Simulation results

Xiaodan Lyu, Emily Berg, Heike Hofmann

4/6/2020

```
rm(list = ls(all = T))
```

Table 1

```
load("intermediate_results/Cor0.9_D60_nsim1000_seed2020_Result.RData")
mcmse.eb <- tapply(colMeans((pred.eb.store-YbarNis.store)^2), nis, mean)
## MC MSE of the proposed EB predictor
sprintf("%.2f", mcmse.eb*10^5)
```

```
## [1] "24.10" "15.92" "9.27"
```

```
dmse.alt <- function(pred.store){
  mcmse <- apply((pred.store-YbarNis.store)^2, 1, tapply, nis, mean)
  dmse <- mcmse-as.vector(mcmse.eb)
  diff <- rowMeans(dmse)*1e+5
  err <- 1.96*sqrt(apply(dmse, 1, var)/1000)*1e+5
  paste0(sprintf("%.2f", diff), " (", sprintf("%.2f", err), ")")
}
dmse.result <- do.call(
  "cbind",
  lapply(list(eb0 = pred.eb0.store,
              pi = pred.pi.store,
              zi = pred.zi.store,
              si = pred.si.store),
         dmse.alt))
dmse.result
```

```
##      eb0           pi            zi            si
## [1,] "1.16 (0.99)" "1.18 (0.99)" "4.27 (0.96)" "71.56 (7.32)"
## [2,] "0.89 (0.72)" "0.89 (0.72)" "3.57 (0.70)" "78.03 (8.17)"
## [3,] "0.63 (0.35)" "0.63 (0.35)" "2.75 (0.36)" "71.28 (6.54)"
```

Table 2

```
dmse.store <- c()
rhos <- c(-0.9, -0.6, -0.3, 0, 0.3, 0.6, 0.9)
```

```
for (rho in rhos){
  load(sprintf("intermediate_results/Cor%s_D60_nsim1000_seed2020_Result.RData", rho))
  mcmse.eb <- apply((pred.eb.store-YbarNis.store)^2, 1, tapply, nis, mean)
  mcmse.eb0 <- apply((pred.eb0.store-YbarNis.store)^2, 1, tapply, nis, mean)
  dmse <- mcmse.eb0-mcmse.eb
  diff <- rowMeans(dmse)*1e+5
  err <- 1.96*sqrt(apply(dmse, 1, var)/1000)*1e+5
  # rej <- ifelse(diff>err, TRUE, FALSE)
  dmse.store <- cbind(dmse.store, paste0(sprintf("%.2f", diff), " (", sprintf("%.2f", err), ")"))
}
colnames(dmse.store) <- rhos
dmse.store
```

```
##      -0.9          -0.6          -0.3          0             0.3
## [1,] "2.42 (0.25)" "0.97 (0.18)" "0.16 (0.12)"  "-0.13 (0.08)" "-0.04 (0.09)"
## [2,] "2.48 (0.23)" "0.78 (0.16)" "0.01 (0.13)"  "-0.13 (0.08)" "-0.04 (0.07)"
## [3,] "1.54 (0.13)" "0.39 (0.10)" "-0.02 (0.06)" "-0.08 (0.04)" "0.02 (0.04)"
##      0.6           0.9
## [1,] "0.44 (0.13)" "1.16 (0.18)"
## [2,] "0.27 (0.12)" "0.89 (0.16)"
## [3,] "0.22 (0.06)" "0.63 (0.09)"
```

Table 3

```
seeds <- 2015:2019
m12boot <- m1biasboot <- m2boot <- mseboot <- c()
ybar <- predeb <- onestep <- c()
for (seed in seeds){
  load(sprintf("intermediate_results/bootstrap/Cor0.9_D60_nsim200_seed%.0f_Result.RData", seed))
  m12boot <- rbind(m12boot, M12boot.store)
  m1biasboot <- rbind(m1biasboot, M1biasboot.store)
  m2boot <- rbind(m2boot, M2boot.store)
  mseboot <- rbind(mseboot, MSEboot.store)
  ybar <- rbind(ybar, YbarNis.store)
  predeb <- rbind(predeb, pred.eb.store)
  onestep <- rbind(onestep, mse.eb.store)
}
mcmse <- colMeans((predeb - ybar)^2)
evalmse <- function(estmse){
  bmse <- apply(estmse, 1, function(x) tapply(x-mcmse, nis, mean))
  cp <- apply(abs(predeb - ybar) <= 1.96*sqrt(estmse), 1, tapply, nis, mean)
  bmse.val <- rowMeans(bmse)*1e+5
  bmse.err <- 1.96*sqrt(apply(bmse, 1, var)/1000)*1e+5
  cp.val <- rowMeans(cp)*100
  cp.err <- 1.96*sqrt(apply(cp, 1, var)/1000)*100
  data.frame(
    bmse = paste0(sprintf("%.2f", bmse.val), " (", sprintf("%.2f", bmse.err), ")"),
    cp = paste0(sprintf("%.2f", cp.val), " (", sprintf("%.2f", cp.err), ")"))
}
output <- do.call("cbind", lapply(list(
  onestep = onestep, boot = mseboot, semiboot = onestep + m2boot),
```

```
  evalmse))
output
```

```
##   onestep.bmse   onestep.cp    boot.bmse      boot.cp semiboot.bmse
## 1 -0.68 (0.51) 94.94 (0.33)  0.47 (0.48) 94.55 (0.33)   0.46 (0.53)
## 2 -0.97 (0.32) 94.88 (0.32)  0.00 (0.30) 94.07 (0.33)  -0.13 (0.33)
## 3 -0.68 (0.19) 94.47 (0.34) -0.16 (0.16) 93.78 (0.32)  -0.12 (0.19)
##    semiboot.cp
## 1 95.36 (0.31)
## 2 95.47 (0.30)
## 3 95.37 (0.31)
```

Table S2

```
bootres <- do.call("cbind", lapply(
  list(m2boot/mseboot, m1biasboot/mseboot, m12boot/mseboot), colMeans))
bootres <- apply(bootres, 2, tapply, nis, mean)
round(bootres*100, 2)
```

```
##    [,1]  [,2]  [,3]
## 5  4.97 -2.14 -0.14
## 10 5.71 -2.99 -0.21
## 20 6.44 -3.08 -0.26
```

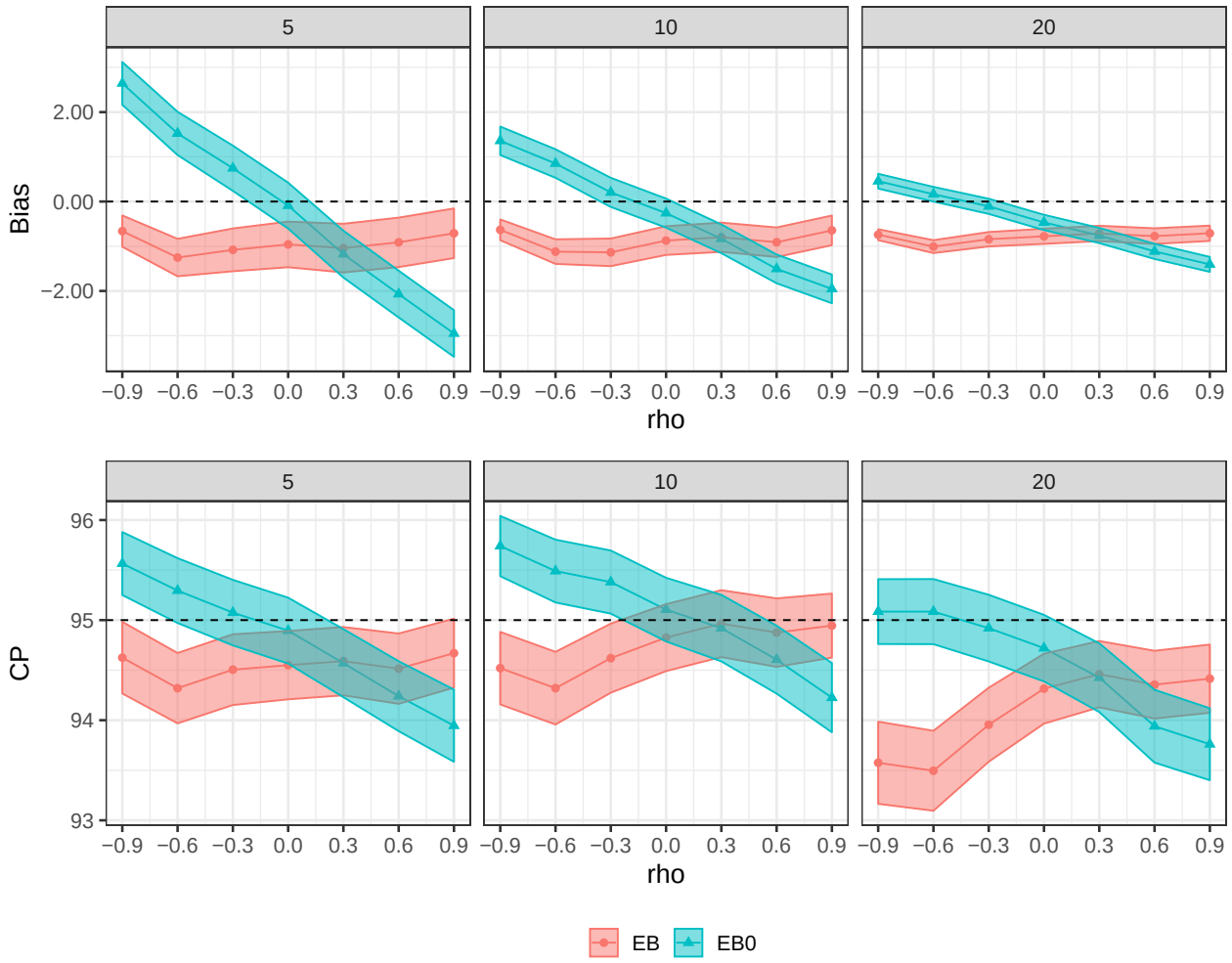
Figure S2

```
library(dplyr)
library(ggplot2)
library(patchwork)
ebeb0.store <- c()
for (rho in rhos){
  load(sprintf("intermediate_results/Cor%s_D60_nsim1000_seed2020_Result.RData", rho))
  mcmse.eb <- tapply(colMeans((pred.eb.store-YbarNis.store)^2), nis, mean)
  mcmse.eb0 <- tapply(colMeans((pred.eb0.store-YbarNis.store)^2), nis, mean)
  bmse <- apply(mse.eb.store, 1, tapply, nis, mean) - as.vector(mcmse.eb)
  bmse0 <- apply(mse.eb0.store, 1, tapply, nis, mean) - as.vector(mcmse.eb0)
  cp <- apply(abs(pred.eb.store - YbarNis.store) <= 1.96*sqrt(mse.eb.store), 1, tapply, nis, mean)
  cp0 <- apply(abs(pred.eb0.store - YbarNis.store) <= 1.96*sqrt(mse.eb0.store), 1, tapply, nis, mean)
  res <- cbind(nis = unique(nis), rho = rho,
               val_bmse = rowMeans(bmse), val_bmse0 = rowMeans(bmse0),
               err_bmse = 1.96*sqrt(apply(bmse, 1, var)/1000),
               err_bmse0 = 1.96*sqrt(apply(bmse0, 1, var)/1000),
               val_cp = rowMeans(cp), val_cp0 = rowMeans(cp0),
               err_cp = 1.96*sqrt(apply(cp, 1, var)/1000),
               err_cp0 = 1.96*sqrt(apply(cp0, 1, var)/1000))
  ebeb0.store <- rbind(ebeb0.store, res)
}
tb <- ebeb0.store %>% as_tibble() %>%
  mutate(nis = forcats::fct_relevel(factor(nis), "5")) %>%
```

```
  mutate_at(vars(val_bmse:err_bmse0), vars(.*1e+5)) %>%
  mutate_at(vars(val_cp:err_cp0), vars(.*100)) %>%
  tidyr::pivot_longer(val_bmse:err_cp0, names_to = c(".value", "metric"), names_pattern = "(.*)_(.*)")
```

```
## Warning: Using quosures is deprecated
## Please use a one-sided formula, a function, or a function name
## This warning is displayed once per session.
```

```
g1 <- tb %>% filter(grepl("bmse", metric)) %>%
  rename(Bias = val) %>%
  ggplot(aes(x = rho, y = Bias, color = metric)) +
  geom_point(aes(shape = metric), size = rel(2)) + geom_line() +
  geom_ribbon(aes(ymin = Bias - err, ymax = Bias + err, fill = metric), alpha = 0.5) +
  geom_hline(yintercept = 0, linetype = "dashed") +
  scale_y_continuous(labels = function(x) sprintf("%.2f", x)) +
  scale_x_continuous(breaks = rhos) +
  guides(color = FALSE, shape = FALSE, fill = FALSE) +
  theme_bw(base_size = 15) +
  facet_grid(~nis)
g2 <- tb %>% filter(grepl("cp", metric)) %>%
  rename(CP = val) %>%
  mutate(metric = recode_factor(metric, "cp" = "EB", "cp0" = "EB0")) %>%
  ggplot(aes(x = rho, y = CP, color = metric)) +
  geom_point(aes(shape = metric), size = rel(2)) + geom_line() +
  geom_ribbon(aes(ymin = CP - err, ymax = CP + err, fill = metric), alpha = 0.5) +
  geom_hline(yintercept = 95, linetype = "dashed") +
  scale_x_continuous(breaks = rhos) +
  theme_bw(base_size = 15) + labs(color = "", shape = "", fill = "") +
  facet_grid(~nis) + theme(legend.position = "bottom")
g1 / g2
```


5
10
20
Bias
2.00
0.00
−2.00
−0.9 −0.6 −0.3 0.0 0.3 0.6 0.9
rho
CP
96
95
94
93
EB
EB0