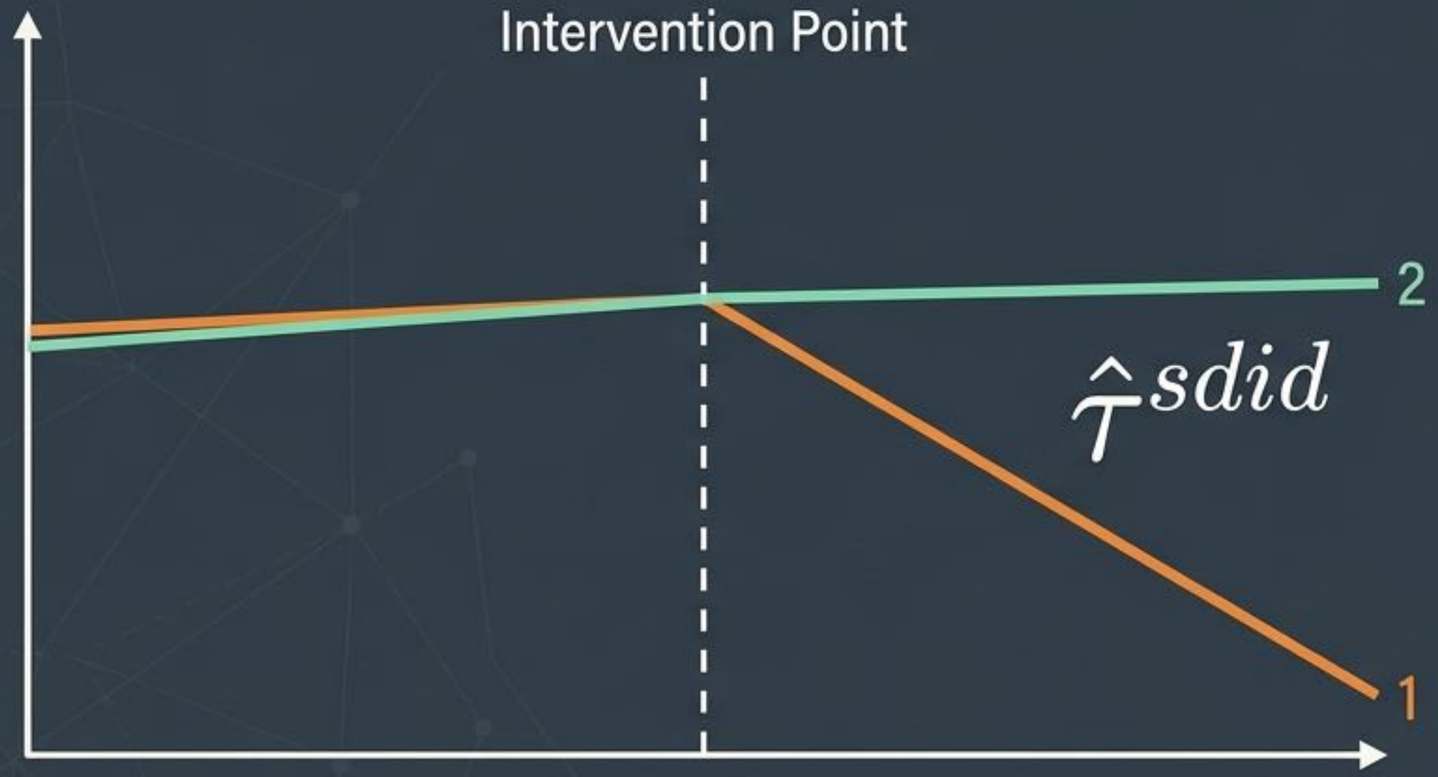
Synthetic Difference-in-Differences

A Unified TWFE Framework for Comparative Case Studies: Re-evaluating Proposition 99



Study Note

Target:

Graduate-level econometric reference guide.

Prerequisites:

- Two-Way Fixed-Effects (TWFE)
- Standard Difference-in-Differences (DiD)
- Synthetic Control (SC)

Source Data:

Abadie et al. (2010) via sdid Stata package.

Defining the Estimand: The Average Treatment Effect on the Treated (ATT)

$$\tau = \frac{1}{N_{tr} T_{post}} \sum_{i:W_i=1} \sum_{t>T_{pre}} [Y_{it}(1) - Y_{it}(0)]$$

Observed Outcome: California's actual cigarette sales after 1989.

Missing Counterfactual: What California would have sold without Prop 99. Every estimator simply imputes this unobserved value.

Study Note

The Prop 99 Panel Data

Problem Frame: We only have one treated unit ($N_{tr} = 1$). We must synthesize the missing data from a donor pool.

Dataset Specs:

- **Units:** 39 US states (1 CA + 38 Controls)
- **Time:** 1970–2000 (19 pre-treatment, 12 post-treatment years)
- **Treatment:** 1989 (Proposition 99 passed)
- **Outcome** (Y_{it}): Cigarette sales in packs per capita.

Three Estimators, One Unified Equation

$$\left(\hat{\tau}, \hat{\mu}, \hat{\alpha}, \hat{\beta}\right) = \arg\min \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \, \hat{\omega}_i \hat{\lambda}_t$$

α_i

Unit Fixed Effect
(Absorbs level gaps)

β_t

Time Fixed Effect

$\hat{\omega}_i$

Optimized

Unit Weight Dial

$\hat{\lambda}_t$

Optimized

Time Weight Dial

Study Note

The Aha Moment

Arkhangelsky et al. (2021) reveal that classical Difference-in-Differences, Synthetic Control, and SDID are mathematically identical.

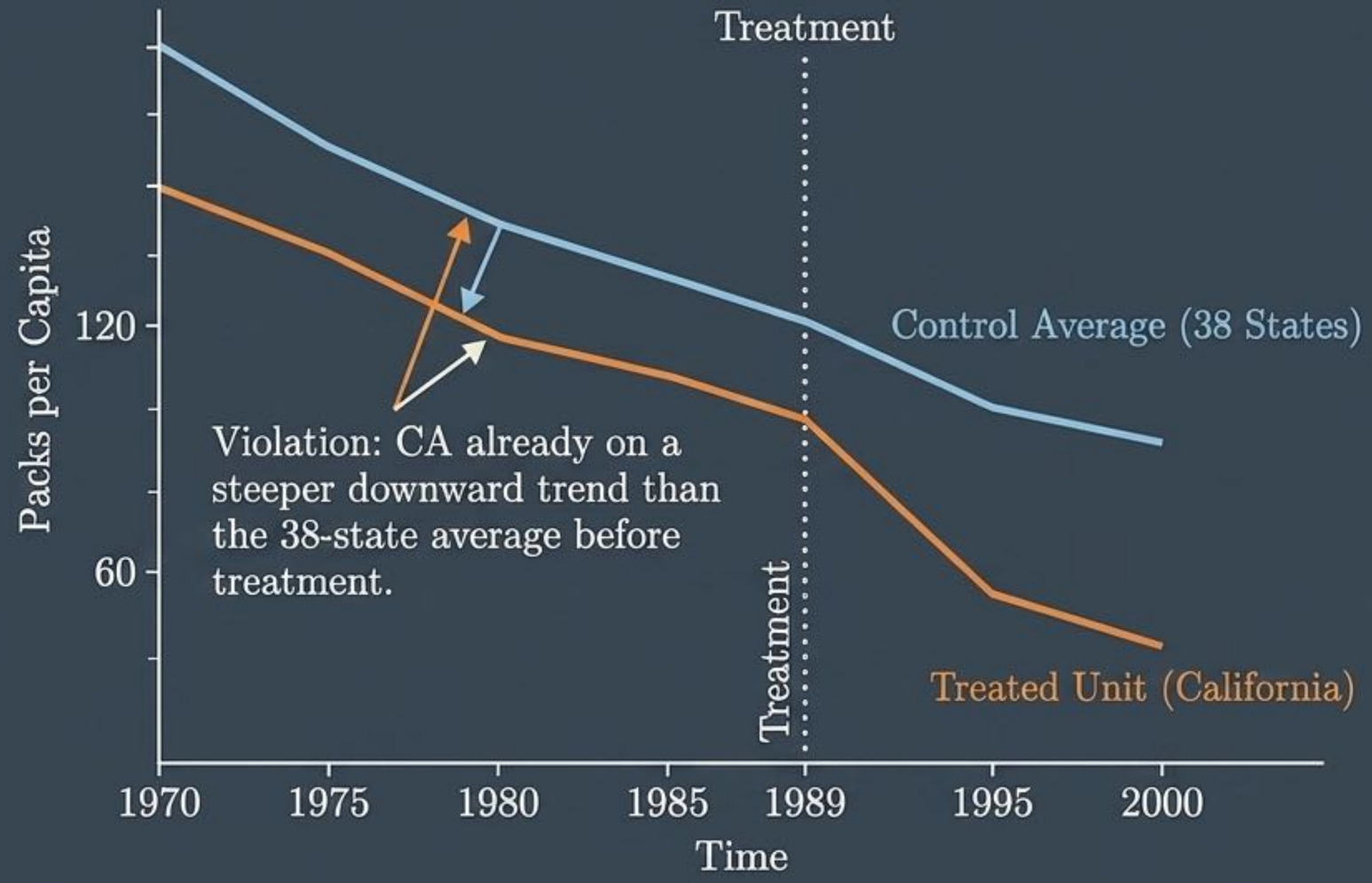
By simply turning the $\hat{\omega}_i$ and $\hat{\lambda}_t$ dials from "uniform" to "optimized", and toggling the unit fixed effect α_i, one framework generates all three.

The Estimator Typology Matrix

Estimator	Unit Weights (ω)	Time Weights (λ)	Unit FE (α_i)	Required Assumption
Difference-in-Differences (DiD)	Uniform	Uniform	Yes	Parallel trends vs. all controls.
Synthetic Control (SC)	Optimized	None	No (Crucial constraint)	Exact match on California's level AND trend.
Synthetic DiD (SDID)	Optimized	Optimized	Yes	Match California's trend (constant level gap allowed).

Insight: SC drops the unit fixed effect (α_i), forcing it to build a synthetic state that perfectly matches both baseline level and trajectory. SDID restores α_i, requiring only a parallel trajectory.

Difference-in-Differences: The Parallel Trends Vulnerability



Analysis

Equation Dials:

- ω = uniform
- λ = uniform

Prop 99 Estimate:

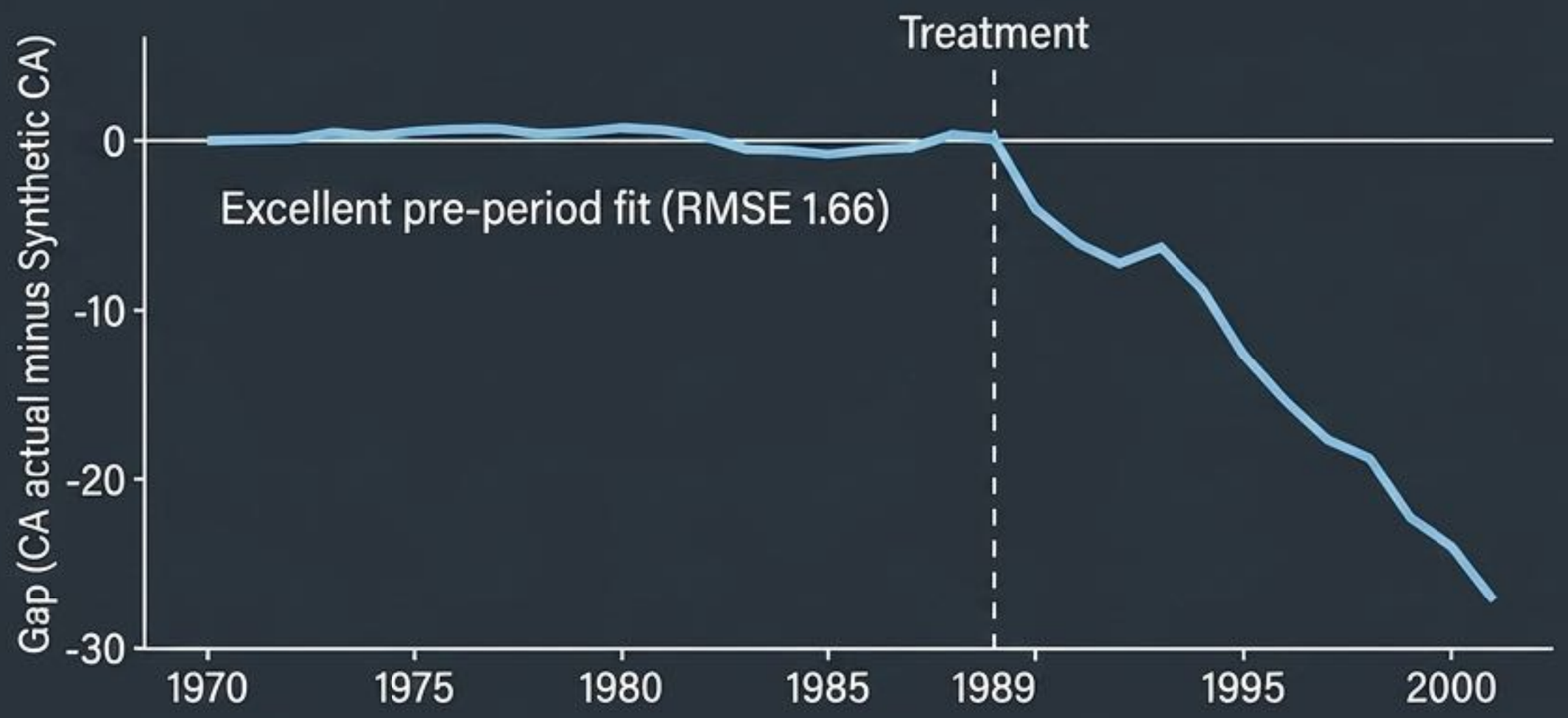
-27.35 packs per capita

Critique: Because the raw control average was drifting away from CA before treatment, the core parallel trends assumption fails. DiD drastically overstates the treatment effect by attributing prior trend differences to the policy.

Synthetic Control: Matching Levels and Trends

$$\text{argmin} \sum\sum (Y_{it} - \mu - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sc}$$

Note: α_i is intentionally dropped to force level matching.



Analysis

- **Weights:**
  Highly concentrated on just 6 donor states (Utah 0.39, Montana 0.23, Nevada 0.21).
- **Prop 99 Estimate:**
  -19.48 packs per capita (synth2)
- **Critique:**
  Dropping the unit fixed effect forces the algorithm to find states that match CA's exact smoking level. This is a highly demanding requirement that shrinks the donor pool drastically.

SDID Unit Weights (ω): Diffuse Matching with Level Gaps

$$\hat{\omega}^{sdid} = \arg\min \sum_{t=1}^{T_{pre}} \left(\omega_0 + \sum \omega_i Y_{it} - Y_{tr,t}\right)^2 + \zeta^2 T_{pre} \|\omega\|_2^2$$

1

Intercept (ω_0): Allows matching the trend without strictly matching the baseline level. The level gap is absorbed by the restored fixed effects.

2

Ridge Penalty (ζ^2): Forces diffuse, distributed weights. Prevents the algorithm from relying idiosyncratically on just 1 or 2 donor states.

Implementation

Donor Pool:
Weight is spread widely across ~20 states, rather than SC's concentrated 6.

Top Donors:

- Nevada (0.12)
- New Hampshire (0.11)
- Connecticut (0.08)
- Delaware (0.07)

Result:
Yields a more stable, less idiosyncratic comparison group than classical SC.

SDID Time Weights (λ): Emphasizing the Relevant Past

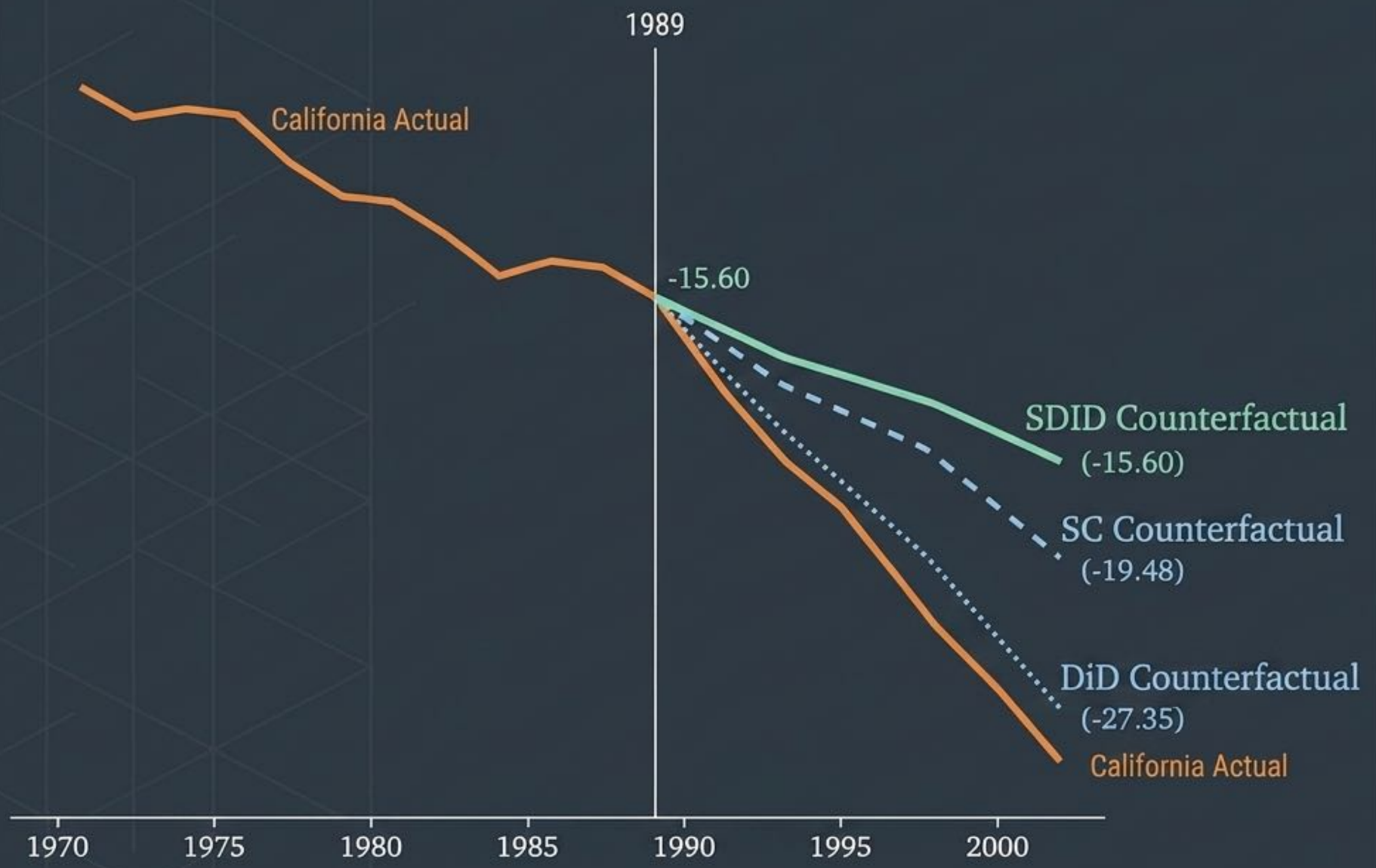
SDID seeks **pre-period years** whose **weighted average aligns with the post-period control averages.** Years that look most like the post-period get the most weight.

Time-Weight Heatmap

The Core Intuition

- Smoking behavior in 1972 tells us very little about the counterfactual outcome for 1995.
- The late 1980s are vastly more predictive of post-treatment behavior.
- Both DiD and SC mathematically fail by incorrectly treating 1972 and 1988 as equally informative.

The Counterfactuals Compared: Proposition 99 Estimates



Synthesis

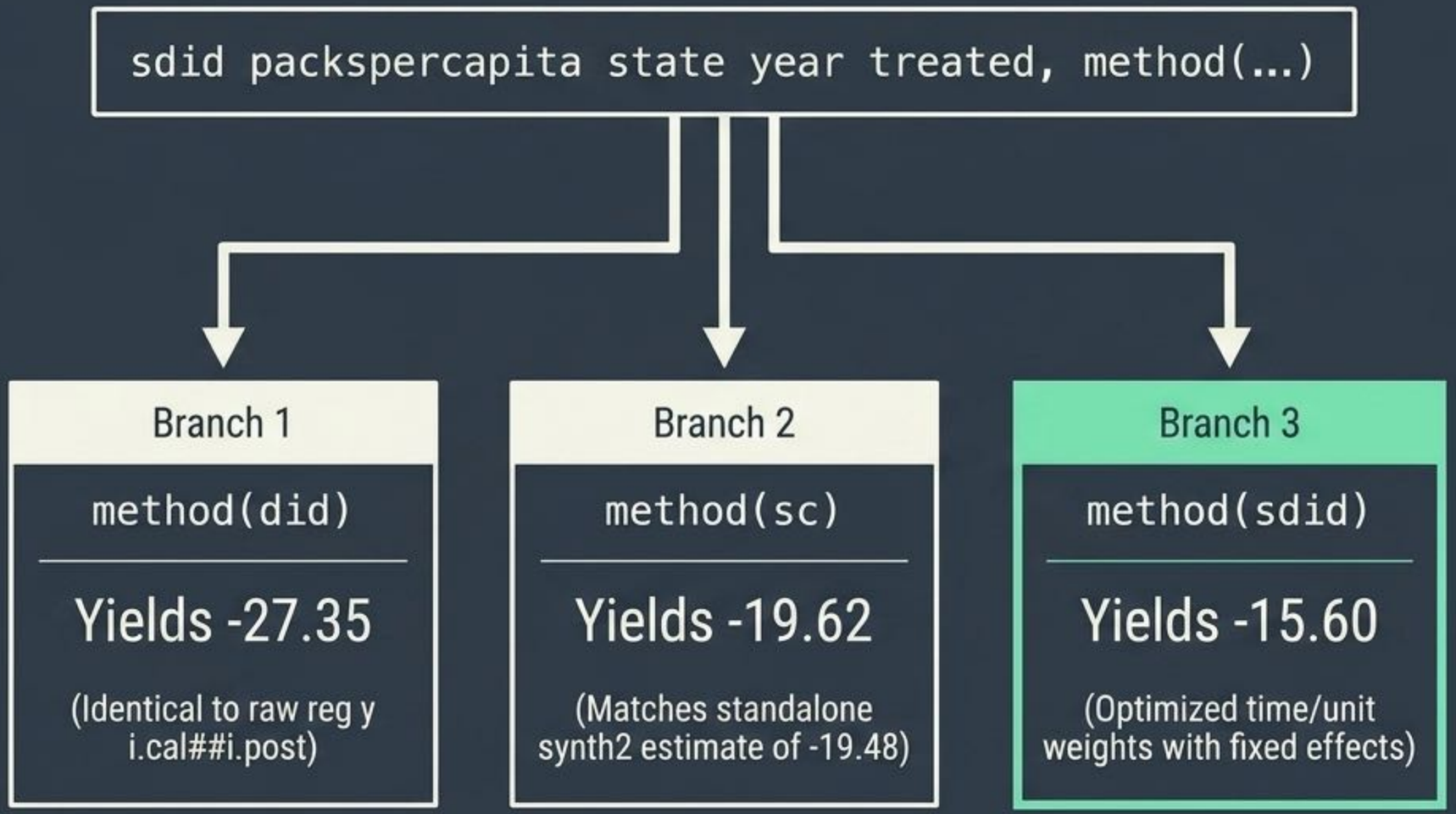
Observation:

SDID's synthetic control line actually sits ABOVE California. It does not attempt to close the structural level gap (which is safely absorbed by fixed effects); it merely ensures a parallel trajectory.

Result:

SDID yields the most conservative and credible estimate: roughly a 20% reduction (-15.60 packs/capita).

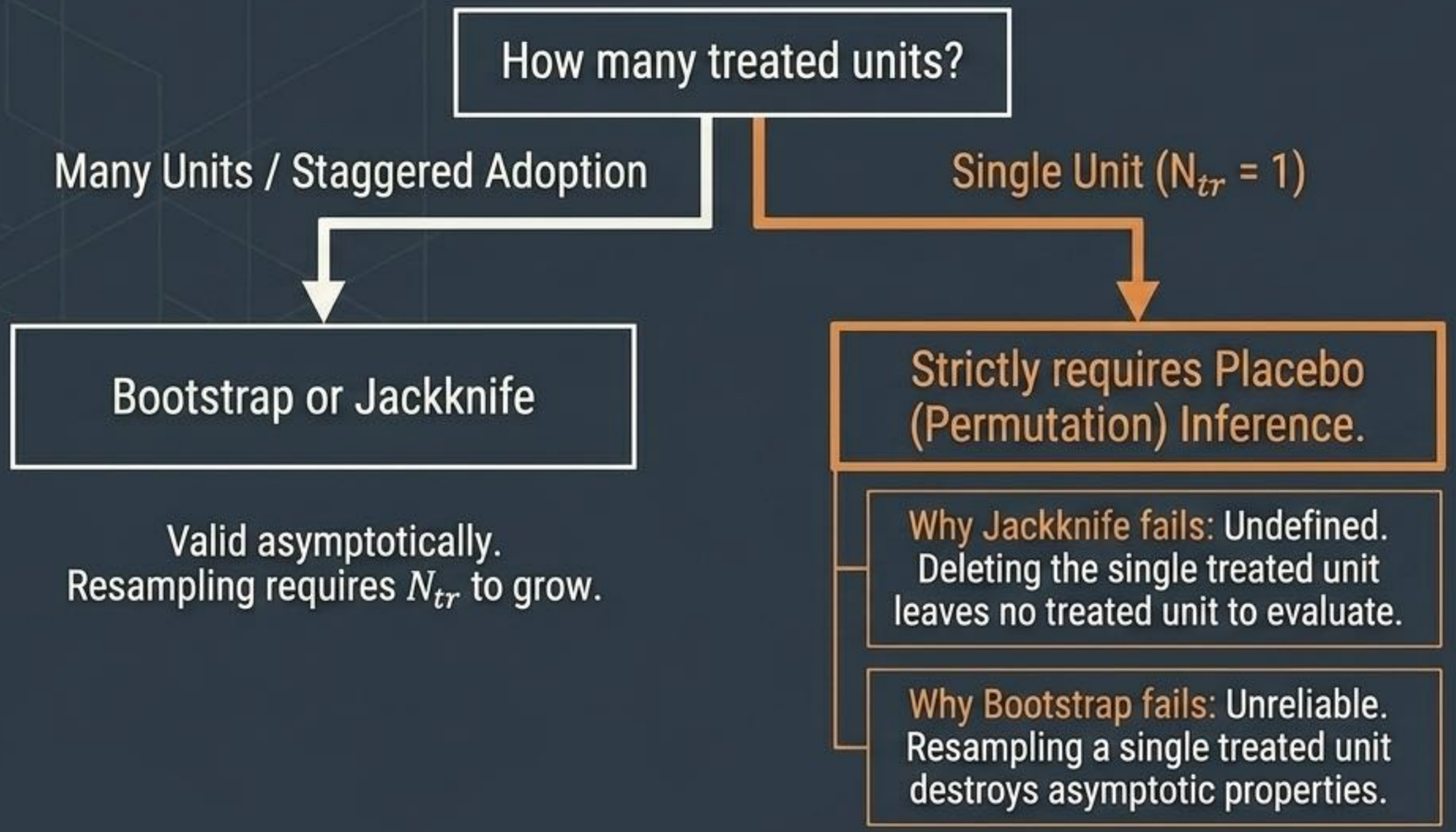
Implementation: One Framework, One Command



Computational Efficiency

Because optimal weights are mathematically computed once and reused across variance estimators, this unified framework is computationally cheap and theoretically perfectly consistent.

Inference Decision Matrix: Dealing with $N_{tr} = 1$



Implementation

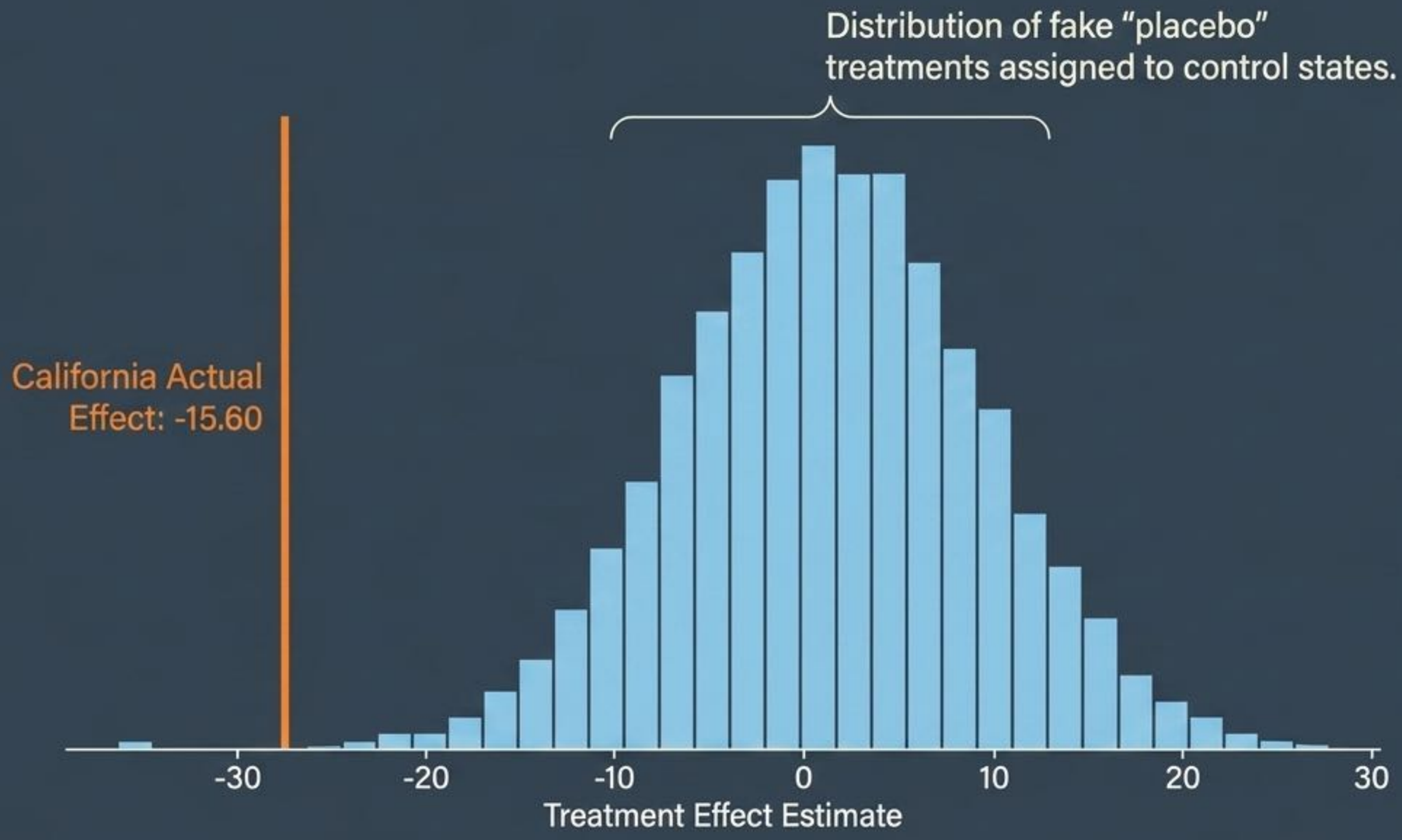
Stata Syntax:

Append vce(placebo) to the base sdid command.

Mechanism:

Repeatedly assign fake "placebo" treatments to control states, re-estimate the mathematical framework, and build a variance distribution to test the actual estimate against.

Placebo Distribution: Permutation Test Mechanics



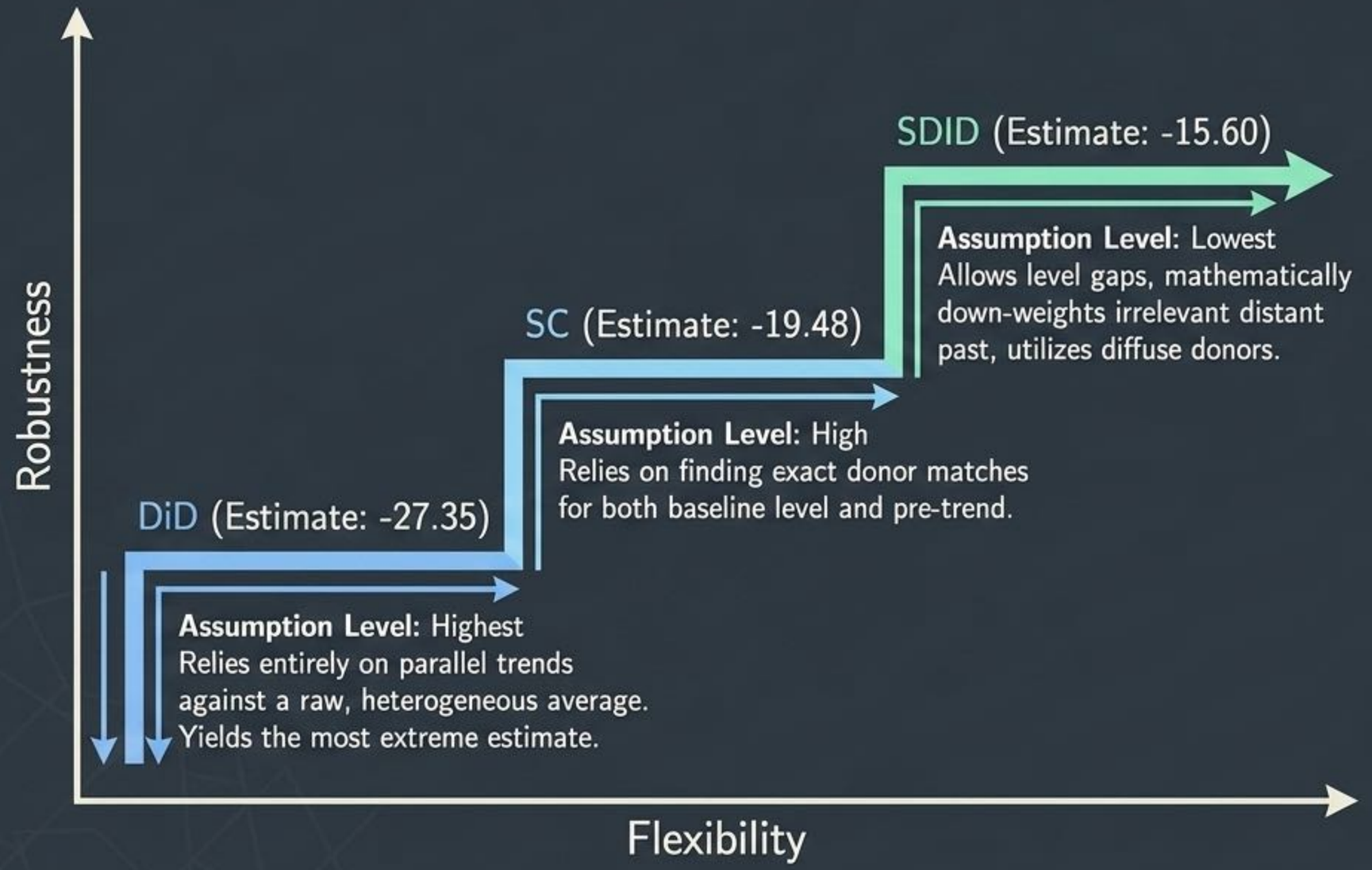
Study Note

Statistical Power

Normal Approximation: SE = 9.88. 95% Confidence Interval is [-35.0, 3.8]. Because the interval includes zero (p=0.114), this demonstrates the inherently low power of any $N_{tr} = 1$ study design.

Rank-Based Permutation: Only 1 out of 38 donor states produced a placebo effect greater than or equal to California's actual magnitude. Permutation p-value: $p = 0.026$. Statistically significant.

Synthesis: The Assumption-Robustness Trade-Off



Conclusion

All methods definitively agree that Proposition 99 reduced smoking in California.

However, as an estimator becomes more flexible and relies less on exact parallel trends, it yields a vastly more conservative, credible estimate.

SDID brackets the empirical truth most reliably at roughly –16 to –20 packs per capita.