

### Front door

In a recent blog I highlighted how regression discontinuity could be an example of Pearl's front door criterion. This is interesting because examples are rare. It occurred to me that given their commonality a number of other natural experiment designs could fall under the front door criterion. Let's recap front door criterion. Basically it is when the exposure is shielded from confounding (C - often unobserved) by a variable. In Pearl's original formulation it is a mediator that is shielded by the exposure. The DAG below illustrates. The shielding variable in my example is time (T) but could be another variable that causes the natural experiment (E). Conditioning on time blocks confounding, and the effect of time on the outcome. This allows us to estimate the effect of the policy / intervention on the outcome (Y).



Y = outcome, E = policy / intervention, T = time, C = confounding

Natural experiments: similarities

Let's compare the following designs:

- Regression discontinuity.
- Interrupted time series (with and without control).
- Difference-in-difference.
- Synthetic control .

Now this blog is not about the technicalities of estimation from these designs. Also I simplify a lot for illustration. For example I use a linear relationship between the outcome and time, and I keep any differences in the pre intervention period between groups fixed. The intervention happened in 2006 in one population and not others.

Normally, regression discontinuity doesn't use time, as this is "weak" design, but conceptually it is the same as an interrupted time series with no control, as part A of the figure below argues.  $Y(1)$  and  $Y(0)$  are the outcome in the intervention group under the intervention (1) and no intervention (0).  $Y(0)$  is counter-to-fact (dashed-line) so we use the observed (solid lines) pre-intervention trend as a proxy ( $Y(0)_{pre}$ ).



Adding a control group ( $Y(com)$ ) that might have a different level of the outcome, but a similar pre-trend, leads us to a difference-in-difference design (panel B). This is similar to adding a control to the interrupted-time-series. Admittedly, difference-in-difference design often uses fewer pre and post time points.

Finally, in panel C synthetic control extends this by weighting a range of controls (represented by the shaded area around line) in the pre intervention period to mimic  $Y(0)_{pre}$  and then proxy  $Y(0)$ .

In all these designs time is a forcing variable, causing the intervention and allowing, with assumptions, confounding control. While we tend to treat each design separately, highlighting commonality may be useful, and the front door criterion is one way?

Does this make sense? Great to hear your views.

[R\\_code-for-replicating-figuresDownload](#)

[Tweet](#)