Ch. 9: Experiments with Longitudinal Elements

Dylan Baker

February 5, 2025

Contents

1 Main Takeaways

1.1 Gains from Pre-Treatment Outcome Measures

Consider the two parameters:

$$\begin{aligned}\tau^{DID} = \tau^{\text{post}} &- \tau^{\text{pre}} \\ &= \sum_{t>0} \left(\mathbb{E}\left[Y_{it} \mid D_{it} = 1\right] - \mathbb{E}\left[Y_{it} \mid D_{it} = 0\right]\right) \\ &\quad - \sum_{t\leq 0} \left(\mathbb{E}\left[Y_{it} \mid D_{it} = 1\right] - \mathbb{E}\left[Y_{it} \mid D_{it} = 0\right]\right)\end{aligned}$$

and

$$\tau^{\text{Post}} = \sum_{t>0} \left(\mathbb{E}\left[Y_{it} \mid D_{it} = 1\right] - \mathbb{E}\left[Y_{it} \mid D_{it} = 0\right]\right)$$

When you have pre-treatment observations, there are instances where it may be beneficial to include them in your specification and instances where it may not be. This depends on how many pre-treatment observations you have and the strength of the autocorrelation parameter.

Under the assumptions of

1. constant variance across time periods,
2. equal variances between all pairs of time points
3. autocorrelation in outcomes does not depend on treatment effects
4. treatment and control groups are identical in cross-sectional size

we have

$$\text{Var}\left[\tau^{\text{Post}}\right] = \frac{2\sigma^2}{n_d}\left(\frac{1+(N_{\text{Post}}-1)\,\rho}{N_{\text{Post}}}\right)$$

and

$$\text{Var}\left[\tau^{DID}\right] = \frac{2\sigma^2}{n_d}\left(\frac{1+(N_{\text{Post}}-1)\,\rho}{N_{\text{Post}}} - \frac{(N_{\text{Pre}}+1)\,\rho - 1}{N_{\text{Pre}}}\right)$$

Below is a guide:

Exhibit 9.2: Intuition for Choosing Difference-in-Means versus Difference-in-Differences

ρ	N_{Pre}	Variance Ordering	Preferred Choice?
0	1	$Var[\tau^{DID}] > Var[\tau^{Post}]$	Difference-in-Means
	5	$Var[\tau^{DID}] > Var[\tau^{Post}]$	Difference-in-Means
	10	$Var[\tau^{DID}] > Var[\tau^{Post}]$	Difference-in-Means
0.25	1	$Var[\tau^{DID}] > Var[\tau^{Post}]$	Difference-in-Means
	5	$Var[\tau^{Post}] > Var[\tau^{DID}]$	Difference-in-Differences
	10	$Var[\tau^{Post}] > Var[\tau^{DID}]$	Difference-in-Differences
0.5	1	$Var[\tau^{DID}] > Var[\tau^{Post}]$	Difference-in-Means
	5	$Var[\tau^{Post}] > Var[\tau^{DID}]$	Difference-in-Differences
	10	$Var[\tau^{Post}] > Var[\tau^{DID}]$	Difference-in-Differences

Figure 1: Choosing Diff-in-Diff versus Diff-in-Means

1.2 Choosing the Optimal Number of Pre- and Post- Periods

The smallest sample size correspond to a given MDE is given by:

$$n_0^* = n_1^* = n^* = \frac{2\left(t_{\alpha/2} + t_\beta\right)^2 \sigma^2}{(MDE)^2}\left(\frac{N_{\text{Pre}} + N_{\text{Post}}}{N_{\text{Pre}} * N_{\text{Post}}}\right)$$

As the number of periods increases from 4 to 16, the minimum sample size required to detect a given MDE decreases four-fold.

Questions

Look into this more.

1.3 Solomon 4-Group Design

One may worry that taking measures repeatedly may influence participants' responses (e.g., if they get tired of responding). One way to get around this is to use the Solomon 4-group design. Under this design, participants are randomly divided into to 2, where half are given a pre-test and half are not. Then, post-test outcomes are measured for all participants and compared based on whether they received a pre-test. If results are similar across groups, they can be pooled.

1.4 Causal Density

"The actual measurement of a long run treatment effect and the discovered underlying mechanisms at work depends on the outcome's causal density.

The causal density of an outcome variable refers to how well the determinants of the outcome are understood. An outcome with low causal density is one where the factors determining the outcome are well understood. For example, farmers have a near-perfect understanding of the determinants of crop yields, yielding a low causal density. In contrast, the determinants of labor market outcomes are much less well understood and thus have high causal density. For example, future employment and wage outcomes are a function of the numerous and interconnected inputs by multiple people and institutions."

"As the time between experiment and observation of the outcome increases, so do the number of potential mechanisms driving the effect and the causal density. For example, consider the question of whether financial incentives increase test scores. Suppose the researcher announces the incentives right before she administers the test. In that case, one can be reasonably sure that the only mechanism through which the treatment works is the student's effort on the test. Conversely, suppose that the researcher announces the financial incentives months before administering the test. Then, the student's effort remains a potential mechanism. However, possible changes to the student's test score could arise through her own study effort or even investment rom her parents, teachers, or peers. This general problem is one of forecasting long-term treatment outcomes, and its applicability is quite widespread."

1.5 Surrogates

When we are interested in the long-run effect of something but may be interested in expanding the policy before we can observe the long-run effect, we may use a surrogate to estimate the long-run effect. In this case, under strong assumptions, we consider the effect of the treatment on the surrogate and the effect of the surrogate on the long-run outcome.

The key assumptions are:

Notes

1. (A9.3, Comparability): $P_i \perp Y_i \mid Y_i^S$
   - That is, the "conditional distribution of the primary outcome given the surrogates is the same in the observational and experimental samples."
2. A linear relationship between the outcome and the surrogate
3. $\text{Var}\left[Y_i^S \mid D_i\right] = \text{Var}\left[Y_i \mid D_i\right]$
4. (A9.4, Surrogacy Condition): $D_i \perp Y_i \mid Y_i^S$
   - "The surrogacy condition requires that the surrogate fully captures the causal link between the treatment and the primary outcome. Under the surrogacy condition, there remains no treatment effect on the outcome after conditioning on the surrogate."

Under these conditions, we get:

$$\tau = \rho_{Y_i, Y_i^S} \cdot \tau^s$$

If the surrogacy condition doesn't hold, we instead get:

$$\tau = \rho_{Y_i, Y_i^S} \cdot \tau^S + \left(\mathbb{E}\left[Y_i \mid Y_i^S, D_i = 1\right] - \mathbb{E}\left[Y_i \mid Y_i^S, D_i = 0\right]\right)$$