Simulation Study

Md Abdul Basit

In this document, we reproduce the results presented in Table 1 (Section 4) of the submitted manuscript. Before executing the following code chunks, please note that–

1. The simulation results include Monte Carlo errors, but we have verified that the fluctuation in results is negligible and does not affect any substantive conclusions drawn from them.

2. Reproducing the simulation results is computationally expensive. The time required to render this document was about 9 hours on an Intel(R) Xeon(R) CPU E5-2650 server with 20 cores and 40 threads.

3. As it takes a long time to run all the codes, some code chunks are not evaluated (`eval=FALSE`), and the results for those code chunks are loaded from `/intermediate_results/` instead. Please set `eval=TRUE` if it is needed to reproduce the results from those chunks.

4. The function `simulation_final_results` doesn't return any result object. Instead, it pastes the results in `/intermediate_results/simulation_results_scenario-I` (for Scenario-I) or `/intermediate_results/simulation_results_scenario-II` (for Scenario-II). The main advantage of storing results in this approach is that even if the simulation is still running, we may observe the results for the completed simulation settings from the csv files stored in /intermediate_results/ instead of waiting for all the settings to finish.

5. In order to reduce the computational time, the number of replications within each simulation setting and the number of bootstrap samples used to calculate the confidence intervals could be reduced to 500 by setting `R = 500` and `B = 500` in `simulation_final_results()`, respectively.

6. If this Rmd file fails to render, delete the `./cache/` folder and then try rendering it again.

Loading required packages and functions

First we load the necessary packages and functions to conduct the simulation study.

```r
library(tidyverse)
library(parallel)
library(kableExtra)
source("code/simulation_functions.R")
```

Obtaining true (population level) partially identified intervals

Next, we obtain the true partially identified intervals for each simulation setting using large scale numerical approximations. The computation of these true intervals should take up about 40 minutes to complete. The output is saved in `/intermediate_results/population_level_partial_intervals.Rdata`.

```r
## The simulation settings
contrast <- rep(rep(list(c(1,-1,0), c(1,0,-1), c(0,1,-1)), each = 6), 2)
lambda <- rep(c(0, 0.1, 0.2, 0.5, 1, 2), 6)
overlap <- rep(c(TRUE, FALSE), each = 18)
```

```r
partial_intervals <- round(
  matrix(
    unlist(
      Map(f = function(x, y, z) {
        pop_level_partial_intervals(seed = 1660, contrast = x, overlap = z,
                                    lambda = y,size = 1e6)
      }, x = contrast, y = lambda, z = overlap)
    ), ncol = 2, byrow = T
  ), 3
)

colnames(partial_intervals) <- c("true_lower","true_upper")

save(partial_intervals,
     file = "intermediate_results/population_level_partial_intervals.Rdata")
```

Simulation results

Next, we set up a total of 36 simulation settings (18 under each scenario) and then reproduce the simulation results for these settings under each scenario separately. The intermediate results are stored in `/intermediate_results/`.

Simulation settings

```r
contrast <- rep(rep(list(c(1,-1,0), c(1,0,-1), c(0,1,-1)), each = 6), 1)
lambda <- rep(c(0, 0.1, 0.2, 0.5, 1, 2), 3)

## Loading the true (population-level) partially identified intervals
load("intermediate_results/population_level_partial_intervals.Rdata")
```

Simulation of Scenario-I (adequate overlap)

```r
simulation_final_results(seed = 1660, overlap = TRUE, contrast = contrast,
                         lambda = lambda, size = 750, B = 1000, R = 1000,
                         parallel = TRUE, alpha = 0.1,
                         partial_intervals = partial_intervals)
```

Simulation of Scenario-II (lack of overlap)

```r
simulation_final_results(seed = 1660, overlap = FALSE, contrast = contrast,
                         lambda = lambda, size = 750, B = 1000, R = 1000,
                         parallel = TRUE, alpha = 0.1,
                         partial_intervals = partial_intervals)
```

Table 1 (Section 4)

```r
scenario_I <- read_csv("intermediate_results/simulation_results_scenario-I.csv")
scenario_II <- read_csv("intermediate_results/simulation_results_scenario-II.csv")

combined_results <- rbind(scenario_I, scenario_II)
```

```r
combined_results %>%
  mutate(
    Overlap = case_when(
      Overlap == 1 ~ "I",
      Overlap == 0 ~ "II"
    ),
    ATE = rep(rep(c("$\\tau_{1, 2}$", "$\\tau_{1, 3}$", "$\\tau_{2, 3}$"), each = 6), 2),
    bias_lower = paste0("$", sprintf("%.2f", bias_lower), "$"),
    bias_upper = paste0("$", sprintf("%.2f", bias_upper), "$"),
    partial_interval = paste0("$", "(", sprintf("%.3f", combined_results$true_lower), ", ",
                              sprintf("%.3f", combined_results$true_upper), ")", "$"),
    point_interval =  paste0("$", "(", sprintf("%.3f", combined_results$point_lower), ", ",
                             sprintf("%.3f", combined_results$point_upper), ")", "$"),
    conf_interval = paste0("$", "(", sprintf("%.3f", combined_results$conf_lower), ", ",
                           sprintf("%.3f", combined_results$conf_upper), ")", "$"),
  ) %>%
  select(Overlap, ATE, lambda, Lambda, bias_lower, bias_upper, non_coverage,
         partial_interval, point_interval, conf_interval) %>%
  kbl(booktabs = T, digits = 3, align = "ccrrrrrrrr", escape = F, format = "latex",
      col.names = linebreak(c("Scenario", "ATE", "$\\lambda$", "$\\Lambda$", "Lower", "Upper",
                    "Non-\n coverage", "Partially identified \n interval",
                    "Point estimate \n interval", "Confidence \n interval"), align = "c"),
      caption = "Simulation results presented in Table 1 (Section 4).") %>%
  add_header_above(c(" " = 4, "% Bias" = 2, " " = 4)) %>%
  kable_styling(latex_options = c("scale_down", "hold_position")) %>%
  collapse_rows(columns = 1:2, valign = "middle", latex_hline = "custom",
  custom_latex_hline = 1:2)
```

Table 1: Simulation results presented in Table 1 (Section 4).

Scenario	ATE	λ	Λ	% Bias Lower	% Bias Upper	Non-coverage	Partially identified interval	Point estimate interval	Confidence interval
		0.0	1.00	2.79	2.79	0.103	$(-0.050, -0.050)$	$(-0.048, -0.048)$	$(-0.116, 0.017)$
		0.1	1.11	3.43	3.01	0.104	$(-0.108, 0.008)$	$(-0.107, 0.010)$	$(-0.171, 0.076)$
		0.2	1.22	3.40	0.65	0.101	$(-0.165, 0.067)$	$(-0.164, 0.067)$	$(-0.228, 0.133)$
	$\tau_{1,2}$	0.5	1.65	2.97	0.59	0.115	$(-0.328, 0.237)$	$(-0.328, 0.238)$	$(-0.387, 0.302)$
		1.0	2.72	6.74	-3.17	0.124	$(-0.559, 0.488)$	$(-0.558, 0.489)$	$(-0.606, 0.543)$
		2.0	7.39	6.67	-5.32	0.110	$(-0.832, 0.801)$	$(-0.832, 0.802)$	$(-0.856, 0.830)$
		0.0	1.00	1.49	1.49	0.105	$(-0.034, -0.034)$	$(-0.035, -0.035)$	$(-0.101, 0.033)$
		0.1	1.11	2.15	1.12	0.110	$(-0.093, 0.025)$	$(-0.093, 0.024)$	$(-0.161, 0.092)$
		0.2	1.22	4.10	2.50	0.111	$(-0.152, 0.083)$	$(-0.152, 0.084)$	$(-0.217, 0.150)$
I	$\tau_{1,3}$	0.5	1.65	5.27	0.69	0.111	$(-0.320, 0.254)$	$(-0.319, 0.255)$	$(-0.380, 0.318)$
		1.0	2.72	7.27	-2.13	0.126	$(-0.558, 0.502)$	$(-0.556, 0.502)$	$(-0.607, 0.556)$
		2.0	7.39	9.38	-5.69	0.110	$(-0.837, 0.808)$	$(-0.836, 0.808)$	$(-0.861, 0.837)$
		0.0	1.00	1.22	1.22	0.092	$(0.015, 0.015)$	$(0.016, 0.016)$	$(-0.052, 0.085)$
		0.1	1.11	0.26	-0.80	0.101	$(-0.045, 0.076)$	$(-0.044, 0.077)$	$(-0.112, 0.144)$
		0.2	1.22	2.31	0.46	0.101	$(-0.106, 0.135)$	$(-0.105, 0.136)$	$(-0.173, 0.202)$
	$\tau_{2,3}$	0.5	1.65	3.40	-0.91	0.093	$(-0.281, 0.306)$	$(-0.280, 0.306)$	$(-0.345, 0.367)$
		1.0	2.72	4.48	-2.57	0.102	$(-0.531, 0.546)$	$(-0.531, 0.547)$	$(-0.583, 0.596)$
		2.0	7.39	9.62	-7.57	0.102	$(-0.827, 0.829)$	$(-0.827, 0.829)$	$(-0.851, 0.853)$
		0.0	1.00	-15.30	-15.30	0.189	$(-0.044, -0.044)$	$(-0.084, -0.084)$	$(-0.261, 0.101)$
		0.1	1.11	-11.34	-19.21	0.205	$(-0.106, 0.019)$	$(-0.140, -0.026)$	$(-0.307, 0.159)$
		0.2	1.22	-7.02	-23.34	0.225	$(-0.167, 0.083)$	$(-0.194, 0.033)$	$(-0.350, 0.219)$
	$\tau_{1,2}$	0.5	1.65	3.27	-35.38	0.274	$(-0.334, 0.273)$	$(-0.341, 0.207)$	$(-0.474, 0.389)$
		1.0	2.72	16.90	-51.30	0.376	$(-0.559, 0.548)$	$(-0.551, 0.475)$	$(-0.647, 0.620)$
		2.0	7.39	30.93	-65.77	0.500	$(-0.824, 0.852)$	$(-0.812, 0.813)$	$(-0.860, 0.869)$
		0.0	1.00	-15.57	-15.57	0.203	$(-0.030, -0.030)$	$(-0.072, -0.072)$	$(-0.243, 0.101)$
		0.1	1.11	-12.37	-19.35	0.217	$(-0.089, 0.031)$	$(-0.126, -0.017)$	$(-0.285, 0.162)$
		0.2	1.22	-8.88	-22.72	0.230	$(-0.147, 0.092)$	$(-0.179, 0.037)$	$(-0.327, 0.218)$
II	$\tau_{1,3}$	0.5	1.65	-0.37	-33.93	0.274	$(-0.307, 0.276)$	$(-0.324, 0.211)$	$(-0.448, 0.386)$
		1.0	2.72	9.74	-48.42	0.355	$(-0.525, 0.544)$	$(-0.525, 0.475)$	$(-0.617, 0.617)$
		2.0	7.39	21.21	-61.01	0.472	$(-0.798, 0.846)$	$(-0.790, 0.809)$	$(-0.840, 0.865)$
		0.0	1.00	-0.26	-0.26	0.129	$(0.014, 0.014)$	$(0.014, 0.014)$	$(-0.110, 0.135)$
		0.1	1.11	2.08	-2.59	0.141	$(-0.037, 0.065)$	$(-0.035, 0.063)$	$(-0.157, 0.183)$
		0.2	1.22	3.60	-5.17	0.154	$(-0.087, 0.116)$	$(-0.084, 0.111)$	$(-0.203, 0.229)$
	$\tau_{2,3}$	0.5	1.65	8.88	-11.67	0.175	$(-0.234, 0.264)$	$(-0.227, 0.254)$	$(-0.334, 0.363)$
		1.0	2.72	17.29	-21.99	0.253	$(-0.455, 0.485)$	$(-0.441, 0.468)$	$(-0.532, 0.557)$
		2.0	7.39	27.26	-34.41	0.335	$(-0.760, 0.781)$	$(-0.748, 0.767)$	$(-0.799, 0.814)$