Causal Effects with Fewer Negatives

Greg Leo

Abstract: I leverage the structure of commutative groups to reformulate traditional difference-in-differences (DiD) expressions into a Difference-of-Sums (DoS). This algebraic restructuring yields equivalent results while reducing the number of "-" symbols by approximately 66%. Furthermore, I demonstrate that DoS attains the lower bound on the number of "-" symbols among all expressions equivalent to DiD. While this article is generally tongue-(in+cheek), in the discussion, I speculate that the DoS expression might meaningfully simplify the interpretation of validity assumptions.

1 Introduction

The Difference in Difference estimator (DiD) is a simple and powerful causal estimator (Athey and Imbens, 2006). It is ubiquitous in empirical economics. 100% of all empirical economics papers written in the past five years use a DiD design.[-1] However, the expression's exclusive use of subtraction may create a significant cognitive burden (Adams et al., 2021).

2 Theory

Let $\mathscr{X} = \{y_{1,1}, y_{1,2}, y_{2,1}, y_{2,2}\}$ be a set of variables that serve as a basis for a free abelian group F. Throughout, I assume evaluations of these variables take place in an arbitrary commutative group $(G,+)$ where subtraction is defined as the addition of the additive inverse: for all $a, b \in G$, $a - b := a + (-b)$.

I define the **set of expressions** $\mathscr{E}$ inductively as the smallest set such that if $x \in \mathscr{X}$, then $x \in \mathscr{E}$, and if $A, B \in \mathscr{E}$, then $(A+B) \in \mathscr{E}$ and $(A-B) \in \mathscr{E}$. The Difference-in-Differences (DiD) and Difference-of-Sums (DoS) Expressions are

$$E_{\text{DiD}} \equiv \left(y_{1,1} - y_{1,2}\right) - \left(y_{2,1} - y_{2,2}\right), \tag{1}$$
$$E_{\text{DoS}} \equiv \left(y_{1,1} + y_{2,2}\right) - \left(y_{1,2} + y_{2,1}\right). \tag{2}$$

The **minus-count function** $N^- : \mathscr{E} \to \mathbb{N}$ is defined recursively: (1) $N^-(x) = 0$ for all $x \in \mathscr{X}$. (2) $N^-(A+B) = N^-(A) + N^-(B)$. (3) $N^-(A-B) = N^-(A) + N^-(B) + 1$.

I define equivalence of two expressions as follows. For any commutative group G and any valuation map $v : \mathscr{X} \to G$, there is a unique extension $\bar{v} : \mathscr{E} \to G$ that respects the group operations. $E_1, E_2 \in \mathscr{E}$ are **equivalent**, denoted $E_1 \equiv E_2$, if $\bar{v}(E_1) = \bar{v}(E_2)$ for all groups G and all valuations v.

Lemma 1. For all $a, b \in G$, $-(a+b) = -a - b$

Proof of Lemma 1. Using associativity and commutativity, $(a+b) + (-a-b) = (a-a) + (b-b) = 0 + 0 = 0$. Thus $-a-b$ is an additive inverse of $a+b$. The result follows from uniqueness of additive inverses. □

Proposition 1. $E_{\text{DiD}} \equiv E_{\text{DoS}}$

Proof of Proposition 1. Let G be any commutative group and v any valuation, $\bar{v}(E_{\text{DiD}}) = \left(v(y_{1,1}) - v(y_{1,2})\right) - \left(v(y_{2,1}) - v(y_{2,2})\right)$. By Lemma 1 and commutativity of G , ... $= \left(v(y_{1,1}) + v(y_{2,2})\right) - \left(v(y_{1,2}) + v(y_{2,1})\right) = \bar{v}(E_{\text{DoS}})$. Since $\bar{v}(E_{\text{DiD}}) = \bar{v}(E_{\text{DoS}})$ for all G, they are equivalent. □

Thus, DiD and DoS are equivalent. Note that $N^-(E_{\text{DiD}}) = 3$, $N^-(E_{\text{DoS}}) = 1$. I now prove that DoS attains the lower bound on the number of minuses of all expressions equivalent to DiD, which follows as a corollary from the following lower bound:

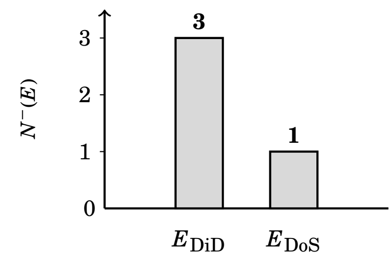

Figure 1: Minus-count comparison for E_{DiD} and E_{DoS}

Proposition 2 (Minus-Count Bound)**.** Let $E \in \mathscr{E}$ be any expression equivalent to E_{DiD}. $N^-(E) \geq N^-(E_{\text{DoS}}) = 1$.

Proof of Proposition 2. Assume for contradiction that there exists an expression $E \in \mathscr{E}$ with $N^-(E) = 0$ such that $E \equiv E_{\text{DiD}}$.

Since $\mathscr{X}$ is a basis for the free abelian group F, evaluating E in F yields a unique linear combination $\bar{v}_F(E) = \sum_{x \in \mathscr{X}} n_x x$, where $n_x \in \mathbb{Z}$. Because $N^-(E) = 0$, the expression E contains no subtraction operations. By structural induction on the grammar of $\mathscr{E}$, it follows that every coefficient n_x must be non-negative ($n_x \geq 0$).

However, the unique basis representation of the Difference-in-Differences expression in F is: $E_{\text{DiD}} = 1y_{1,1} - 1y_{1,2} - 1y_{2,1} + 1y_{2,2}$. Specifically, the coefficient for $y_{1,2}$ is -1.

Since $E \equiv E_{\text{DiD}}$, the uniqueness of the basis representation requires $n_{y_{1,2}} = -1$. This contradicts the requirement that $n_{y_{1,2}} \geq 0$. Therefore, any equivalent expression must have $N^-(E) \geq 1$. □

3 Discussion

The DoS expression might meaningfully simplify the interpretation of validity assumptions.

References

Adams, G.S., Converse, B.A., Hales, A.H., Klotz, L.E., 2021. People systematically overlook subtractive changes. Nature 592, 258–261.

Athey, S., Imbens, G.W., 2006. Identification and inference in nonlinear difference-in-differences models. Econometrica 74, 431–497.

[-1] This figure has been rounded up to the nearest 100%.