

The Science of Incrementality

Demonstrating technical mastery in distinguishing correlation from causation to drive marketing efficiency and product growth.

The Three Pillars of Measurement



Experimental

Randomized assignments (RCTs) and Switchback designs. The gold standard for eliminating selection bias in native platforms.



Quasi-Experimental

Leveraging natural variations. DiD and Synthetic Control methods for when randomization is restricted by business constraints.

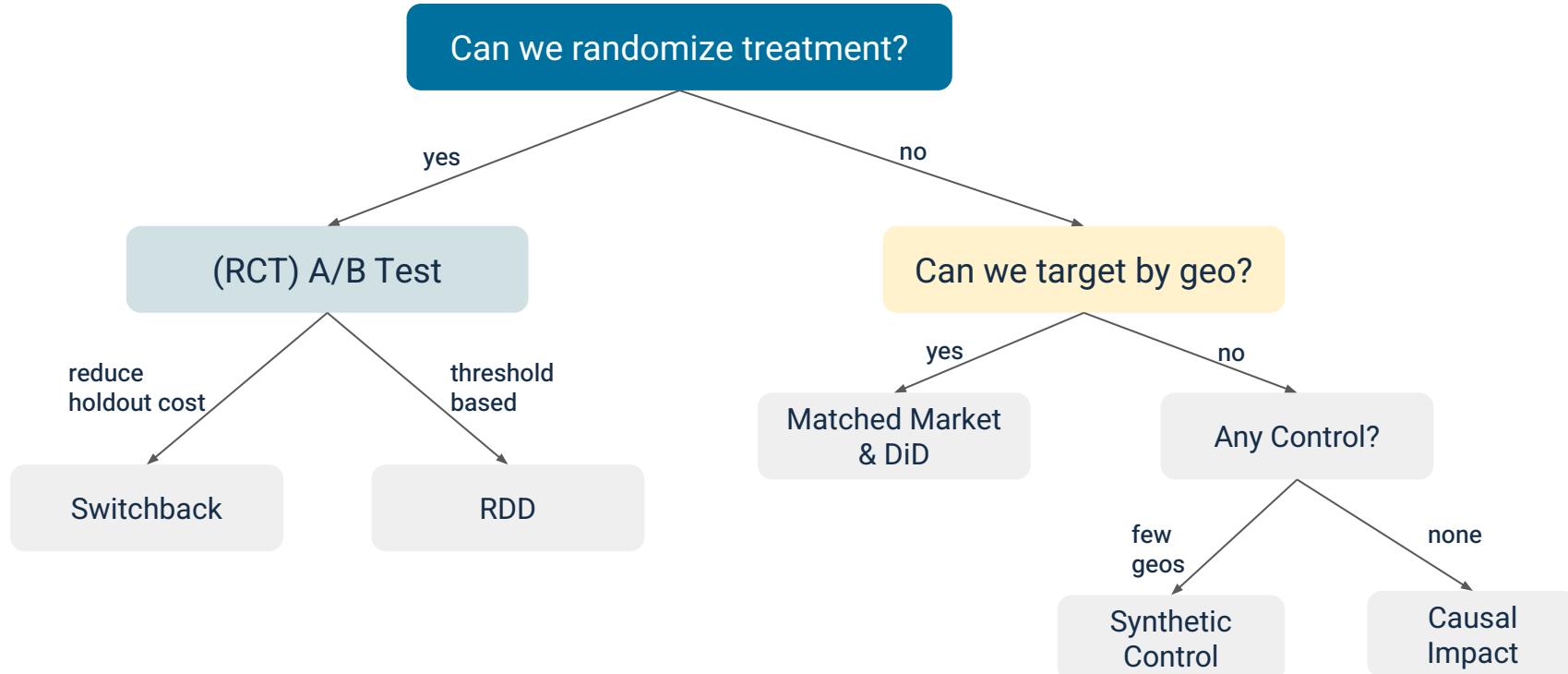


Model-Based Inference

Counterfactual forecasting via Bayesian Structural Time Series (Causal Impact) and observational bias correction via Propensity Score Matching and Inverse Propensity Weighting.

Which Incrementality Method Should We Use?

A decision framework based on what you can control in your experiment setup



A/B Testing (RCT)

Commonly applicable for measuring **Display, Video, Search, Email, SMS, DM**

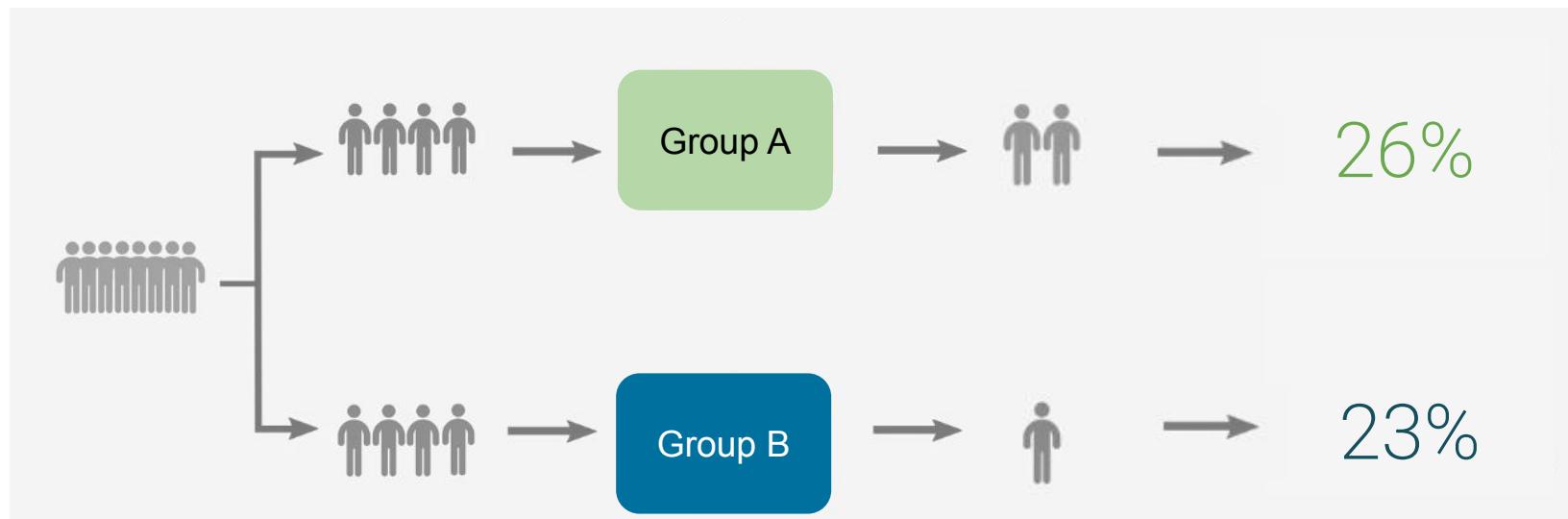
1. Randomly split eligible audience into 2 equal groups



2. Push ad to **Group A**, not Group B



3. (ATE) Incr. Conversion% =
 $+3\% \pm CI$



ATE: Average Treatment Effect for the overall population

© 2026 Selina (Xian) Xiao

Switchback / Crossover

Pulses of ON/OFF over time within the same unit. Best for **Marketplaces** (e.g., Uber, DoorDash).

1. Alternate treatment ON/OFF across time periods



2. Each unit serves as its own control at different times



3. (ATE) Incr. Conversion% =
 $+3\% \pm CI$



Regression Discontinuity Design (RDD)

Treatment is determined by a threshold/cutoff, e.g., **loyalty program** tiers, spend-based promotions.

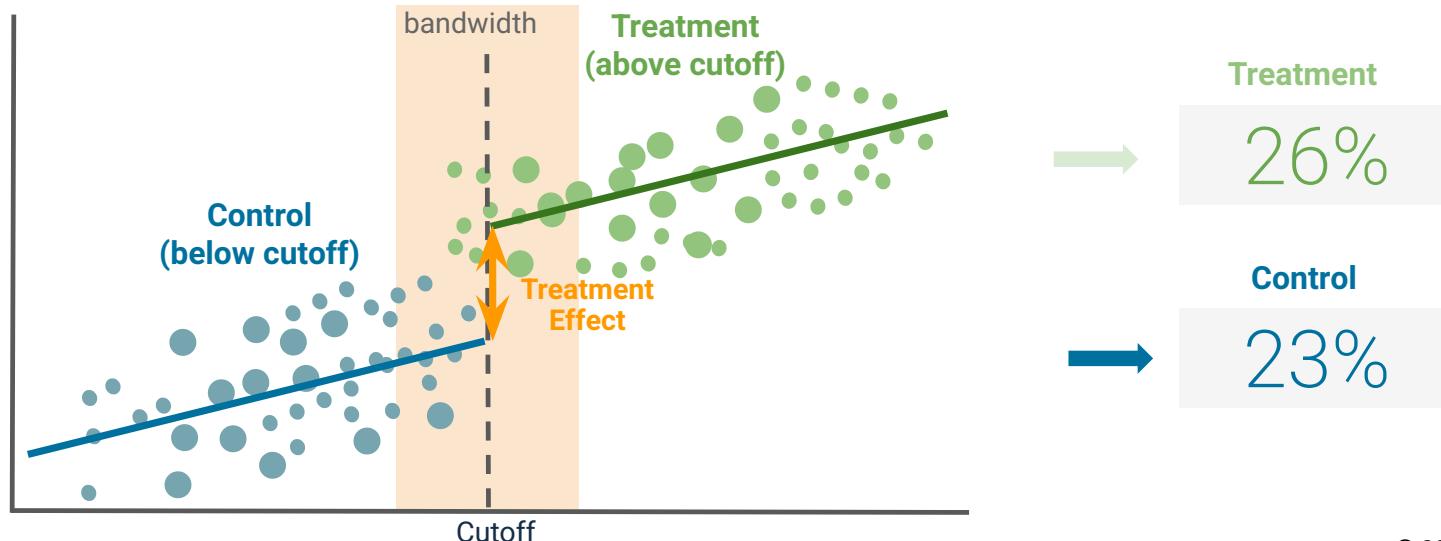
1. Treatment assigned by a cutoff (e.g., VIP = \$1,000+ annual spend)



2. Compare individuals just above vs. just below the cutoff (within a bandwidth)



3. (LATE) Incr. Conversion% =
 $+3\% \pm \text{CI}$

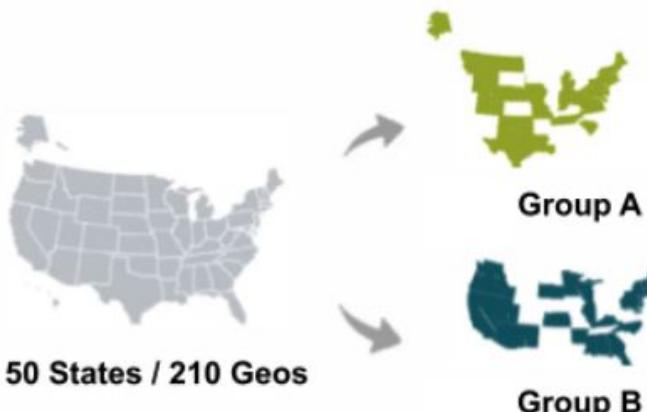


LATE: Local Average Treatment Effect, the causal effect at the cutoff, not generalizable to the overall population

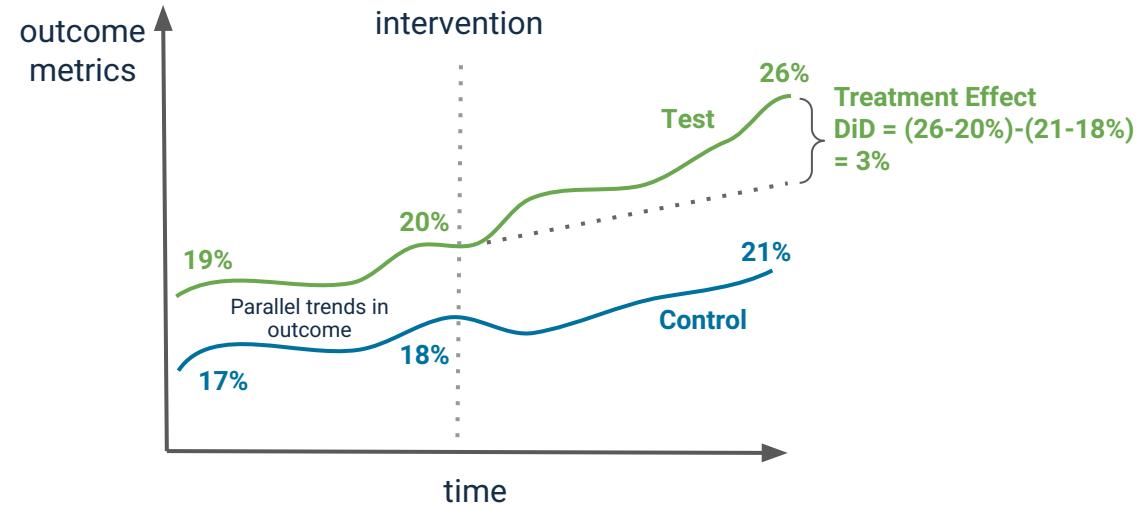
Matched Market & DiD

Commonly applicable for measuring user-level (**digital, CRM**) and geo holdout-applicable channels (e.g., **CTV/OTT, Podcasts**)

1. Select two groups of geos with parallel trends in outcome metrics



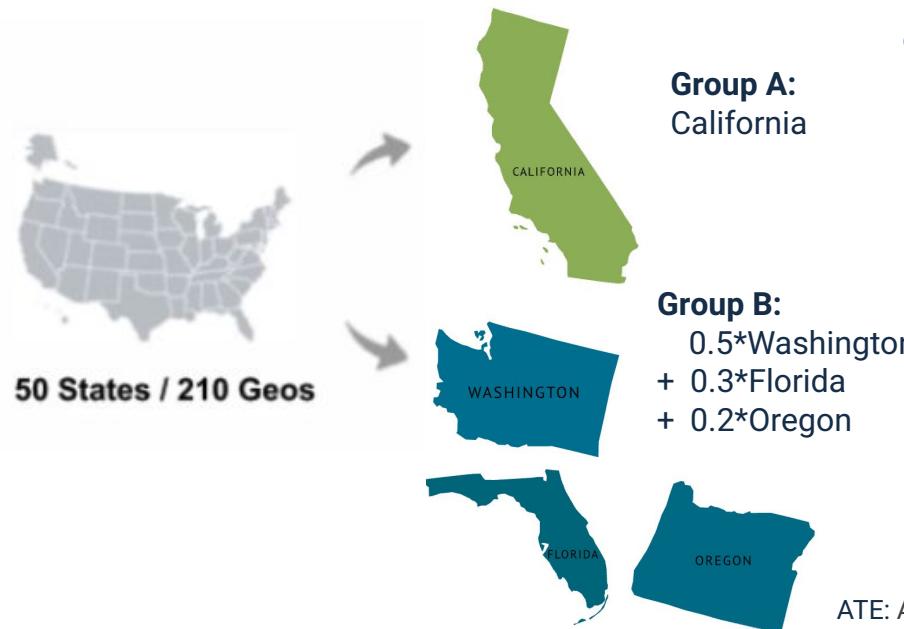
2. Push ad to **Group A**, not Group B



Synthetic Control

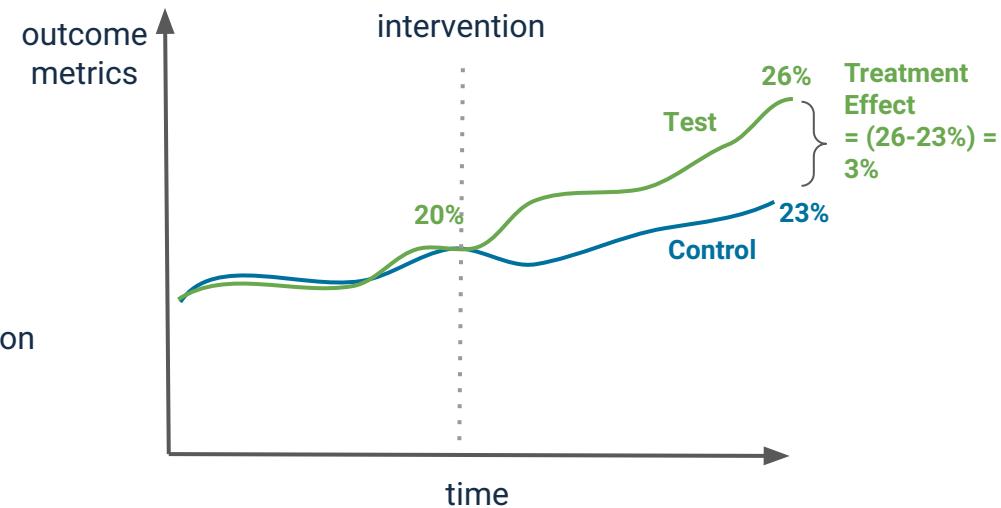
Commonly applicable for measuring user-level (**digital, CRM**) and geo holdout-applicable channels (e.g., **CTV/OTT, Podcasts**)

1. Create a weighted combination of geos that closely matches Group A's behavior



2. Push ad to **Group A**, not Group B

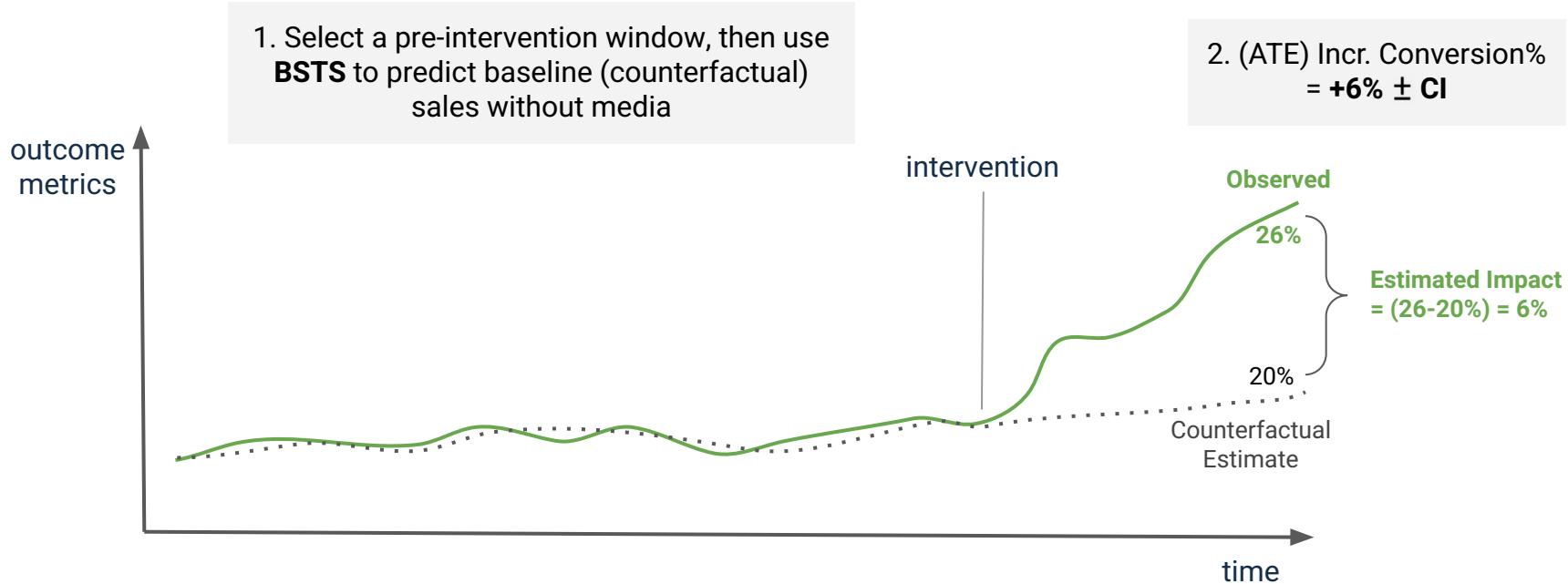
3. (ATE) Incr. Conversion%
= **+3% ± CI**



ATE: Average Treatment Effect for the overall population

Causal Impact & BSTS

Mostly applicable to offline or hard-to-track channels (e.g., **Linear TV, Influencers**), effective at detecting large, significant changes



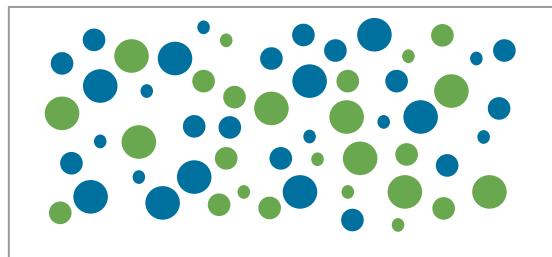
ATE: Average Treatment Effect for the overall population

© 2026 Selina (Xian) Xiao

Bias Correction (PSM + IPW)

Correcting for non-random treatment assignment in observational data: match each treated user to K similar controls by propensity score, and weight each record to represent the overall population

1. User-level with ad exposure data across channels



● Ad Exposure
● No Ad Exposure

2. (ATT) PSM to select matching **Group A** and Group B



● Group A: Ad Exposure
● Group B: No Ad Exposure

3. (ATE) IPW to quantify
Incr. Conversion%
= **+3% ± CI**



Incr.
Conv%

26%

23%

Ad Exposure Effect
= $(26\% - 23\%) = 3\%$