2026 AGIFORS RM STUDY GROUP

A scalable blueprint for airline RM experiments: switchback design for seat pricing tests

Hermon Kweon, PhD (ADC)

Nicolas Ferrari Ortiz (LATAM)

From pricing decisions to causal evidence

Revenue management teams change prices often. Randomized experiments help measure which changes actually move revenue.



A uniform 10% price increase



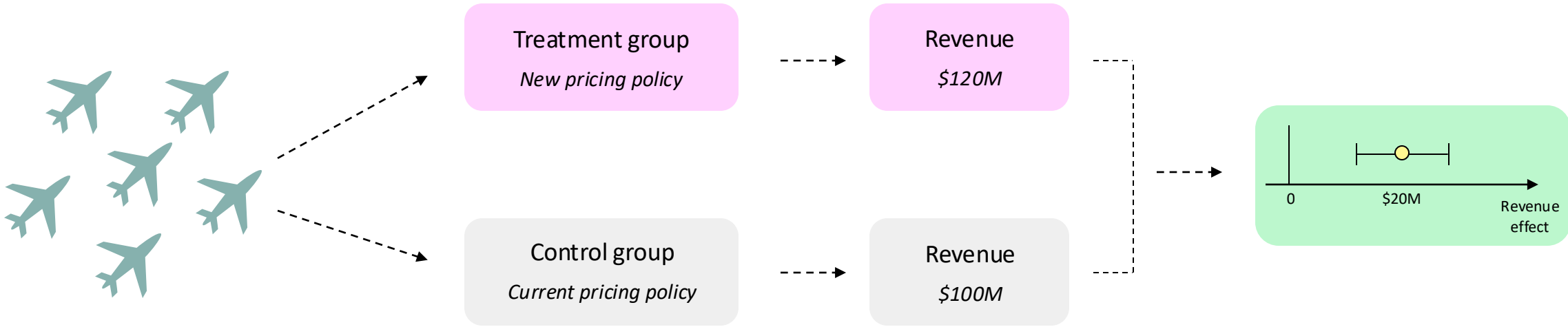
A new pricing optimizer



A larger price gap between fare families

Pricing policies are randomized by route or flight, not by passenger

Typically, a pricing policy must apply consistently within **a flight or route for a defined period**. All bookings in that unit share the same treatment assignment.



Randomization → Execution & Data Collection → Statistical Analysis

Higher precision reduces the time needed to reach the target signal-to-noise ratio

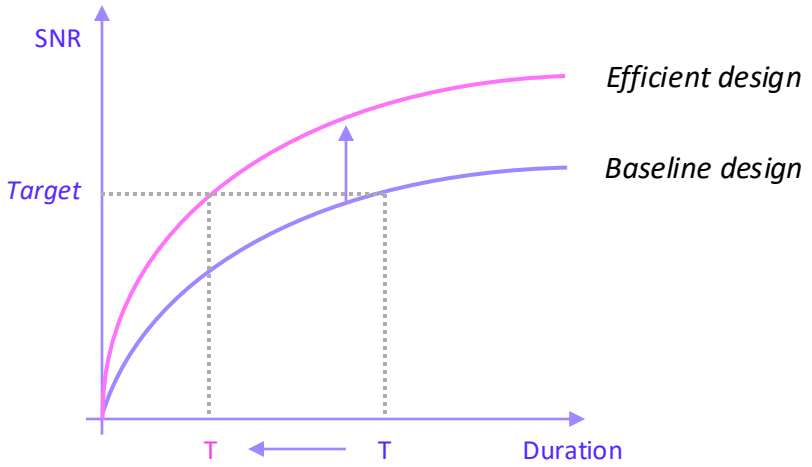
An observed difference is not enough. Statistical inference assesses the estimated effect relative to its uncertainty.

$$\text{Observed Signal to Noise Ratio (SNR)} = \frac{\text{Estimated Effect}}{\textbf{Standard Error (SE)}}$$

Lower SE = Higher precision ⇒ Higher SNR

When planning an experiment, choose a duration (T) so that the effect of interest is expected to reach the target SNR. The required duration depends on the expected SE. This is where precision matters.

$$\text{SNR at } T = \frac{\text{Effect size to detect}}{\text{Expected SE at } T}$$



Lower SE at the same duration means the target SNR is reached sooner.

The shorter the duration, the faster the decision

How can a smarter design reduce the runtime of route-level experiments?

Problem

Treatment assignment at flight or route level

⇒ Fewer independent comparisons

⇒ Lower precision (higher SE)

⇒ Longer runtime required

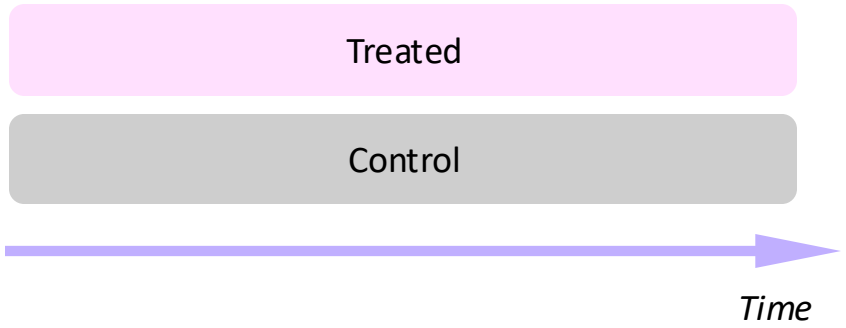
Solution

Switchback: Switch treatment assignment over time

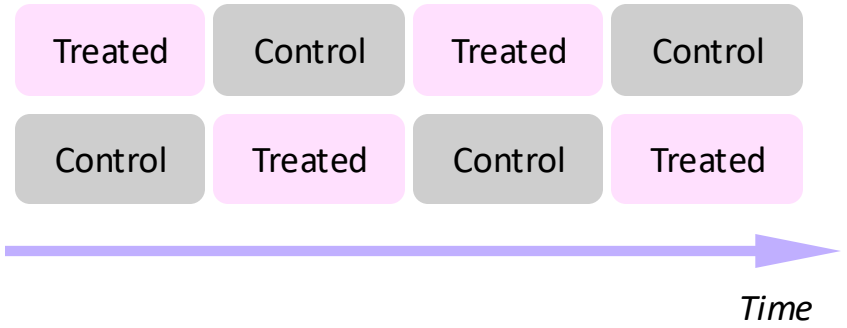
⇒ Repeated comparisons within the same unit

⇒ Higher precision (lower SE)

⇒ Shorter runtime required



Fixed assignment



Switchback assignment

Ancillary pricing lets us isolate the precision question

This talk studies how different switchback designs improve precision relative to fixed assignment and shorten the duration of ancillary pricing experiments.

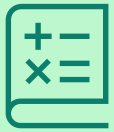
Why ancillary pricing?

For seat ancillaries, it is plausible that the pricing policy affects purchases mainly in the current booking period.

What this lets us study

Focus on the precision gains from different switchback designs, without the added trade-off from effects that persist after treatment switches.

Overview: three ways to evaluate switchback designs



Stylized model

- Why should switchbacks improve precision?
- Derive predictions about when and why precision gains occur



Simulation

- Do the theoretical predictions hold under controlled conditions?
- Vary the demand structure and compare the precision of alternative designs



Placebo tests with real data

- Do the predicted gains persist in real airline data?
- Apply hypothetical assignments to historical outcomes and compare standard errors

Structure of the data and experiment setup (fixed design)

Each group-period outcome combines many bookings that share the same treatment assignment.

PNR	OD	Booking time	Revenue	Group
…	SEALAX	Sep 3	20$	A
…	SEAYVR	Sep 5	7$	B
…	SFOSEA	Sep 15	14$	B
…	JFKSEA	Sep 22	22$	A
…	SEAPDX	Sep 30	10$	A

Aggregate by assignment unit and period

Experiment period (booking period)

		Sep W1	Sep W2	Sep W3	Sep W4
Group A	OD1	300$	305$	..	..
	OD2	285$	..	..	..
	OD3	..	..	..	..
Group B	OD4	305$	300$	..	..
	OD5	280$	..	..	..
	OD6	..	..	..	..

Treated

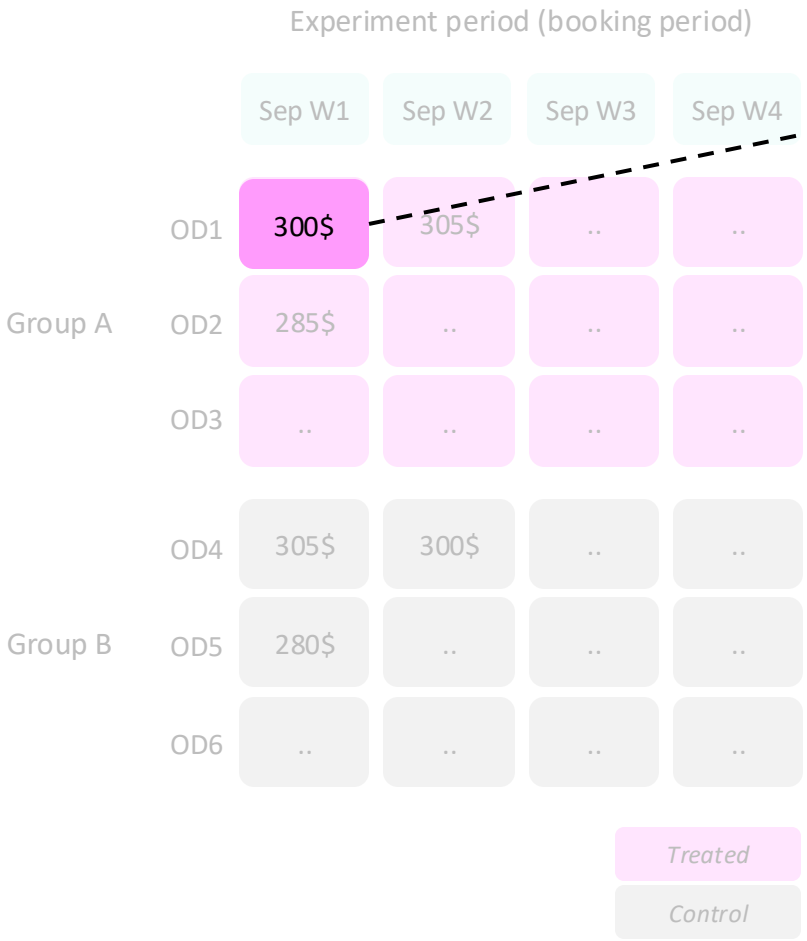
Control

Booking-level transaction data

Analysis dataset

Where does variation in OD-period revenue come from?

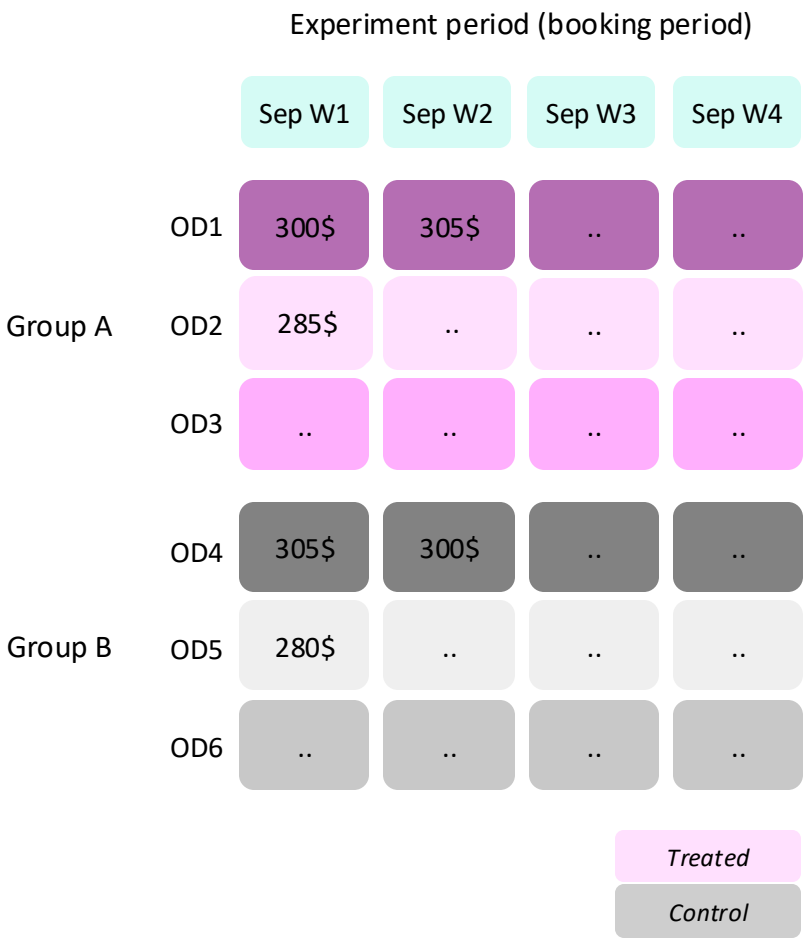
A stylized model separates the sources of variation in each OD and period.



OD-period revenue =

- Typical revenue for the OD
- Common period conditions
- OD-specific time-varying conditions
- Treatment effect
- Remaining variation

OD-period revenue varies because ODs have different baselines



OD-period revenue =

- *Distance*
- *Domestic / International*
- *Hub / non-hub market*
- *Typical passenger mix*

Typical revenue for the OD

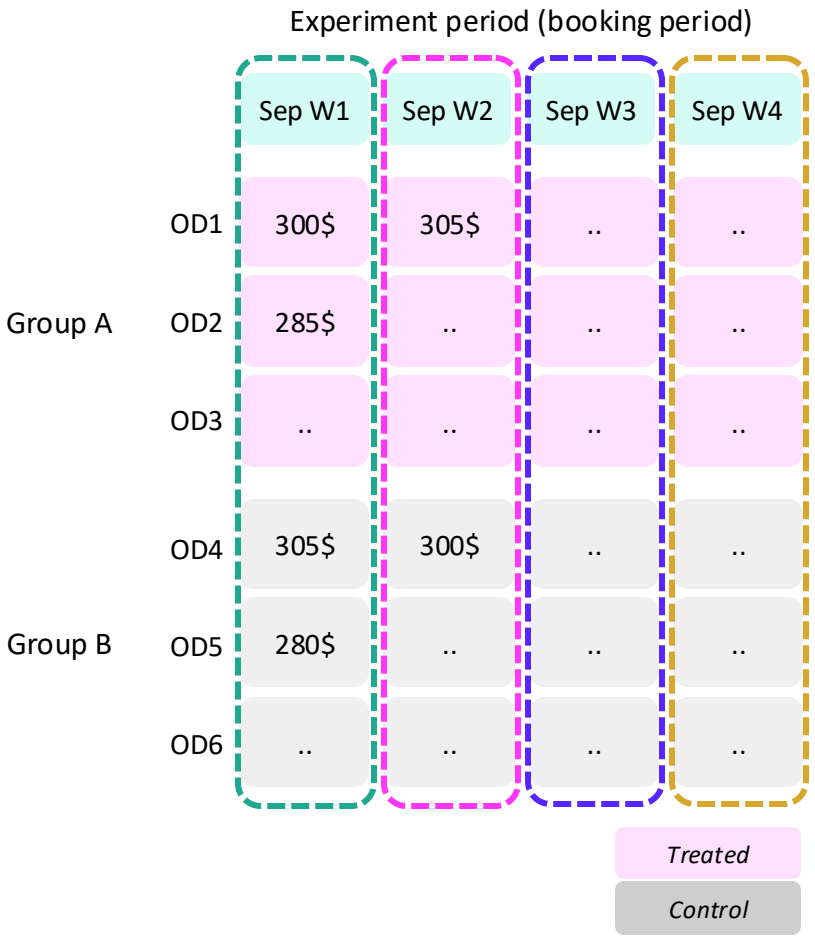
Common period conditions

OD-specific time-varying conditions

Treatment effect

Remaining variation

OD-period revenue varies because periods face common conditions



OD-period revenue =

- *Common seasonal patterns*
- *Day-of-week effects*
- *Network-wide promotions*

Typical revenue for the OD

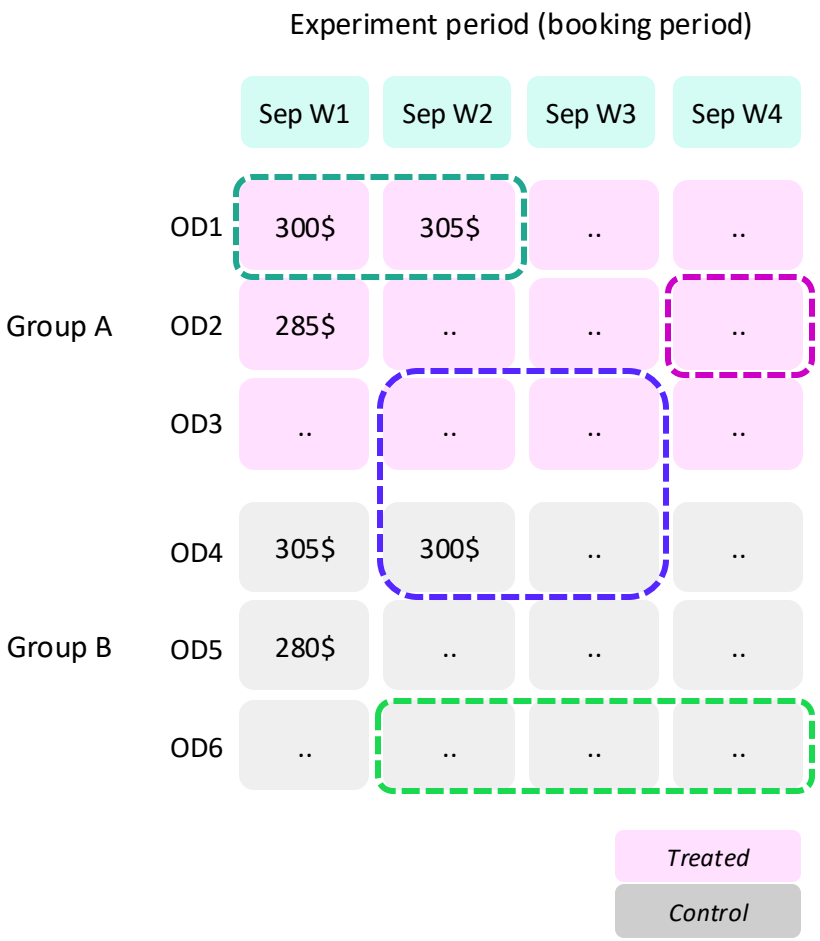
Common period conditions

OD-specific time-varying conditions

Treatment effect

Remaining variation

OD-period revenue varies because OD conditions change over time



OD-period revenue =

- *Local events*
- *Holiday demand that varies by OD*
- *Temporal shifts in passenger mix*

Typical revenue for the OD

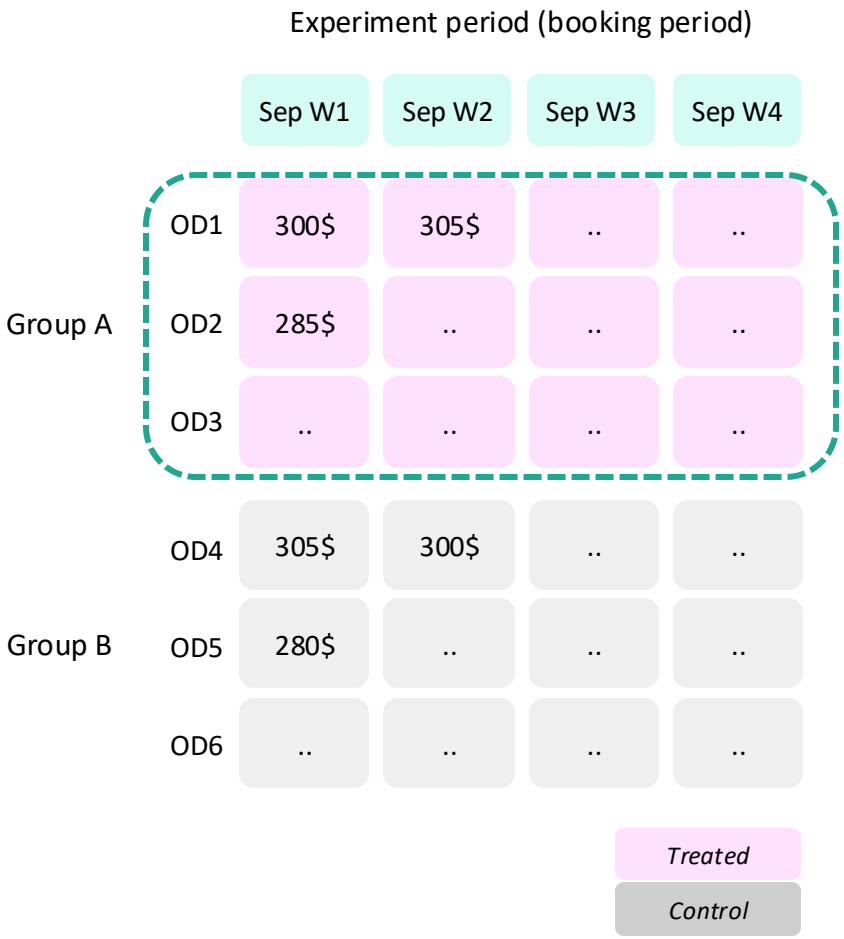
Common period conditions

OD-specific time-varying conditions

Treatment effect

Remaining variation

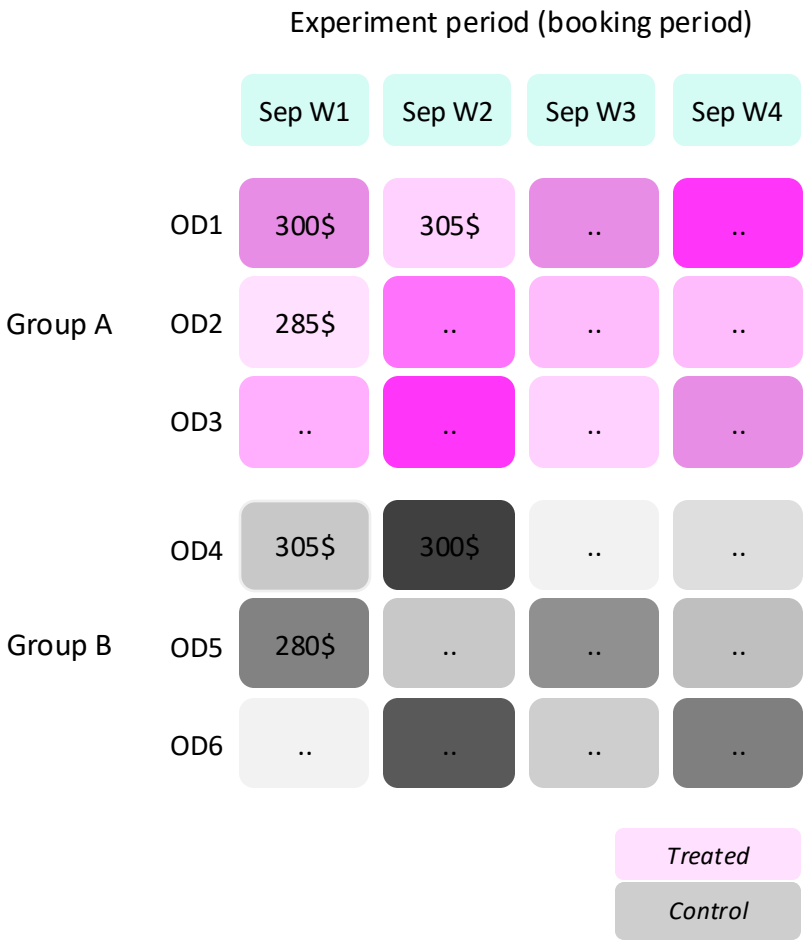
OD-period revenue varies because treatment changes revenue



OD-period revenue =

- Typical revenue for the OD
- Common period conditions
- OD-specific time-varying conditions
- **Treatment effect**
- Remaining variation

OD-period revenue also contains unexplained variation

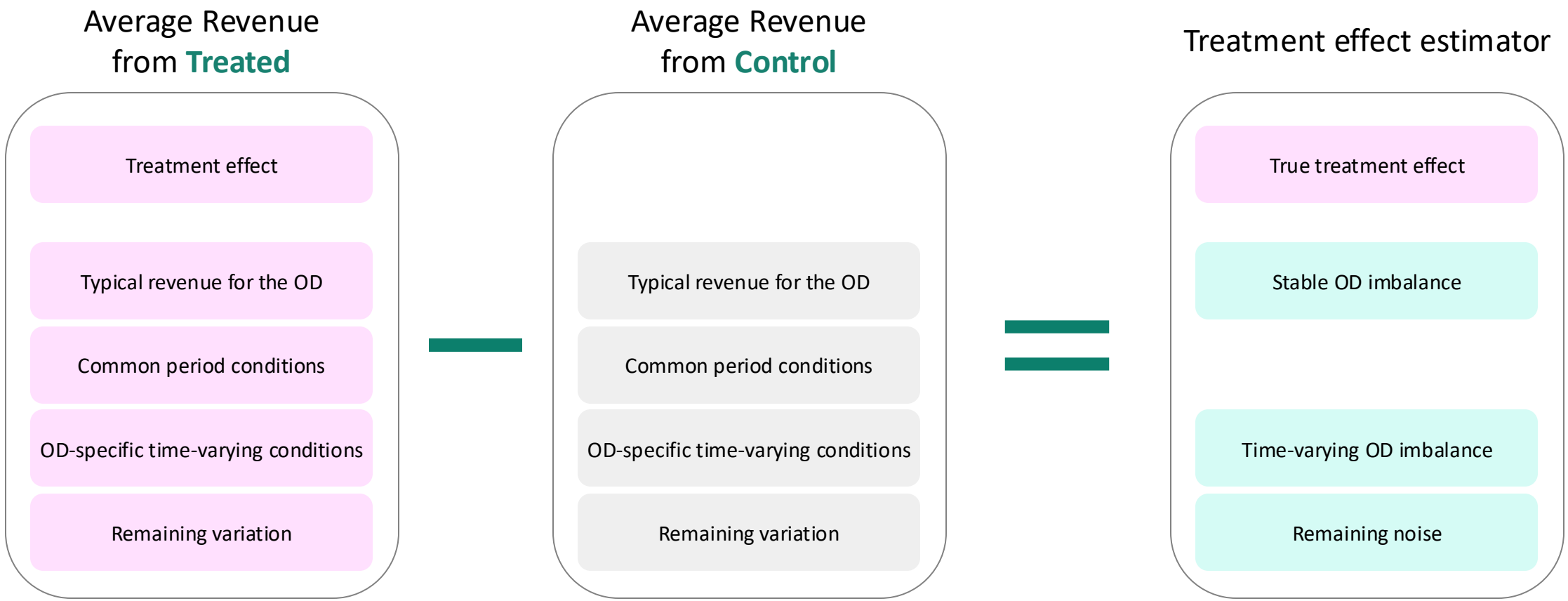


OD-period revenue =

- Typical revenue for the OD
- Common period conditions
- OD-specific time-varying conditions
- Treatment effect
- Remaining variation

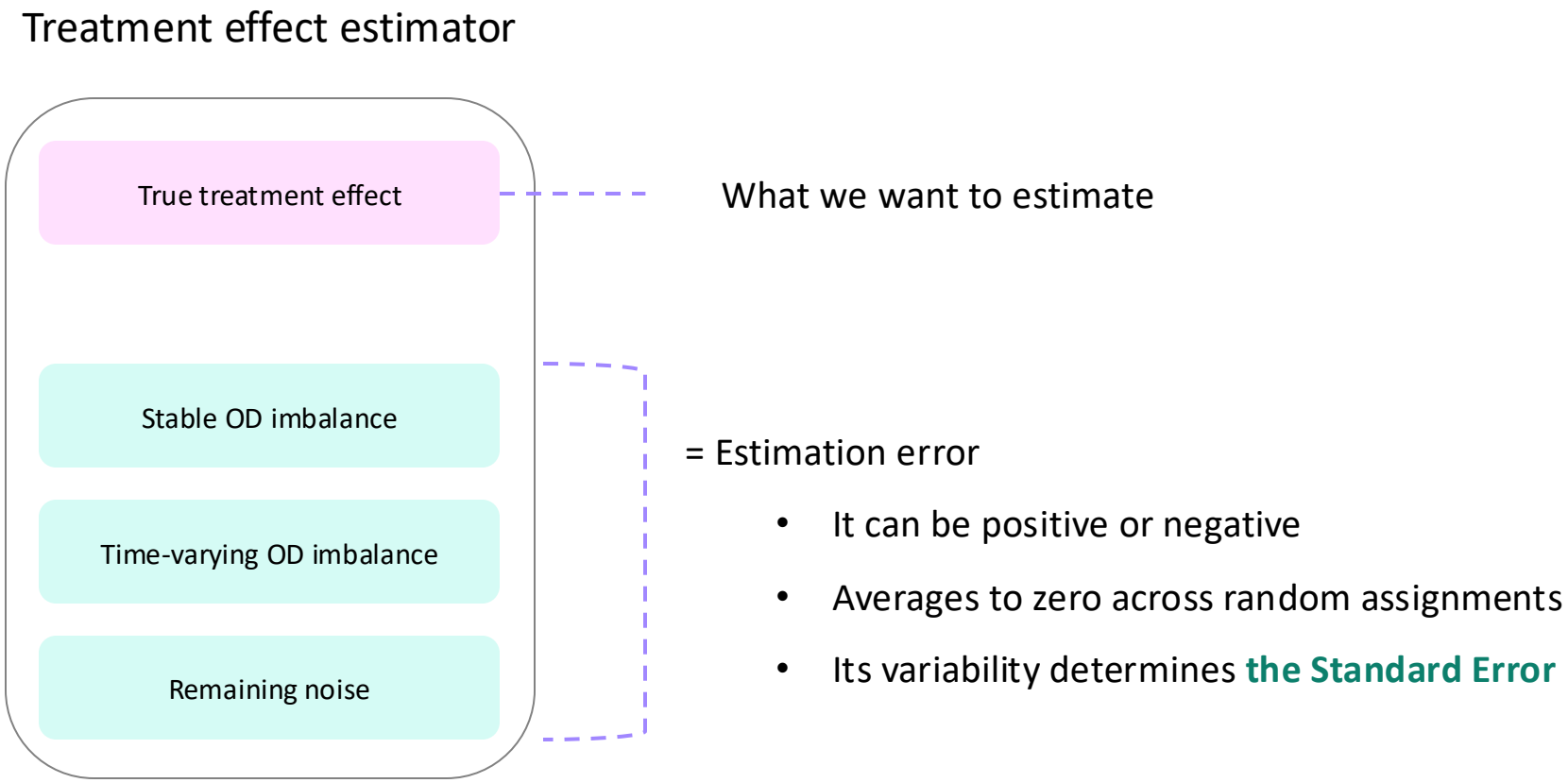
The treatment effect estimate is treated revenue minus control revenue

Common period conditions cancel, but a realized assignment can leave differences between the two groups.



The estimate equals the true effect plus estimation error

The error averages to zero across random assignments, but its variability determines the standard error.



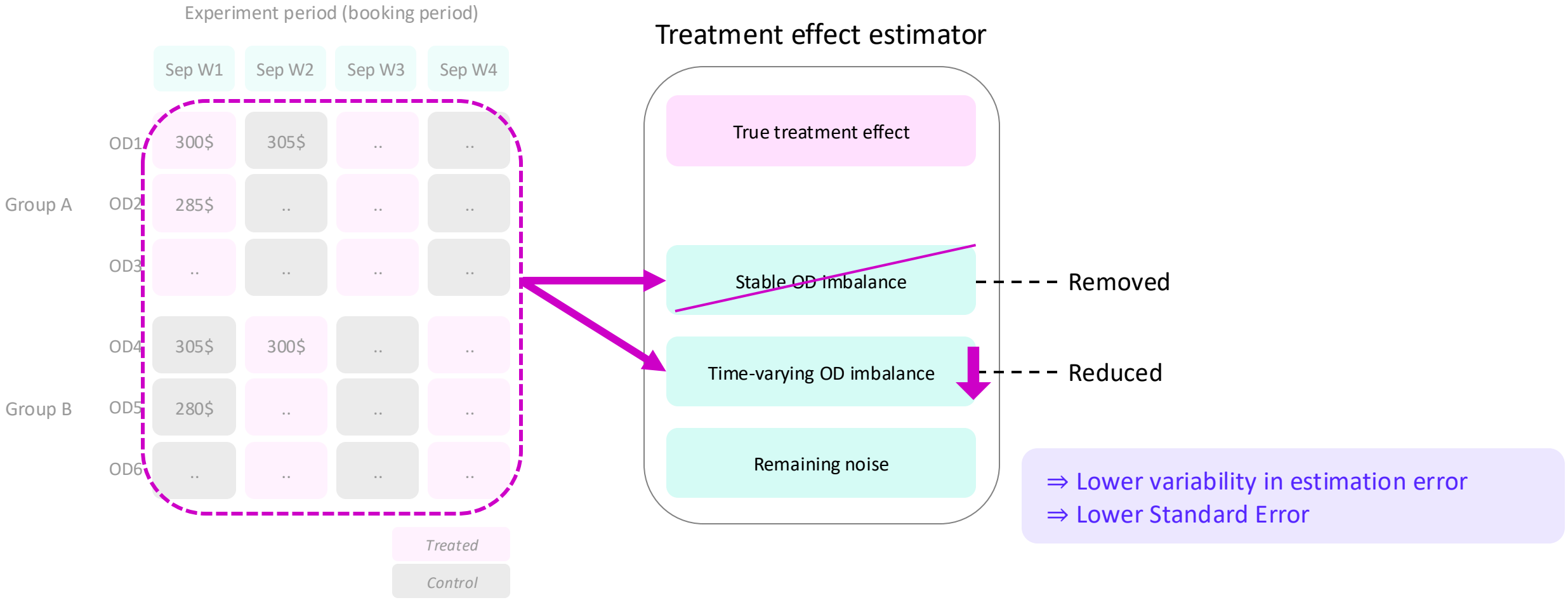
Switchbacks alternate treatment over time

Each OD is observed under both policies, creating repeated within-OD treatment contrasts.



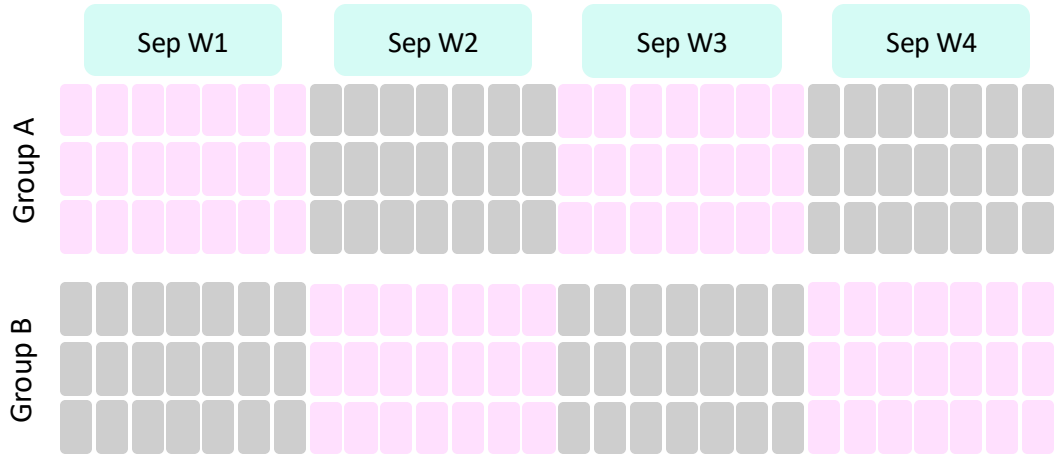
Switchbacks create repeated treated minus control contrasts within each OD

Stable OD differences cancel in the subtraction, while nearby comparisons reduce persistent time-varying imbalance.

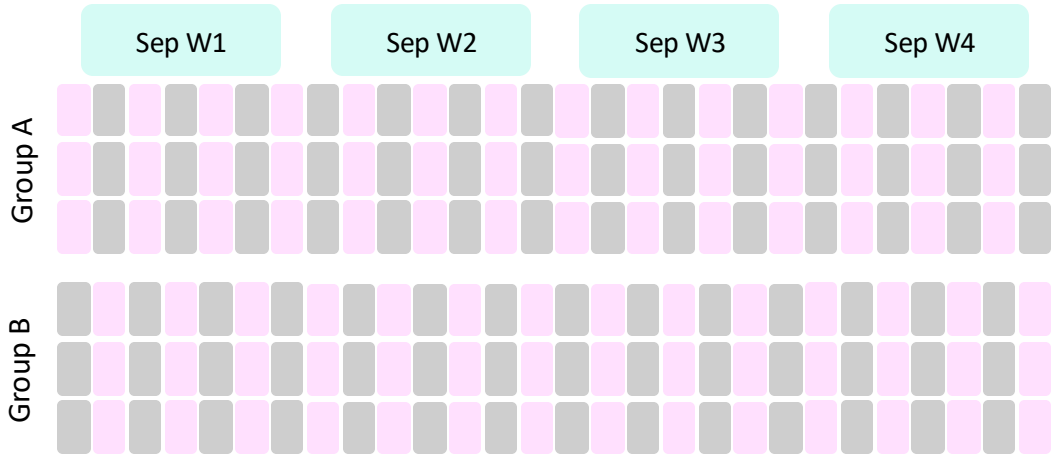


Switching frequency determines how local the comparison is

Daily switching creates closer comparisons. When nearby periods share similar OD conditions, more time-varying imbalance cancels and the SE falls.



Weekly switching



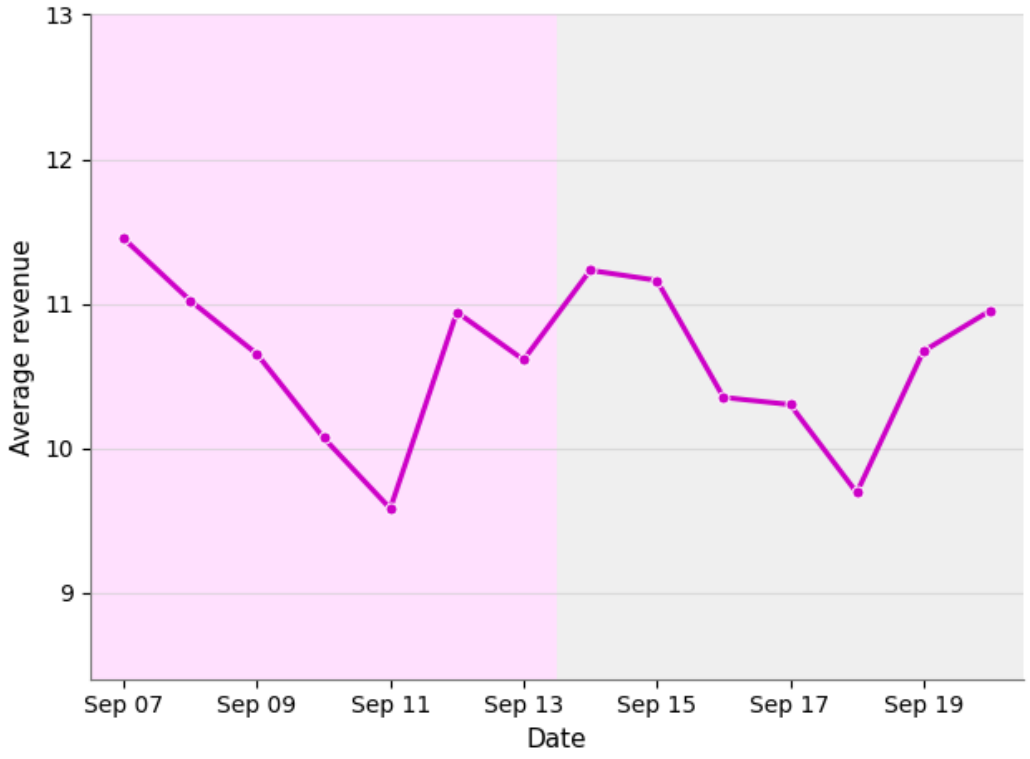
Daily switching

Treated

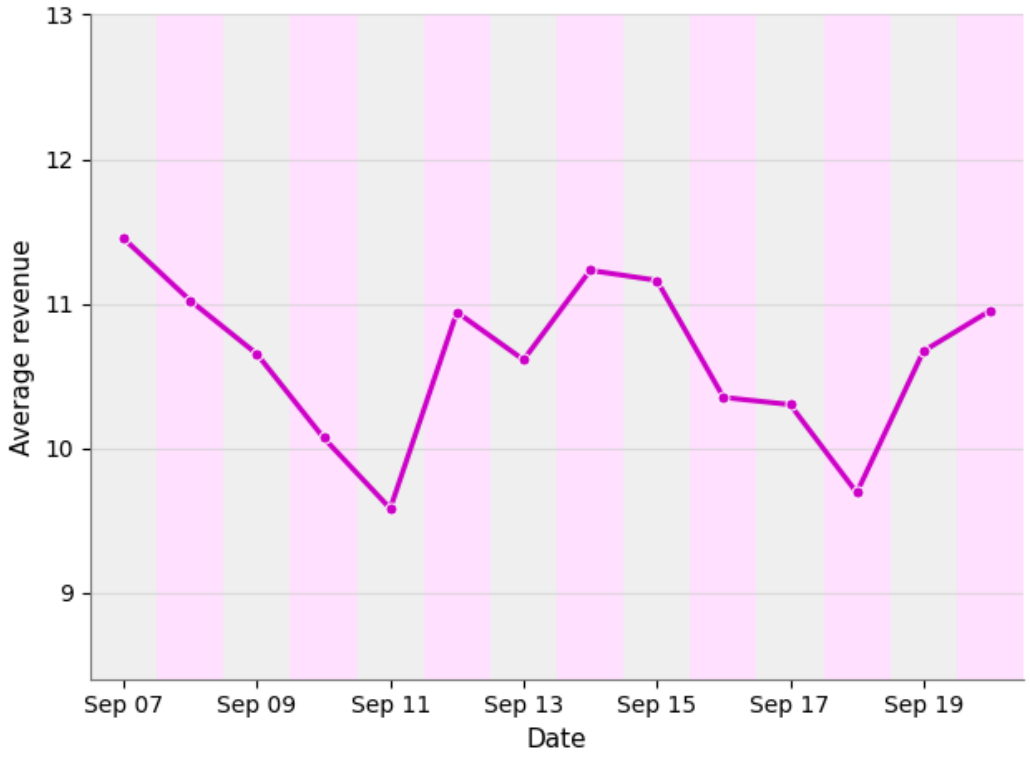
Control

When OD conditions repeat weekly, weekly switching captures most of the gain

Treatment and control observations seven days apart face similar OD conditions, so daily switching should add little.



Weekly switching



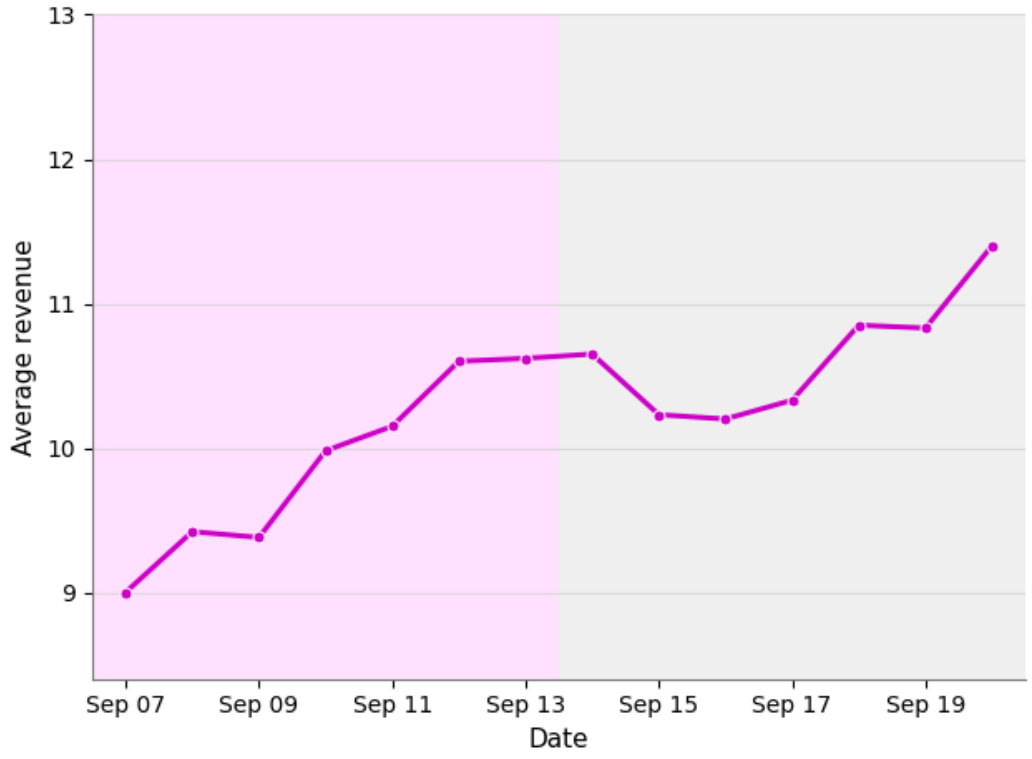
Daily switching

Treated

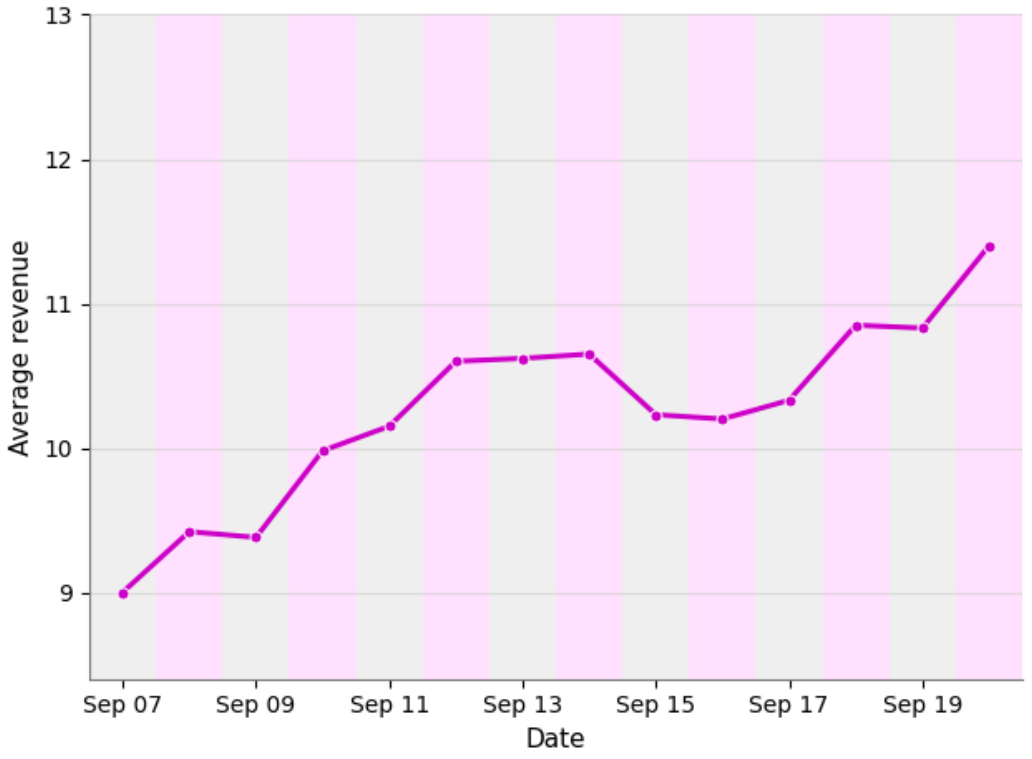
Control

Daily switching adds more when nearby OD conditions are more alike

Adjacent treated and control days face similar conditions, while observations farther apart do not. Daily switching therefore creates a cleaner comparison and improves precision relative to weekly switching.



Weekly switching



Daily switching

Treated

Control

The theoretical intuition leads to testable predictions

The predictions describe when switching improves precision and when switching more often improves it further.

Prediction	Test	Expected pattern
Switching should improve precision	Compare fixed and switchback assignment	Switchbacks reduce SE
OD-specific day-to-day persistence should make daily switching more useful	Increase OD-specific day-to-day persistence	Daily gains more over weekly

Simulations isolate the mechanisms behind the predictions

We vary one demand feature at a time, apply the same assignment designs, and compare precision.

Calibrated synthetic data

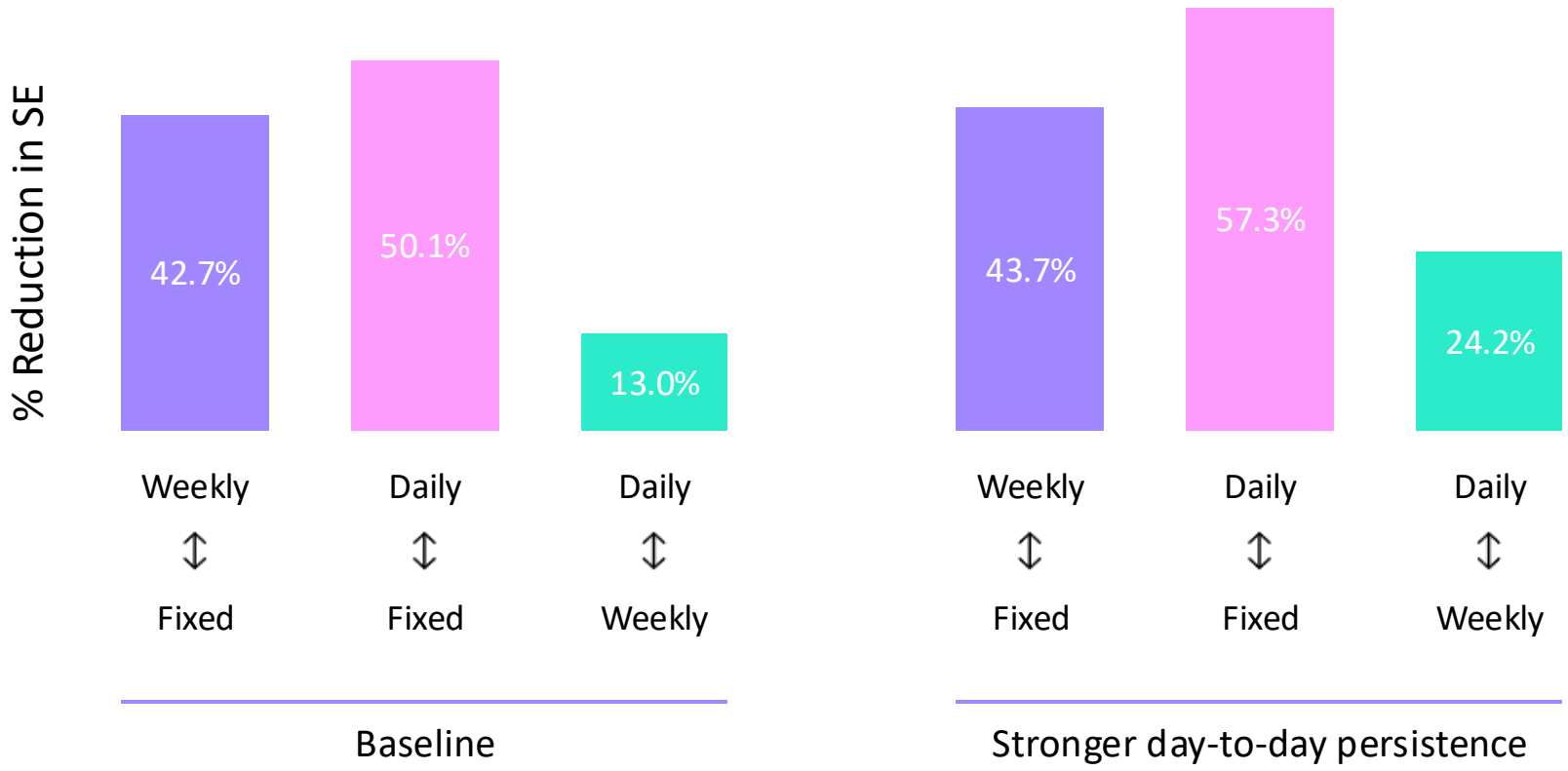
- Simulate daily revenue for each OD
- Reproduce key demand patterns:
  - Stable OD differences
  - Day-of-week patterns
  - Common demand shocks
  - Long horizon seasonality
  - OD-specific day-to-day persistence

Analysis

- 2 Scenarios
  - Baseline
  - Stronger day-to-day persistence
- Apply 3 assignment designs
  - Fixed, Weekly / Daily switching
- Repeat 200 Monte-Carlo runs and obtain mean SE

Simulation results match the predicted precision patterns

Switchbacks reduce SE in both scenarios. Stronger day-to-day persistence increases the advantage of daily switching.



Historical placebo tests bring the predictions to real LATAM Airlines data

The placebo tests show how fixed, weekly, and daily assignment perform under observed demand patterns.

Data

- 80 International routes of LATAM in 2024
- Economy brand
- Outcome: Seat selection revenue

Placebo test

- Treatment – Control assignment without real treatment effect
- 1,000 iterations

Prediction

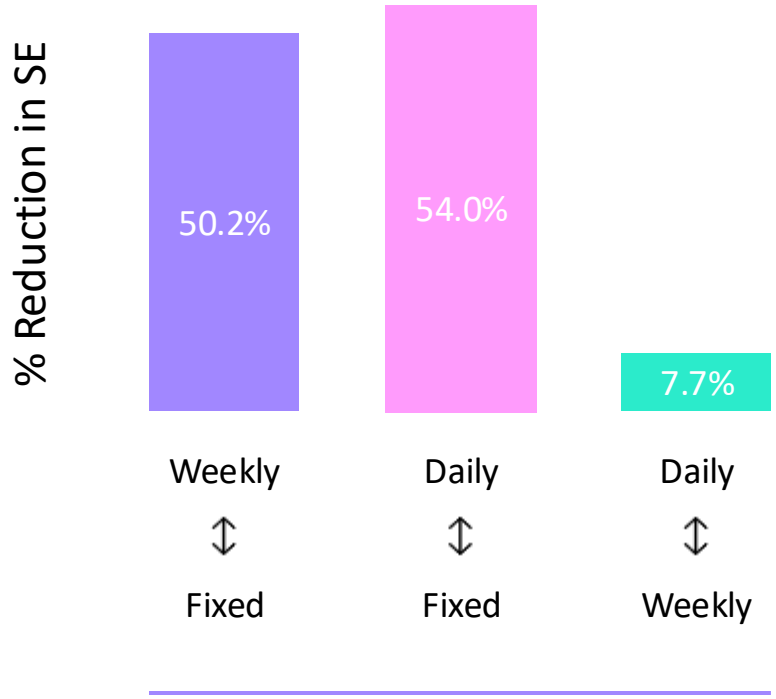
- Weekly persistence is observed to be a dominant factor

⇒ Switchback reduces SE substantially

⇒ Daily v.s. Weekly switching has similar effects on SE

Historical data confirm large switchback gains and little benefit from daily switching

At 8 weeks, weekly and daily switching reduce SE by about 50% relative to fixed assignment and perform nearly identically.



Precision gains translate into faster decisions and more learning

- 50% lower SE means roughly 75% less required runtime
- An experiment that would normally need four months could reach the same precision in roughly one month
- Shorter experiments speed up decisions and compound more learning

For further details:

Improving Precision in Cluster-Level Pricing Experiments with Temporal Switchbacks

Ferrari-Ortiz et al. 2026,

Forthcoming in *the Proceedings of the 35th ACM International Conference on Information and Knowledge Management (CIKM 2026).*