

ECON 8000/9000 Empirical Energy Economics
PS7: Replication Exercise: Event Study
Christy Zhou

Instruction and Notes:

- Deadline: Sunday 4/19 @ 5pm
- Submission: Please compile all results into a single PDF.
- Group: Everyone can work together, but only **up to two students** can submit the assignment together. If so, make sure you include both names at the top when you submit.
- Here I write instructions using Stata. You can use other software if you like.
- This study has a replication package online. I recommend you follow my instructions below, which are slightly less complicated than the online package.

Q1. Replication Severnini (2022) “The Power of Hydro Dams: Historical Evidence from the US over the 20th Century”, EJ

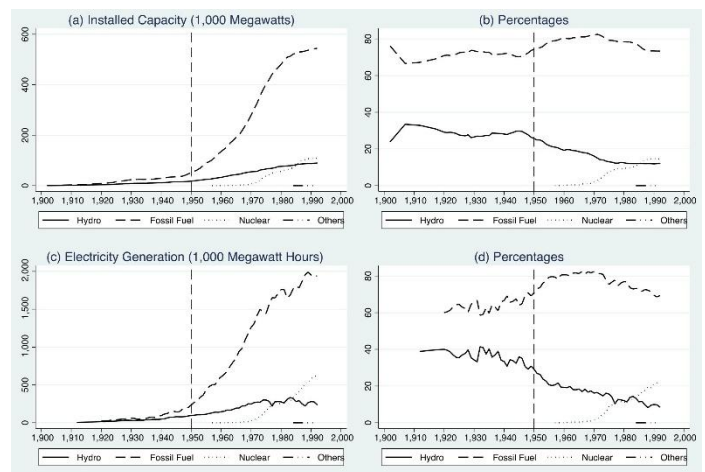
In this exercise, we will focus on producing event study estimates, in particular, the main results in Figure 4. We will first do a visualization exercise to analyze the temporal variation. We will then produce the main event study results in a few ways.

Q1.A Visualization Exercise

First, we will demonstrate the update of hydropower and reproduce Figure 1.

Key Info:

- Dataset: “[data_for_fig1.dta](#)”
- This is a yearly dataset.
- There are 4 types of energy sources:
 - Hydro (in variables ***_hydro**)
 - Fossil fuel (in variables ***_ffuel**)
 - Nuclear (in variables ***_nuclear**)
 - Other fuel (in variables ***_other_fuel**)
- There are four types of measures
 - Capacity in GW (in **cap_gw_***)
 - Capacity percentage (in **cap_p***)
 - Electricity generation in GW (in **gen_gwh_***)
 - Electricity generation percentage (in **gen_p***)



Tasks:

1. Reproduce each panel of Figure 1 which cap the plot at year 1992. You can do something like:
 - `twoway ///`
`(line var1 year if year <= 1992, options) ///`
`(line var2 year if year <= 1992, options) ///`
`(line var3 year if year <= 1992, options) ///`
`(line var4 year if year <= 1992, options), other options`
`graph export "path/filename.pdf", replace`
 - You can do it in other ways as well.
2. After plotting all four panels, describe what you observe.

Q2. Event Study

In this exercise, we will study the main event in Figure 4. This is obtained using an ES under a typical TWFE specification. We will also examine cohort-specific estimates and the plot using the Callaway and Sant'Anna (CS) estimator.

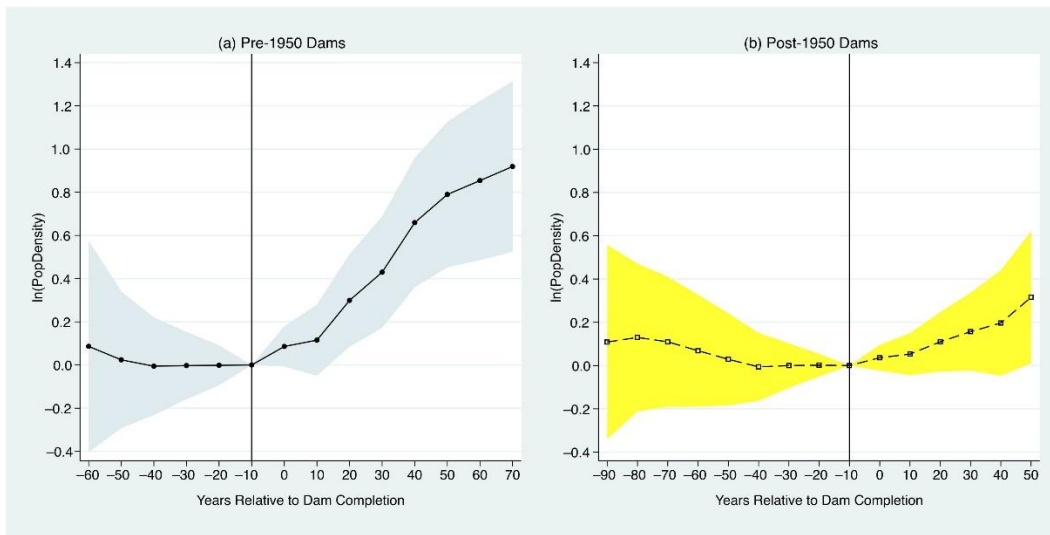


Fig. 4. *Impact of Hydro Dams on Population Density: Short- and Long-Run Estimates. (a) Pre-1950 Dams. (b) Post-1950 Dams.*

Info:

- Dataset: `"data_for_fig4.dta"`
- Key vars:
 - outcomes: `Indensity` (main outcome)
 - key cross-sectional identifier: `county_fips`
 - key temporal identifier: `year`
 - treatment timing (also the treatment cohort): `year_treat`
 - event time
 - In years: `evtime`
 - In decade: `evtime_decade` (this is `evtime/10`)
 - In dummies:
 - Dummies before event time = 0: `evtime*b` such as `evtime20b`
 - Dummy for even time = 0: `evtime0`
 - Dummies for post event time = 0: `evtime*a` such as `evtime20a`
 - Other vars for control
 - Please read page 437 under Eq(1) to see definitions of the geographic controls: `npff_*`, `reg_fips`, `latc1`, `latc2`, `latc3`, `lonc1`, `lonc2`, `lonc3`, `mt_*`, `mp_*`, `railtrack`, `waterway`, `distnearmsa_miles`, `inst1994`
 - Figure 4 separate analysis for pre-1995 dams, and post-1995 dams:
 - Figure 4 panel A: Limit sample to `"if (treated_b1950==1 | falsetreat==1)"`
 - Figure 4 panel B: Limit sample to `"if (treated_a1950==1 | falsetreat==1)"`

Tasks:

- 1. TWFE regression for both panels in Fig 4:
 - Run a `reghdfe` of `Indensity` on `"evtime100b-evtime20b evtime0-evtime80a"`. The coefficients for these dummies are what we will estimate
 - Control for `"npff_*`
 - Absorb the following in `abs()`:
 - `"i.county_fips i.year c.reg_fips#i.year c.latc1#i.year c.latc2#i.year c.latc3#i.year c.lonc1#i.year c.lonc2#i.year c.lonc3#i.year c.mt_summer1#i.year c.mt_fall1#i.year c.mt_winter1#i.year c.mt_spring1#i.year c.mp_summer1#i.year c.mp_fall1#i.year c.mp_winter1#i.year c.mp_spring1#i.year c.railtrack#i.year c.waterway#i.year c.distnearmsa_miles#i.year c.inst1944#i.year"`
 - You may want to write them in a macro, say, under `$x_fe`, as the list is long

- Recall for Figure 4 panel A: Limit sample to `"if (treated_b1950==1 | falsetreat==1)"`
- Recall for Figure 4 panel B: Limit sample to `"if (treated_a1950==1 | falsetreat==1)"`
- **Task 3.A Plot Panel A**
 - Task 3.A.1: Run a `csdid` and store the outcome.
 - Method 1:
 - Re-run Task 1 with `savefe`. Given we use default names, I recommend put `"cap drop __hdfc*"` before the `reghdfe` with `savefe` option
 - Run CSDID that similar to this `"csdid Indensity $x_ctrl __hdfc* $if_cond, ivar(county_fips) time(year) gvar(year_treat) cluser(county_fips) notyet long2"`. The key is to include the FEs (`__hdfc*`) that from `reghdfe`.
 - Method 2:
 - Re-run Task 1 without `"evtime100b-evtime20b evtime0-evtime80a"`. Generate residual `"Indensity_resid"` from the `predict` command.
 - Run CSDID using `Indensity_resid`, similar to this `"csdid Indensity_resid $x_ctrl $if_cond, ivar(county_fips) time(year) gvar(year_treat) cluser(county_fips) notyet long2"`.
 - You should observe coefficients for each cohort (4 of them) and weight.
 - What kind of group heterogeneity do you observe?
 - Task 3.A.2: Store the table version of the event study
 - For both methods:
 - Run `"estat event"` first just to take a look at the result
 - Run `"estat event, estore(a_name_you_choose)"`
 - Store result `"a_name_you_choose"` to a table
 - Task 3.A.3: Plot Event Study (ES)
 - For both methods, we can plot 3 ways:
 - Version 1: All years
 - Re-run 3.A.1 (aka `reghfe` with `savefe` and `csdid` with `__hdfc*` etc)
 - Run `"estat event"`
 - Run `"csdid_plot"` and store the graph
 - Version 2: A narrower window
 - Re-run 3.A.1 (aka `reghfe` with `savefe` and `csdid` with `__hdfc*` etc)
 - Run `"estat event, window(-50,80)"`
 - Run `"csdid_plot, xlabel(-50(10)80)"` and store the graph
 - Version 3:
 - Re-run Version 2 but remove `"long2"` in the option
 - Compare Version 1 vs Version 2 and your cohort-specific results from Task 3.A.1
 - Explain which cohort drives the estimates that you observe in the event study graph?
 - Compare Version 2 vs Version 3:
 - What happened to the estimate in the event year -10?
 - Do you know why? Speculate
 - Compare Version 2 with Figure 4 Panel A
- **Task 3.B Plot Panel B**
 - Repeat Task 3.A above for Panel B.
 - For version 2, choose `"window(-80 60)"` and `"xlabel(-80(10)60)"` for the plot
 - Compare Version 2 with Figure 4 Panel B