EI Difference in Difference Simulation

Michael Weaver

13/12/2021

Define Function to generate EI data

```
sim_ei_dd = function(n = 100) {
  out = list()
  #Fraction veterans, x
  x = rnorm(n) %>% plogis()

  #Fraction veteran support in time 0, that varies with x
  #veteran support at x = 0
  v00 = rnorm(1)
  #contextual effect for veterans
  v10 = rnorm(1)*x
  #error
  e_v0 = rnorm(1)
  #Fraction veteran support suffrage
  B_vi_0 = (v00 + v10 * x + e_v0)  %>% plogis

  #Fraction civilian support in time 0, that varies with x
  #civilian support at x = 0
  c00 = rnorm(1)
  #contextual effect for civilian
  c10 = rnorm(1)*x
  #error
  e_c0 = rnorm(1)
  #Fraction civilian support suffrage
  B_ci_0 = (c00 + c10 * x + e_c0)  %>% plogis

  #Suffrage at time 0:
  s0 = x*B_vi_0 + (1-x)*B_ci_0

  #Select a single uniform t1 - t0 shift for civilians, veterans,
  #among set of possible shifts
  diff_c_u = runif(1, 0 - min(B_ci_0), 1 - max(B_ci_0))
  diff_v_u = runif(1, 0 - min(B_vi_0), 1 - max(B_vi_0))

  #Select different t1-t0 shifts for civilians, veterans
  #independent of x, among shifts possible for all units
  diff_c_i = runif(n, 0 - min(B_ci_0), 1 - max(B_ci_0))
  diff_v_i = runif(n, 0 - min(B_vi_0), 1 - max(B_vi_0))

  #Select different t1-t0 shifts for civilians, veterans
  #among possible shifts in each unit (could be dependent on x)
```

```
  diff_c_alt = runif(n, 0 - B_ci_0, 1 - B_ci_0)
  diff_v_alt = runif(n, 0 - B_vi_0, 1 - B_vi_0)

  #Get suffrage vote at time 1
  s1_u = x*(B_vi_0 + diff_v_u ) + (1-x)*(B_ci_0 + diff_c_u)
  s1_i = x*(B_vi_0 + diff_v_i ) + (1-x)*(B_ci_0 + diff_c_i)
  s1_alt = x*(B_vi_0 + diff_v_alt ) + (1-x)*(B_ci_0 + diff_c_alt)

  #Truth
  out$att_u = mean(diff_v_u - diff_c_u)
  out$att_i = mean(diff_v_i - diff_c_i)
  out$att_alt =mean(diff_v_alt - diff_c_alt)

  #Truth: correlation between B_vi_0, B_ci_0 and x
  out$beta_vx_0 = lm(B_vi_0 ~ x) %>% coef %>% .[2]
  out$beta_cx_0 = lm(B_ci_0 ~ x) %>% coef %>% .[2]

  #Truth: correlation between B_vi_0, B_ci_0 and x
  out$beta_vx_diff_i = lm(diff_v_i ~ x) %>% coef %>% .[2]
  out$beta_cx_diff_i = lm(diff_c_i ~ x) %>% coef %>% .[2]
  out$beta_vx_diff_alt = lm(diff_v_alt ~ x) %>% coef %>% .[2]
  out$beta_cx_diff_alt = lm(diff_c_alt ~ x) %>% coef %>% .[2]

  #Ecological Regression Estimate
  out$att_hat_u = lm(s1_u - s0 ~ x) %>% coef %>% .[2]
  out$att_hat_i = lm(s1_i - s0 ~ x) %>% coef %>% .[2]
  out$att_hat_alt = lm(s1_alt - s0 ~ x) %>% coef %>% .[2]
return(as.data.table(out))
}
```

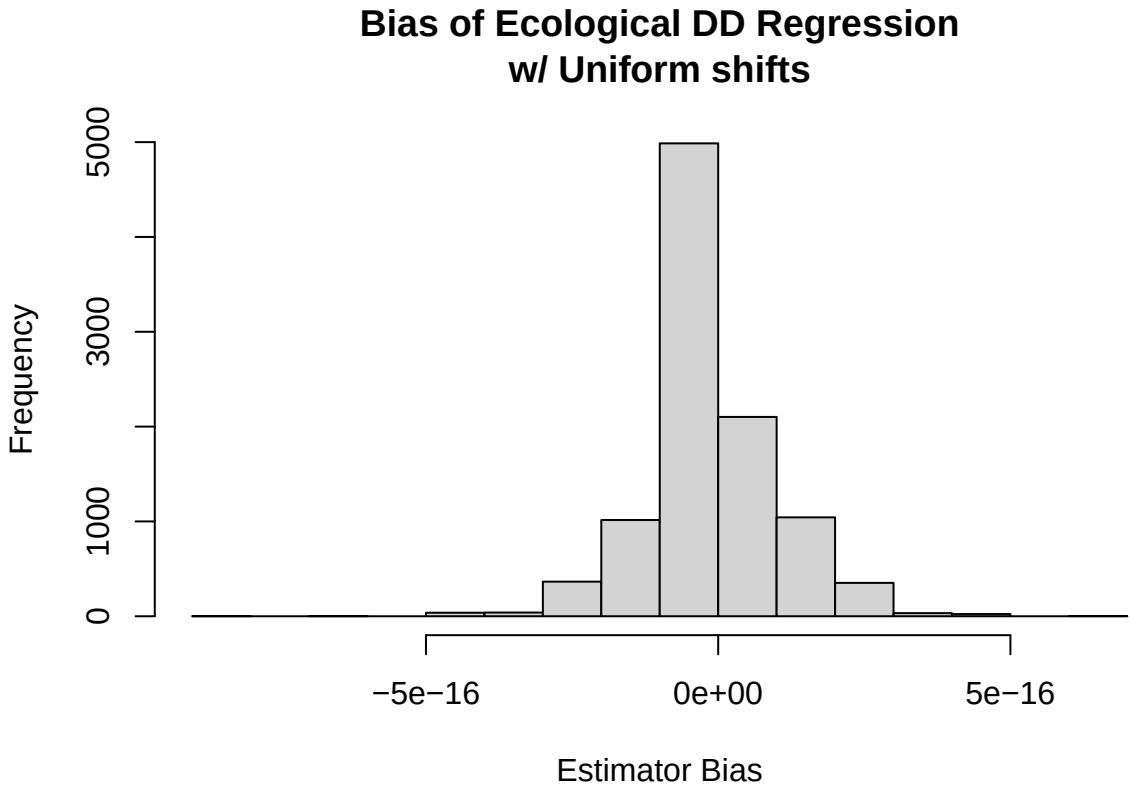
Generate 10000 simulations, $n = 100$

```
sim_data = lapply(1:10000, function(x) sim_ei_dd(100)) %>% rbindlist
```

Properties of ecological DD regression

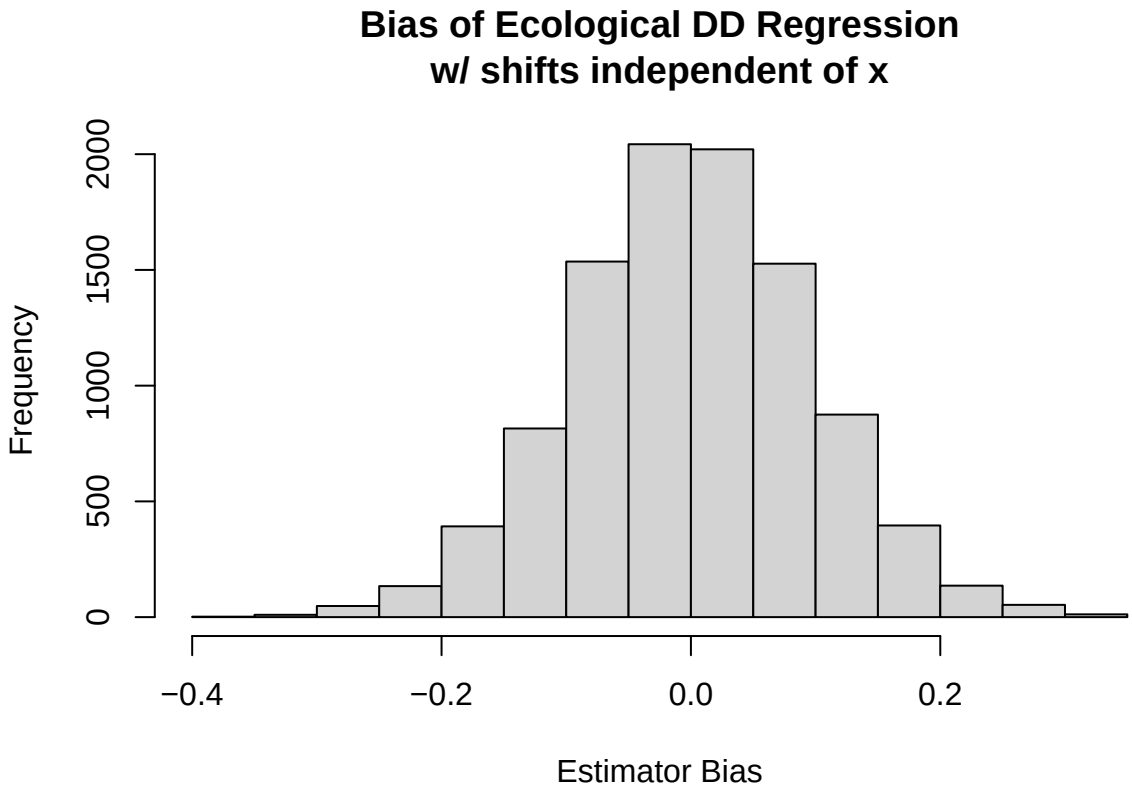
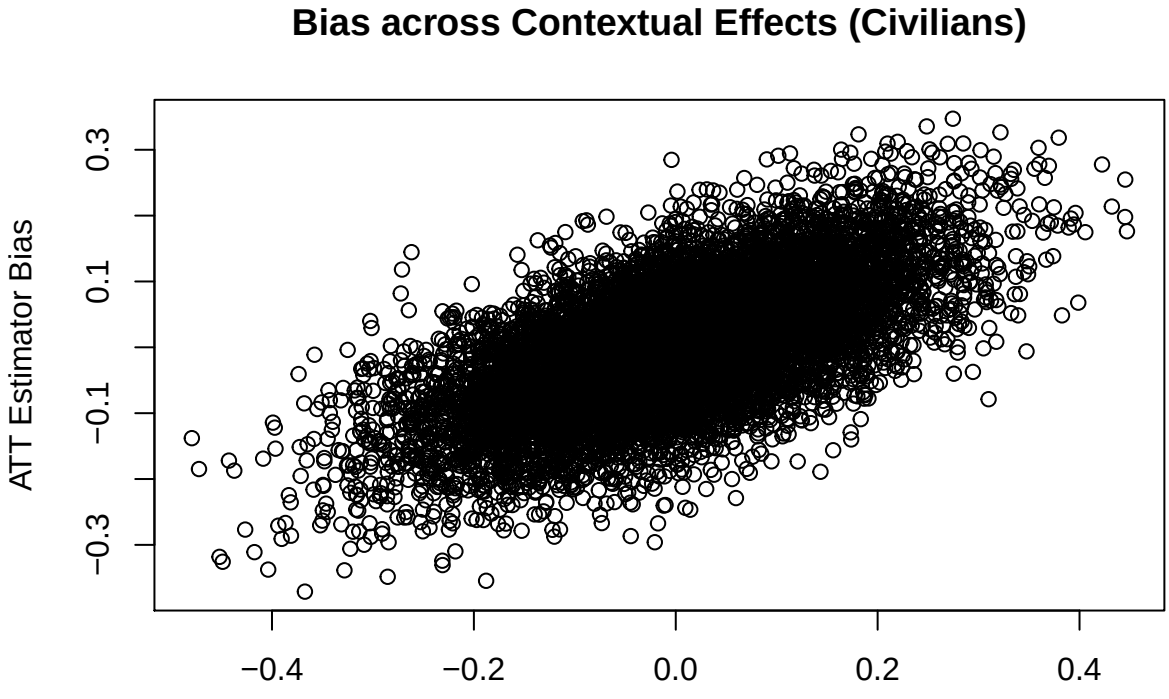
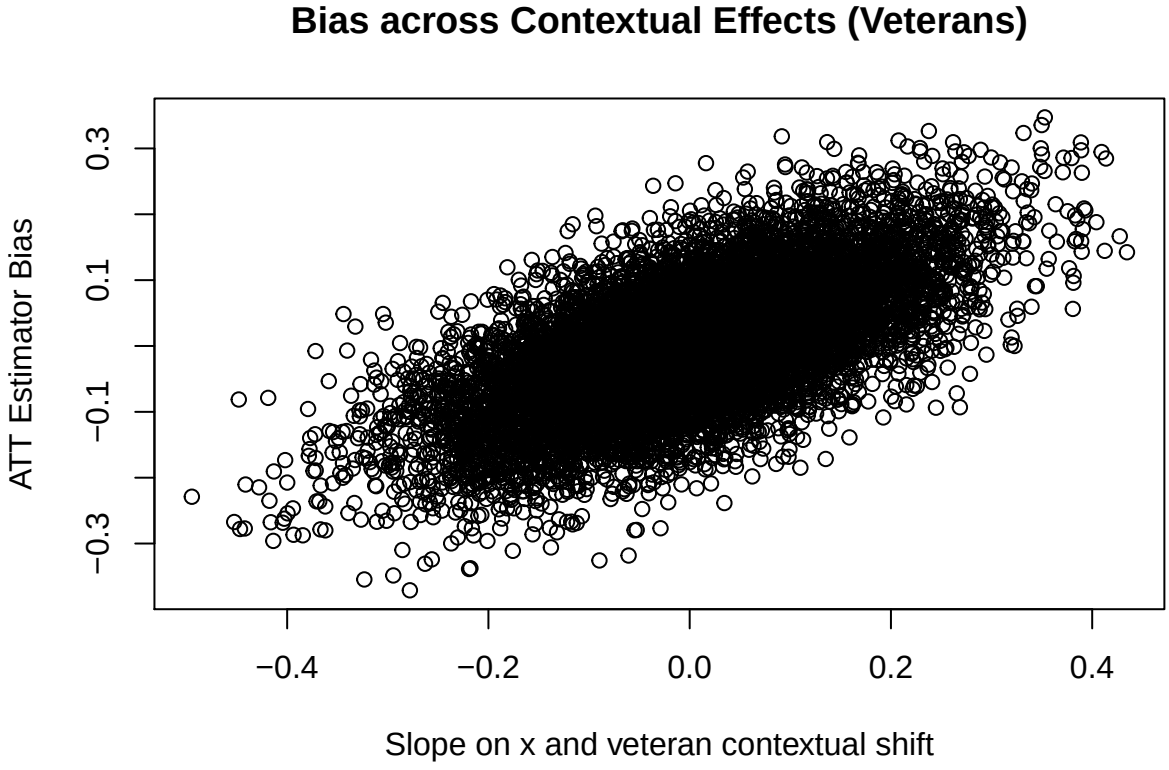
With a *uniform* shift from t_0 to t_1

- Average bias is $-8.1423041 \times 10^{-19}$



With a shift from t_0 to t_1 *independent* of X (no contextual effects)

- Average bias is 9.3934516×10^{-4}



With contextual shifts for veterans or civilians