0.92	93.1	94.0	94.1	93.5	94.4	95.0
1.0	100	0.0	0.00	93.2	94.7	94.7	93.2	94.7	95.0
1.0	100	0.5	0.46	92.9	94.3	94.2	93.1	94.3	94.6
1.0	100	1.0	0.92	92.5	93.5	93.8	93.1	94.0	94.4
1.0	200	0.0	0.00	93.2	94.8	94.7	93.2	94.8	94.9
1.0	200	0.5	0.46	92.6	94.1	93.9	92.8	94.1	94.3
1.0	200	1.0	0.92	91.3	92.3	92.5	91.8	92.7	93.0
1.0	500	0.0	0.00	93.4	95.0	94.9	93.4	95.0	95.0
1.0	500	0.5	0.45	91.2	92.9	92.7	91.3	93.1	93.2
1.0	500	1.0	0.92	87.8	89.1	89.2	88.5	89.7	90.1
2.0	20	0.0	0.01	94.3	94.7	94.3	93.5	94.0	95.1
2.0	20	0.5	0.46	94.2	94.6	94.0	93.4	93.8	94.8
2.0	20	1.0	0.92	94.5	94.8	94.3	93.9	94.1	95.1
2.0	50	0.0	0.00	94.3	94.9	94.6	93.9	94.6	95.0
2.0	50	0.5	0.44	94.4	94.9	94.6	94.0	94.6	95.1
2.0	50	1.0	0.89	94.2	94.7	94.4	94.0	94.6	94.9
2.0	100	0.0	0.00	94.2	94.9	94.8	94.1	94.8	95.0
2.0	100	0.5	0.44	94.0	94.6	94.4	93.9	94.5	94.8
2.0	100	1.0	0.90	93.6	94.2	94.1	93.7	94.3	94.4
2.0	200	0.0	0.00	94.3	95.0	94.8	94.2	94.9	95.0
2.0	200	0.5	0.44	94.0	94.6	94.4	93.9	94.6	94.6
2.0	200	1.0	0.90	93.3	93.8	93.8	93.4	94.0	94.1
2.0	500	0.0	0.00	94.2	95.0	94.9	94.2	94.9	94.9
2.0	500	0.5	0.44	93.2	94.0	93.8	93.2	94.0	94.0
2.0	500	1.0	0.89	90.8	91.6	91.4	91.1	91.9	91.9

Note: α_g and β_g are group effects in the true model and fitting model, respectively. κ is the shape parameter in the true model. $\hat{\beta}_g$ is the maximum likelihood estimator of β_g. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 8: Coverage probability (%) and mean of $\hat{\beta}_g$ in simulation 1 when the true model is the log-normal regression model. The parameter q of the censoring mechanism is 0.2.

				Wald			LR		
κ	n	α_g	$\hat{\beta}_g$	Naive	Robust	Boot	w	w_K	w^*
0.5	20	0.0	0.00	86.5	89.3	92.4	88.3	91.3	95.2
0.5	20	0.5	0.49	86.9	89.4	92.5	88.7	91.2	94.0
0.5	20	1.0	0.98	87.5	89.2	91.9	89.5	91.2	93.5
0.5	50	0.0	0.00	87.4	92.1	93.2	88.2	92.9	95.0
0.5	50	0.5	0.49	87.7	92.2	93.2	88.7	93.0	94.3
0.5	50	1.0	0.99	88.1	91.7	92.9	89.2	92.6	93.2
0.5	100	0.0	0.00	87.6	93.3	93.8	88.1	93.7	94.8
0.5	100	0.5	0.49	88.4	93.0	93.5	88.9	93.4	94.0
0.5	100	1.0	0.99	88.2	92.8	93.3	88.8	93.3	93.2
0.5	200	0.0	0.00	87.6	93.9	94.0	87.8	94.1	94.9
0.5	200	0.5	0.49	87.3	93.5	93.7	87.5	93.7	93.7
0.5	200	1.0	0.99	88.0	93.3	93.3	88.4	93.6	92.9
0.5	500	0.0	0.00	87.5	94.4	94.4	87.6	94.5	94.8
0.5	500	0.5	0.49	87.3	94.3	94.2	87.4	94.4	94.2
0.5	500	1.0	0.99	86.6	92.7	92.8	86.8	93.0	91.7
1.0	20	0.0	0.00	88.3	90.9	92.8	89.6	92.6	95.1
1.0	20	0.5	0.48	88.5	91.0	92.8	89.8	92.1	94.7
1.0	20	1.0	0.95	89.4	91.2	92.7	90.6	92.3	94.4
1.0	50	0.0	0.00	89.3	93.0	93.7	89.8	94.0	95.0
1.0	50	0.5	0.48	89.4	93.0	93.6	90.1	93.5	94.7
1.0	50	1.0	0.95	89.8	92.8	93.3	90.7	93.4	94.3
1.0	100	0.0	0.00	89.6	93.9	94.1	89.9	94.4	94.9
1.0	100	0.5	0.48	89.7	93.6	93.8	90.1	94.0	94.4
1.0	100	1.0	0.95	89.5	92.8	93.0	90.2	93.4	93.9
1.0	200	0.0	0.00	89.6	94.3	94.3	89.8	94.7	94.8
1.0	200	0.5	0.48	89.2	93.6	93.7	89.4	93.9	94.2
1.0	200	1.0	0.95	88.3	91.9	91.9	88.8	92.4	92.7
1.0	500	0.0	0.00	89.7	94.7	94.7	89.8	94.8	94.8
1.0	500	0.5	0.48	88.6	93.5	93.2	88.8	93.6	93.6
1.0	500	1.0	0.95	84.3	89.3	89.3	84.8	89.7	90.1
2.0	20	0.0	0.00	90.6	92.6	93.5	91.2	93.6	95.1
2.0	20	0.5	0.45	91.1	92.9	93.6	91.6	93.4	95.0
2.0	20	1.0	0.91	91.2	92.9	93.6	92.0	93.6	94.9
2.0	50	0.0	0.00	91.5	94.0	94.2	91.8	94.4	94.9
2.0	50	0.5	0.45	91.5	93.8	94.0	91.7	94.2	94.9
2.0	50	1.0	0.90	91.2	93.4	93.7	91.6	93.8	94.6
2.0	100	0.0	0.00	91.8	94.5	94.5	92.0	94.8	95.0
2.0	100	0.5	0.45	91.7	94.2	94.3	91.8	94.4	94.7
2.0	100	1.0	0.90	90.9	93.0	93.1	91.1	93.3	93.6
2.0	200	0.0	0.00	92.0	94.7	94.7	92.1	95.0	94.9
2.0	200	0.5	0.45	91.3	93.9	93.9	91.4	94.0	94.2
2.0	200	1.0	0.90	89.9	92.1	92.2	90.2	92.4	92.7
2.0	500	0.0	0.00	92.0	94.9	94.7	92.0	94.8	94.9
2.0	500	0.5	0.45	90.2	93.2	92.9	90.3	93.3	93.3
2.0	500	1.0	0.90	85.0	88.4	88.4	85.4	88.7	89.0

Note: α_g and β_g are group effects in the true model and fitting model, respectively. κ is the shape parameter in the true model. $\hat{\beta}_g$ is the maximum likelihood estimator of β_g. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 9: Coverage probability (%) and mean of $\hat{\beta}_g$ in simulation 1 when the true model is the log-normal regression model. The parameter q of the censoring mechanism is 0.4.

				Wald			LR		
κ	n	α_g	$\hat{\beta}_g$	Naive	Robust	Boot	w	w_K	w^*
0.5	20	0.0	0.00	87.2	89.3	92.7	89.1	91.2	95.2
0.5	20	0.5	0.49	87.9	89.6	92.9	89.8	91.3	94.4
0.5	20	1.0	0.98	89.8	89.6	93.4	91.4	91.4	94.9
0.5	50	0.0	0.00	88.3	92.2	93.4	89.1	93.0	95.0
0.5	50	0.5	0.49	89.1	91.9	93.3	89.8	92.8	94.4
0.5	50	1.0	0.98	90.5	91.6	93.2	91.6	92.7	94.9
0.5	100	0.0	0.00	88.4	93.3	93.8	88.8	93.7	94.8
0.5	100	0.5	0.49	88.8	92.9	93.3	89.2	93.5	94.3
0.5	100	1.0	0.98	90.7	92.9	93.4	91.7	93.5	94.8
0.5	200	0.0	0.00	88.5	93.9	94.1	88.7	94.1	94.8
0.5	200	0.5	0.49	89.1	93.6	93.8	89.4	93.9	94.2
0.5	200	1.0	0.98	90.5	92.7	92.9	91.1	93.2	94.4
0.5	500	0.0	0.00	88.3	94.4	94.4	88.4	94.5	94.8
0.5	500	0.5	0.49	88.1	93.5	93.6	88.4	93.6	93.9
0.5	500	1.0	0.98	89.1	92.2	92.1	89.7	92.8	93.6
1.0	20	0.0	0.00	90.0	91.8	93.3	90.9	92.6	95.1
1.0	20	0.5	0.47	90.5	91.8	93.5	91.2	92.6	95.0
1.0	20	1.0	0.96	91.4	92.1	93.6	92.3	92.9	95.0
1.0	50	0.0	0.00	91.0	93.6	94.1	91.4	94.0	95.0
1.0	50	0.5	0.47	91.2	93.6	94.0	91.6	94.0	94.9
1.0	50	1.0	0.95	91.7	92.9	93.5	92.6	93.9	95.0
1.0	100	0.0	0.00	91.2	94.2	94.3	91.4	94.4	94.9
1.0	100	0.5	0.47	90.9	93.6	93.7	91.4	93.9	94.4
1.0	100	1.0	0.94	91.0	92.4	92.9	91.6	93.2	94.2
1.0	200	0.0	0.00	91.4	94.6	94.6	91.4	94.7	94.9
1.0	200	0.5	0.47	90.7	93.6	93.6	90.9	93.9	94.2
1.0	200	1.0	0.94	90.0	92.0	92.2	90.8	92.8	93.6
1.0	500	0.0	0.00	91.3	94.8	94.7	91.3	94.8	94.9
1.0	500	0.5	0.47	89.5	92.8	93.0	89.7	93.0	93.4
1.0	500	1.0	0.94	85.7	88.3	88.3	86.5	88.9	89.8
2.0	20	0.0	0.00	92.8	93.8	94.0	92.5	93.6	95.1
2.0	20	0.5	0.45	92.9	93.8	93.9	92.6	93.6	95.1
2.0	20	1.0	0.91	92.9	93.6	94.1	92.9	93.7	95.2
2.0	50	0.0	0.00	93.1	94.5	94.5	93.0	94.4	95.0
2.0	50	0.5	0.45	93.0	94.3	94.4	92.9	94.2	95.0
2.0	50	1.0	0.90	92.3	93.4	93.6	92.7	93.6	94.2
2.0	100	0.0	0.00	93.3	94.8	94.8	93.3	94.8	95.1
2.0	100	0.5	0.44	92.9	94.2	94.1	93.1	94.3	94.5
2.0	100	1.0	0.89	91.3	92.5	93.0	91.9	93.2	93.4
2.0	200	0.0	0.00	93.4	95.0	94.9	93.4	95.0	95.1
2.0	200	0.5	0.44	92.3	93.8	93.8	92.5	93.9	94.0
2.0	200	1.0	0.89	90.1	91.5	91.8	90.7	92.1	92.2
2.0	500	0.0	0.00	93.2	94.8	94.7	93.2	94.8	94.9
2.0	500	0.5	0.44	91.1	93.0	92.7	91.3	93.0	93.1
2.0	500	1.0	0.89	85.4	87.1	87.3	85.9	87.7	88.0

Note: α_g and β_g are group effects in the true model and fitting model, respectively. κ is the shape parameter in the true model. $\hat{\beta}_g$ is the maximum likelihood estimator of β_g. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 10: Relative bias (%) of $\hat{\beta}_g$ in simulation 1.

			$q = 0.2$			$q = 0.4$		
κ	n	α_g	Weibull	Log-logistic	Log-normal	Weibull	Log-logistic	Log-normal
0.5	20	0.5	0.4	−3.8	−1.5	0.6	−4.1	−2.4
0.5	20	1.0	0.8	−4.0	−1.7	1.8	−2.8	−2.0
0.5	50	0.5	2.1	−5.5	−1.4	2.8	−6.1	−2.3
0.5	50	1.0	1.4	−5.1	−1.5	2.4	−4.5	−2.0
0.5	100	0.5	0.3	−5.2	−1.3	0.4	−5.9	−2.4
0.5	100	1.0	0.3	−5.0	−1.5	0.4	−4.4	−2.2
0.5	200	0.5	0.2	−5.3	−1.2	0.5	−5.6	−2.1
0.5	200	1.0	0.1	−5.1	−1.4	0.5	−4.4	−2.0
0.5	500	0.5	0.2	−5.2	−1.3	0.0	−5.7	−2.2
0.5	500	1.0	0.1	−5.1	−1.4	0.0	−4.6	−2.0
1.0	20	0.5	−0.3	−7.8	−4.5	0.5	−6.9	−5.6
1.0	20	1.0	0.1	−8.0	−4.6	2.1	−5.4	−4.3
1.0	50	0.5	1.0	−10.2	−4.9	1.7	−9.0	−6.0
1.0	50	1.0	0.7	−9.6	−4.9	1.4	−7.8	−5.4
1.0	100	0.5	0.2	−10.2	−4.9	0.0	−8.8	−6.8
1.0	100	1.0	0.1	−9.5	−5.1	0.1	−7.7	−6.0
1.0	200	0.5	0.0	−9.9	−4.8	0.3	−8.9	−6.5
1.0	200	1.0	0.0	−9.4	−5.0	0.2	−7.9	−6.0
1.0	500	0.5	0.1	−9.7	−4.9	0.0	−9.1	−6.5
1.0	500	1.0	0.1	−9.4	−5.1	0.1	−8.1	−6.1
2.0	20	0.5	−0.6	−10.9	−9.8	−0.7	−7.7	−9.6
2.0	20	1.0	−0.4	−11.3	−9.3	−2.3	−7.5	−8.7
2.0	50	0.5	0.4	−15.1	−9.8	0.5	−11.6	−10.7
2.0	50	1.0	0.2	−13.8	−9.6	0.3	−10.7	−10.2
2.0	100	0.5	0.0	−14.1	−10.4	−0.2	−11.4	−12.0
2.0	100	1.0	−0.1	−13.2	−10.1	0.0	−10.5	−11.0
2.0	200	0.5	0.0	−13.3	−9.9	0.1	−11.5	−11.4
2.0	200	1.0	0.0	−12.9	−10.0	0.1	−10.4	−10.9
2.0	500	0.5	0.0	−13.2	−9.9	0.0	−11.5	−11.5
2.0	500	1.0	0.0	−12.8	−10.0	0.0	−10.9	−11.1

Note: α_g and β_g are group effects in the true model and fitting model, respectively. κ is the shape parameter in the true model. q is the parameter of the censoring mechanism. Relative bias (%) is $100 \times$ (mean of $\hat{\beta}_g - \alpha_g)/\alpha_g$, where $\hat{\beta}_g$ is the maximum likelihood estimator of β_g.

Table 11: Test size (%) in simulation 2. The important binary covariate x_1 is omitted.

			Wald			LR		
$\exp(\eta)$	n	α_1	Naive	Robust	Boot	w	w_K	w^*
$q = 0.2$								
Weibull	20	−2	4.57	6.34	6.02	4.25	6.01	4.93
	20	−1	7.09	7.87	6.67	6.31	7.12	5.14
	20	1	11.59	9.80	7.40	10.08	8.34	4.96
	20	2	12.80	9.10	7.00	11.65	7.90	4.94
	200	−2	3.35	5.00	5.08	3.34	4.98	4.91
	200	−1	5.07	5.09	5.14	4.99	5.00	4.85
	200	1	8.64	5.30	5.31	8.50	5.16	4.91
	200	2	9.28	5.10	5.08	9.17	5.00	4.81
Log-normal	20	−2	5.11	6.86	6.31	4.69	6.35	4.96
	20	−1	10.37	9.25	7.25	9.00	7.95	5.07
	20	1	14.02	10.45	7.70	12.29	8.70	5.05
	20	2	12.98	8.85	7.03	11.80	7.69	4.95
	200	−2	4.18	5.21	5.30	4.13	5.18	5.08
	200	−1	8.69	5.67	5.62	8.53	5.51	5.18
	200	1	11.75	5.71	5.60	11.60	5.54	5.08
	200	2	9.96	5.29	5.26	9.85	5.20	5.03
$q = 0.4$								
Weibull	20	−2	5.09	6.26	5.94	5.14	6.32	4.94
	20	−1	7.34	7.92	6.55	6.72	7.29	5.04
	20	1	10.81	9.38	7.04	9.51	8.10	4.96
	20	2	13.50	9.47	7.15	12.36	8.25	5.02
	200	−2	4.16	4.94	5.03	4.18	4.95	4.88
	200	−1	5.38	5.17	5.22	5.34	5.10	4.92
	200	1	8.39	5.40	5.37	8.29	5.29	4.98
	200	2	10.26	5.25	5.21	10.18	5.14	4.88
Log-normal	20	−2	5.14	6.33	6.07	5.12	6.30	5.02
	20	−1	9.78	9.08	6.97	8.62	7.98	5.08
	20	1	13.41	10.37	7.43	11.83	8.76	5.03
	20	2	14.11	9.45	7.11	12.95	8.25	4.93
	200	−2	4.39	5.12	5.19	4.38	5.12	5.04
	200	−1	8.02	5.56	5.53	7.91	5.44	5.12
	200	1	11.14	5.69	5.56	10.95	5.51	5.06
	200	2	10.89	5.41	5.35	10.78	5.32	5.07

Note: α_1 is the coefficient of the omitted covariate x_1. q is the parameter of the censoring mechanism. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 12: Coverage probability (%) and mean of $\hat{\beta}_g$ in simulation 2. The important binary covariate x_1 is omitted. The parameter q of the censoring mechanism is 0.2.

					Wald			LR		
exp(η)	n	α_1	α_g	$\hat{\beta}_g$	Naive	Robust	Boot	w	w_K	w^*
Weibull	20	−2	0.0	0.00	95.4	93.7	94.0	95.7	94.0	95.1
	20	−2	0.5	0.51	95.4	93.4	94.0	95.6	93.8	95.4
	20	−2	1.0	1.02	95.0	93.8	94.1	95.2	93.9	96.0
	20	−1	0.0	0.00	92.9	92.1	93.3	93.7	92.9	94.9
	20	−1	0.5	0.50	92.8	92.2	93.3	93.5	92.8	95.1
	20	−1	1.0	1.00	92.7	91.9	93.4	93.5	92.8	95.7
	20	1	0.0	0.00	88.4	90.2	92.6	89.9	91.7	95.0
	20	1	0.5	0.50	89.0	90.8	93.0	90.5	92.1	94.6
	20	1	1.0	0.98	89.4	90.7	92.9	90.8	92.1	94.2
	20	2	0.0	0.00	87.2	90.9	93.0	88.4	92.1	95.1
	20	2	0.5	0.47	87.1	90.8	92.7	88.3	92.1	94.7
	20	2	1.0	0.92	87.0	90.6	92.6	88.5	92.1	93.9
	200	−2	0.0	0.00	96.6	95.0	94.9	96.7	95.0	95.1
	200	−2	0.5	0.51	96.2	94.7	94.8	96.2	94.7	95.5
	200	−2	1.0	1.02	95.7	94.5	94.3	95.6	94.4	96.0
	200	−1	0.0	0.00	94.9	94.9	94.9	95.0	95.0	95.1
	200	−1	0.5	0.50	94.8	95.0	95.0	94.9	95.0	95.8
	200	−1	1.0	1.00	94.4	94.7	94.6	94.6	94.8	96.2
	200	1	0.0	0.00	91.4	94.7	94.7	91.5	94.8	95.1
	200	1	0.5	0.49	91.0	94.3	94.4	91.1	94.4	94.9
	200	1	1.0	0.98	90.3	93.6	93.6	90.7	94.0	94.3
	200	2	0.0	0.00	90.7	94.9	94.9	90.8	95.0	95.2
	200	2	0.5	0.47	89.3	94.4	94.3	89.5	94.4	94.6
	200	2	1.0	0.92	84.6	90.7	90.7	85.3	91.3	91.4
Log-normal	20	−2	0.0	0.00	94.9	93.1	93.7	95.3	93.7	95.0
	20	−2	0.5	0.50	94.6	93.1	93.5	95.0	93.5	95.2
	20	−2	1.0	1.01	94.3	93.0	93.3	94.5	93.3	95.8
	20	−1	0.0	0.00	89.6	90.8	92.7	91.0	92.1	94.9
	20	−1	0.5	0.49	89.7	90.5	92.4	90.9	91.7	94.4
	20	−1	1.0	0.98	89.9	90.3	92.4	91.3	91.8	94.5
	20	1	0.0	0.00	86.0	89.6	92.3	87.7	91.3	95.0
	20	1	0.5	0.49	86.2	89.5	92.4	87.9	91.3	94.3
	20	1	1.0	0.97	86.8	89.6	92.1	88.5	91.1	93.3
	20	2	0.0	0.00	87.0	91.2	93.0	88.2	92.3	95.0
	20	2	0.5	0.47	86.5	90.7	92.8	87.6	92.1	94.8
	20	2	1.0	0.92	86.0	90.4	92.7	87.6	91.8	93.8
	200	−2	0.0	0.00	95.8	94.8	94.7	95.9	94.8	94.9
	200	−2	0.5	0.50	96.1	95.1	95.0	96.1	95.1	95.8
	200	−2	1.0	1.01	95.9	95.0	94.9	96.0	95.0	96.4
	200	−1	0.0	0.00	91.3	94.8	94.4	91.5	94.5	94.8
	200	−1	0.5	0.49	91.6	94.3	94.5	91.7	94.5	94.9
	200	−1	1.0	0.99	91.7	94.0	94.0	92.0	94.3	95.0
	200	1	0.0	0.00	88.2	94.3	94.4	88.4	94.5	94.9
	200	1	0.5	0.49	88.2	93.9	93.9	88.4	94.2	94.4
	200	1	1.0	0.97	87.4	92.7	92.7	87.9	93.2	93.2
	200	2	0.0	0.00	90.0	94.7	94.7	90.1	94.8	95.0
	200	2	0.5	0.47	88.5	93.6	93.8	88.8	93.9	94.1
	200	2	1.0	0.92	83.6	89.7	89.7	84.0	90.4	90.7

Note: α_g and β_g are group effects in the true model and fitting model, respectively. α_1 is the coefficient of the omitted covariate x_1. $\hat{\beta}_g$ is the maximum likelihood estimator of β_g. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 13: Coverage probability (%) and mean of $\hat{\beta}_g$ in simulation 2. The important binary covariate x_1 is omitted. The parameter q of the censoring mechanism is 0.4.

					Wald			LR		
$\exp(\varepsilon)$	n	α_1	α_g	$\hat{\beta}_g$	Naive	Robust	Boot	w	w_K	w^*
Weibull	20	−2	0.0	0.00	94.9	93.7	94.1	94.9	93.7	95.1
	20	−2	0.5	0.51	94.6	93.8	93.9	94.2	93.4	95.0
	20	−2	1.0	1.01	94.2	93.9	93.8	93.6	93.3	95.1
	20	−1	0.0	0.00	92.7	92.1	93.5	93.3	92.7	95.0
	20	−1	0.5	0.50	92.3	91.9	93.4	92.9	92.4	94.7
	20	−1	1.0	1.00	92.6	91.9	93.3	92.8	92.2	95.0
	20	1	0.0	0.00	89.2	90.6	93.0	90.5	91.9	95.0
	20	1	0.5	0.49	90.4	91.3	93.2	91.5	92.4	95.0
	20	1	1.0	0.98	91.3	91.3	93.2	92.1	92.2	95.1
	20	2	0.0	0.00	86.5	90.5	92.9	87.6	91.8	95.0
	20	2	0.5	0.43	87.4	91.0	93.0	88.6	92.2	95.1
	20	2	1.0	0.88	88.7	90.8	93.1	90.3	92.7	94.9
	200	−2	0.0	0.00	95.8	95.1	95.0	95.8	95.1	95.1
	200	−2	0.5	0.51	95.2	94.8	94.6	95.1	94.7	95.1
	200	−2	1.0	0.99	94.5	94.7	94.5	94.5	94.6	95.1
	200	−1	0.0	0.00	94.6	94.8	94.8	94.7	94.9	95.1
	200	−1	0.5	0.50	94.2	94.5	94.7	94.2	94.5	95.4
	200	−1	1.0	0.99	94.0	94.5	94.7	94.4	94.8	95.8
	200	1	0.0	0.00	91.6	94.6	94.6	91.7	94.7	95.0
	200	1	0.5	0.48	91.1	93.8	94.0	91.3	94.1	94.3
	200	1	1.0	0.97	91.5	92.9	93.2	92.3	93.5	94.4
	200	2	0.0	0.00	89.7	94.8	94.8	89.8	94.9	95.1
	200	2	0.5	0.43	86.3	91.7	91.9	86.7	92.1	92.3
	200	2	1.0	0.87	80.0	85.0	85.3	81.3	86.5	87.2
Log-normal	20	−2	0.0	0.00	94.9	93.7	93.9	94.9	93.7	95.0
	20	−2	0.5	0.51	94.1	93.1	93.4	94.1	92.9	94.8
	20	−2	1.0	1.01	93.8	93.4	93.6	93.5	93.2	95.2
	20	−1	0.0	0.00	90.2	90.9	93.0	91.4	92.0	94.9
	20	−1	0.5	0.49	90.4	90.8	93.1	91.5	92.0	94.7
	20	−1	1.0	0.98	90.7	90.9	92.9	91.6	91.7	94.9
	20	1	0.0	0.00	86.6	89.6	92.6	88.2	91.2	95.0
	20	1	0.5	0.48	87.4	89.8	92.7	88.9	91.3	94.6
	20	1	1.0	0.96	89.0	90.1	93.0	90.3	91.4	94.4
	20	2	0.0	0.00	85.9	90.6	92.9	87.0	91.8	95.1
	20	2	0.5	0.43	85.6	89.9	92.6	87.0	91.4	94.6
	20	2	1.0	0.87	86.4	89.5	92.1	88.6	91.5	93.8
	200	−2	0.0	0.00	95.6	94.9	94.8	95.6	94.9	95.0
	200	−2	0.5	0.51	95.4	95.0	95.0	95.5	95.1	95.5
	200	−2	1.0	1.00	94.8	94.8	94.7	94.8	94.9	95.7
	200	−1	0.0	0.00	92.0	94.4	94.5	92.1	94.6	94.9
	200	−1	0.5	0.49	92.0	94.3	94.2	92.3	94.4	94.9
	200	−1	1.0	0.98	92.3	93.7	93.8	92.7	94.1	95.2
	200	1	0.0	0.00	88.9	94.3	94.4	89.0	94.5	94.9
	200	1	0.5	0.48	88.5	93.3	93.5	88.9	93.7	94.3
	200	1	1.0	0.95	88.2	91.4	91.5	89.2	92.2	93.2
	200	2	0.0	0.00	89.1	94.6	94.6	89.2	94.7	94.9
	200	2	0.5	0.43	84.4	91.1	91.3	84.9	91.7	92.0
	200	2	1.0	0.86	74.6	81.8	82.5	76.4	83.8	84.5

Note: α_g and β_g are group effects in the true model and fitting model, respectively. α_1 is the coefficient of the omitted covariate x_1. $\hat{\beta}_g$ is the maximum likelihood estimator of β_g. LR is the likelihood ratio test, w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w by Kent (1982), w^* is the adjusted statistic by $\hat{E}_{g_0}[w]$ (proposal), and Boot is the test based on the percentile bootstrap confidence interval.

Table 14: Relative bias (%) of $\hat{\beta}_g$ in simulation 2.

			$q = 0.2$		$q = 0.4$	
n	α_1	α_g	$\exp(\eta)$: Weibull	$\exp(\eta)$: Log-normal	$\exp(\eta)$: Weibull	$\exp(\eta)$: Log-normal
20	−2	0.5	1.3	0.1	2.2	1.3
20	−2	1.0	1.6	0.6	1.0	1.0
20	−1	0.5	−0.1	−1.8	−0.5	−2.4
20	−1	1.0	−0.4	−1.9	−0.1	−2.1
20	1	0.5	−0.9	−2.7	−2.9	−4.4
20	1	1.0	−2.0	−3.2	−2.5	−4.0
20	2	0.5	−5.3	−6.1	−13.1	−13.8
20	2	1.0	−7.8	−8.0	−11.5	−12.9
200	−2	0.5	1.6	0.9	1.2	1.1
200	−2	1.0	1.6	0.9	−0.5	−0.1
200	−1	0.5	−0.1	−1.1	−0.8	−2.3
200	−1	1.0	−0.3	−1.4	−1.0	−2.3
200	1	0.5	−1.8	−2.7	−3.5	−4.9
200	1	1.0	−2.3	−3.2	−3.2	−4.7
200	2	0.5	−6.7	−6.7	−14.1	−14.7
200	2	1.0	−8.4	−8.5	−12.9	−14.3

Note: α_g and β_g are group effects in the true model and fitting model, respectively. q is the parameter of the censoring mechanism. α_1 is the coefficient of the omitted covariate x_1. Relative bias (%) is $100 \times$ (mean of $\hat{\beta}_g - \alpha_g)/\alpha_g$, where $\hat{\beta}_g$ is the maximum likelihood estimator of β_g.

2 Kaplan-Meier curves for empirical data

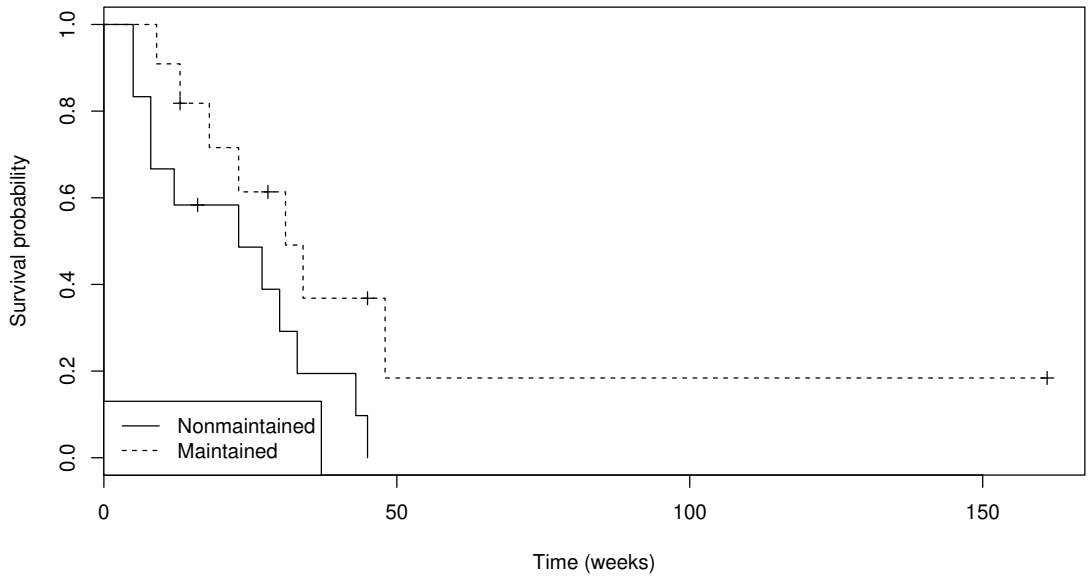

Figure 1: The Kaplan-Meier curves for each treatment group for acute myelogenous leukemia data.

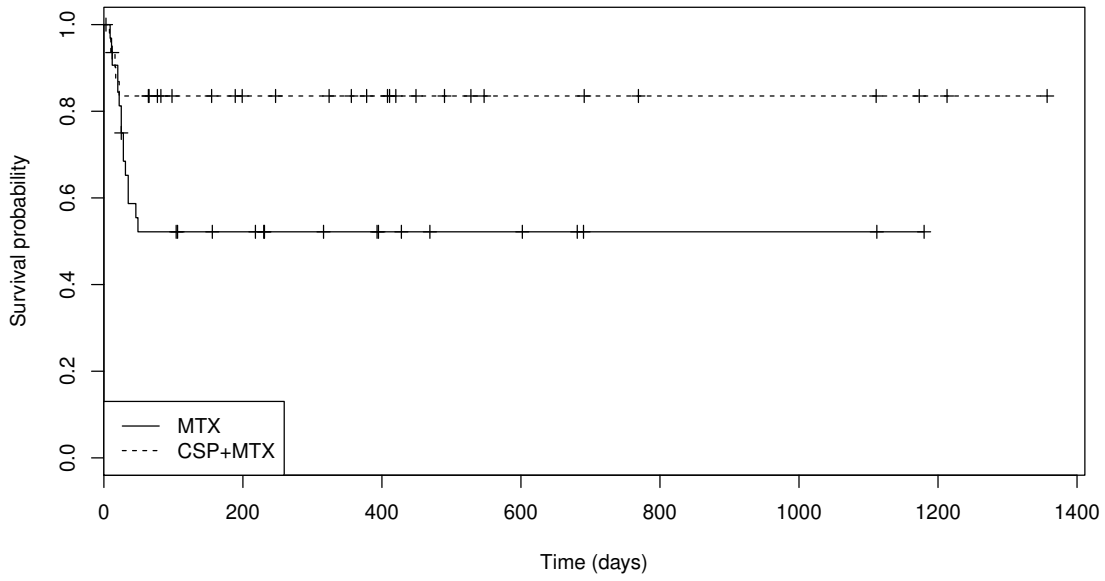

Figure 2: The Kaplan-Meier curves for each treatment group for severe aplastic anemia data.

3 R code for our proposed method

3.1 Input

The `R` file "sample_code.R" contains the function `AFT_Bartlett` with arguments `indat`, `fitdist`, `B`, and `seed`. `indat` is a data frame intended to be analyzed, which has following three variables.

time: event time or censoring time

event: indicator variable, which takes one if the subject experiences event and zero if the subject experiences censoring

group: group indicator variable (e.g., 0 = placebo group; 1 = active group)

`fitdist` specifies a fitting distribution of the error term. In the function `AFT_Bartlett`, we can specify `exponential`, `weibull`, `loglogistic`, and `lognormal`; other distributions are not supported. `B` is the number of resampling arising from the approximation (1) in the main text. `seed` is a random seed.

3.2 Output

The function `AFT_Bartlett` provides the results of naive and robust Wald tests and chi-squared tests based on w, w_K, and $w^* = w/\hat{E}_{g_0}[w]$, where w is the ordinary likelihood ratio statistic, w_K is the adjusted statistic of w proposed by Kent (1982), and w^* is the adjusted statistic using the non-parametric bootstrap method. In the Wald tests, a point estimate of the treatment effect, its standard error, z-value, and p-value are presented. On the other hands, in the likelihood ratio tests, test statistic (w, w_K, or $w^* = w/\hat{E}_{g_0}[w]$), degree of freedom, and p-value are presented.

3.3 Example

We provide an example of an analysis for the acute myelogenous leukemia dataset used in the main article. In the following `R` code, we analyze the dataset by fitting Weibull distribution. Let the number of resampling B be 1,000 and let the random seed be 5678.

```
# observed times
time <- c(9,13,13,18,23,28,31,34,45,48,161,5,5,8,8,12,16,23,27,30,33,43,45)
# event indicators
event <- c(1,1,0,1,1,0,1,1,0,1,0,1,1,1,1,1,0,1,1,1,1,1,1)
# group indicators
group <- c(1,1,1,1,1,1,1,1,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0)

leukemiadat <- data.frame(time=time,event=event,group=group)

AFT_Bartlett(leukemiadat,'weibull',1000,5678)
```

The results are shown as follows. Here, `Loglik` is the maximum value of the log-likelihood function.

```
$`Information`
                                value
number of resampling             1000
fitting distribution          weibull
Loglik               -80.5216452034027

$Wald
        Estimate Std. Error        z          p
Naive  0.9293416  0.3825019 2.429639 0.01511385
Robust 0.9293416  0.4867046 1.909457 0.05620314
```

```
$LR
                w df          p
Ordinary  5.314048  1 0.02115415
Kent      3.282175  1 0.07003608
Bootstrap 3.034376  1 0.08151789
```