Meta-Staggered DiD

Till Tietz

2022-10-16

Simulating Bias

This document outlines sources of bias in estimating staggered DiD designs, with unbalanced panels.

We'll set up simulations with varying ammounts of treatment effect heterogeneity. Our simulated data will have 600 units clustered into 3 groups. Two of the groups will have 5 identical treatment onsets. The third group will have 3 treatment onsets.

```
# homogenous
homogenous <- list(
  n = 400,
  n_shifted = 200,
  n_years = 9,
  n_years_shifted = 4,
  n_groups = 5,
  n_groups_shifted = 3,
  tes = rep(0.25, 5),
  tes_shifted = rep(0.25, 3),
  tes_slope = rep(0, 5),
  tes_slope_sifted = rep(0, 3),
  unit_fes = 0.1,
  unit_fes_shifted = 0.1,
  year_fes = rep(0.2, 9),
  year_fes_shifted = rep(0.2, 4),
  states = 1:3
)

heterogenous_years <- list(
  n = 400,
  n_shifted = 200,
  n_years = 9,
  n_years_shifted = 4,
  n_groups = 5,
  n_groups_shifted = 3,
  tes = rep(0.25, 5),
  tes_shifted = rep(0.25, 3),
  tes_slope = rep(0, 5),
  tes_slope_sifted = rep(0, 3),
  unit_fes = 0.1,
  unit_fes_shifted = 0.1,
  year_fes = (1:9 * 0.1),
  year_fes_shifted = c(0.2, 0.3, 0.4, 0.5),
```

```
  states = 1:3
)

heterogenous_units <- list(
  n = 400,
  n_shifted = 200,
  n_years = 9,
  n_years_shifted = 4,
  n_groups = 5,
  n_groups_shifted = 3,
  tes = rep(0.25, 5),
  tes_shifted = rep(0.25, 3),
  tes_slope = rep(0, 5),
  tes_slope_shifted = rep(0, 3),
  unit_fes = rnorm(400),
  unit_fes_shifted = rnorm(200),
  year_fes = (1:9 * 0.1),
  year_fes_shifted = c(0.2, 0.3, 0.4, 0.5),
  states = 1:3
)

heterogenous_full <- list(
  n = 400,
  n_shifted = 200,
  n_years = 9,
  n_years_shifted = 4,
  n_groups = 5,
  n_groups_shifted = 3,
  tes = (1:5 * 0.1),
  tes_shifted = c(0.3, 0.4, 0.5),
  tes_slope = (1:5 * 0.1),
  tes_slope_shifted = c(0.3, 0.4, 0.5),
  unit_fes = rnorm(400),
  unit_fes_shifted = rnorm(200),
  year_fes = (1:9 * 0.1),
  year_fes_shifted = c(0.2, 0.3, 0.4, 0.5),
  states = 1:3
)

sim_params <- list(homogenous,heterogenous_years,heterogenous_units,heterogenous_full)
```

```
lapply(sim_params, function(p)
  run_simulation(params = p, n = 25))%>%
  unlist()%>%
  as.data.frame()%>%
  magrittr::set_colnames("MSE")%>%
  magrittr::set_rownames(c("homogenous",
                           "heterogeneity years",
                           "heterogeneity years + units",
                           "heterogeneity tau"))%>%
  knitr::kable()
```

	MSE
homogenous	0.0035749
heterogeneity years	0.0038439
heterogeneity years + units	0.0030254
heterogeneity tau	0.0177553

Pooling units with different treatment timings biases results when treatment effects vary substantially by treatment timing.

Estimation via Meta-Analysis

We simulate heterogenous data

```
het_minor <- list(
  n = 400,
  n_shifted = 200,
  n_years = 7,
  n_years_shifted = 4,
  n_groups = 5,
  n_groups_shifted = 3,
  tes = (1:5 * 0.1),
  tes_shifted = c(0.3, 0.4, 0.5),
  tes_slope = (1:5 * 0.1),
  tes_slope_sifted = c(0.3, 0.4, 0.5),
  unit_fes = rnorm(400),
  unit_fes_shifted = rnorm(200),
  year_fes = (1:7 * 0.1),
  year_fes_shifted = c(0.2, 0.3, 0.4, 0.5),
  states = 1:3
)


het_major <- list(
  n = 400,
  n_shifted = 200,
  n_years = 7,
  n_years_shifted = 4,
  n_groups = 5,
  n_groups_shifted = 3,
  tes = (1:5 * 0.1),
  tes_shifted = c(1.3, 2.4, 3.5),
  tes_slope = (1:5 * 0.1),
  tes_slope_sifted = c(0.9, 1.2, 2.5),
  unit_fes = rnorm(400),
  unit_fes_shifted = rnorm(200),
  year_fes = (1:7 * 0.1),
  year_fes_shifted = c(0.2, 0.3, 0.4, 0.5),
  states = 1:3
)
```

Analysis:

```
estimates <- lapply(list(het_minor, het_major), function(d) {
  estimate_pooled <- fixest::feols(y ~ treat | unit + year, data = d)

  est <- lapply(list(c(1, 2), c(3)), function(i) {
    fixest::feols(y ~ treat |
                    unit + year, data = dplyr::filter(d, state %in% i)) %>%
      broom::tidy() %>%
      dplyr::filter(term == "treat")
  }) %>%
    dplyr::bind_rows() %>%
    dplyr::select(estimate, std.error) %>%
    magrittr::set_colnames(c("ATT", "SE")) %>%
    dplyr::mutate(model = as.character(1:2))


  estimate_meta <- meta::metagen(
    TE = est$ATT,
    seTE = est$SE,
    studlab = est$model,
    data = est,
    fixed = FALSE,
    random = TRUE,
    method.tau = "REML",
    hakn = TRUE
  )

  return(list(estimate_pooled, estimate_meta))

})
```

heterogeneity	effect	ATT	SE	ci.lower	ci.upper
minor	TRUE	0.4833883	NA	NA	NA
minor	pooled	0.7355339	0.0675431	0.6031494	0.8679184
minor	meta	0.6500832	0.0626678	0.5272566	0.7729097
major	TRUE	1.0416714	NA	NA	NA
major	pooled	1.4055415	0.0945217	1.2202791	1.5908040
major	meta	0.9945723	0.0678979	0.8614949	1.1276496

```
run_simulation_meta(runs = 20)%>%
  knitr::kable()
```

heterogeneity	MSE_pooled	MSE_meta
minor	0.0181960	0.0153294
major	0.1267235	0.0121361