Conservative Inference in Switchback Experiments

Jose Blanchet **Peter W. Glynn** **Ramesh Johari** **Linjia Wu** **Wenqian Xing**
Stanford University

Abstract

Switchback experiments are widely used in dynamic systems—such as ridesharing platforms and online marketplaces—to evaluate interventions under interference. However, the standard pipeline of estimating the average treatment effect (ATE) with a difference-in-means (DM) estimator can exhibit systematic bias in dynamic settings with evolving system state, due to intertemporal dependence ("carryover effects"). In this paper, we study this bias in a continuous-time Markov chain model of switchback experiments with stochastically monotone dynamics and state-monotone rewards; these are reasonable representations of mean-reverting and auto-regressive systems. We show the DM estimator systematically *underestimates* the true ATE, because it targets an average of transient treatment effects rather than the ATE itself. Using the Ornstein-Uhlenbeck process as a tractable example, we derive closed-form expressions for bias and variance; this analysis shows that standard approaches overestimate the true variance. Taken together, these effects mean that standard switchback experiment analysis yields *overly conservative* inference. We validate our theory using a ride-sharing simulation with real-world calibration.

1 INTRODUCTION

A/B testing is widely used to evaluate policy differences (or *treatment effects*), such as pricing on ride-sharing platforms or inventory management in supply chains. However, estimates from standard randomized experiments are often biased due to *interference*: either across neighbors (network spillovers) or across time (carryover effects). In ride-sharing systems, e.g., pricing policies in one period affect driver availability in subsequent periods, creating temporal spillovers.

To mitigate interference, platforms often adopt *switchback experiments*, i.e., alternating treatment and control across time/regions and estimating interval-level effects. A typical industrial approach uses a difference-in-means (DM) estimator between switchback intervals and reports Huber–White standard errors (Kastelman and Ramesh, 2018). However, when outcomes evolve with system (e.g., pricing efficacy varying with driver availability), this pipeline not only exhibits estimation bias (e.g., Cooprider and Nassiri (2023); Nassiri (2025) note treatment can be overestimated), but when combined with the variance estimator can exhibit overly *conservative inference*. Further, the role of design features (e.g., interval length, washout) and their qualitative impact on the DM estimation pipeline is not well understood. In this paper, we investigate these phenomena by modeling the underlying system *stochastically monotone* dynamics; this is a broad class of dynamics, including mean-reverting autoregressive processes. We show these dynamics lead to conservative inference in switchback experiments, and also use the structure to study the role of experimental design.

An Industrial Approach. We summarize a standard pipeline for switchback experiments, e.g., as presented by DoorDash (Kastelman and Ramesh, 2018). Policies are randomized at the *region*×*time* level (Figure 1): each 30-minute interval in each region is independently assigned to treatment or control. Treatment effects are then estimated by comparing average outcomes (i.e., delivery time) between treated and control region-time units. Specifically, for region-time unit i, let $y_{\kappa,i}$ be the average outcome under policy $\kappa \in \{0,1\}$ (treatment = 1, control = 0). With N_1 treated and N_0 control units, the standard analysis fits $y = X\beta + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma^2 I)$ is a Gaussian error and

$$y = \left(y_{1,1}, \ldots, y_{1,N_1}, y_{0,1}, \ldots, y_{0,N_0}\right)^\top,$$

$$X = \left(x_1, \ldots, x_{N_1+N_0}\right), \quad x_i = \begin{cases} (1,1), & i \le N_1, \\ (1,0), & i > N_1. \end{cases}$$

Proceedings of the 29th International Conference on Artificial Intelligence and Statistics (AISTATS) 2026, Tangier, Morocco. PMLR: Volume 300.

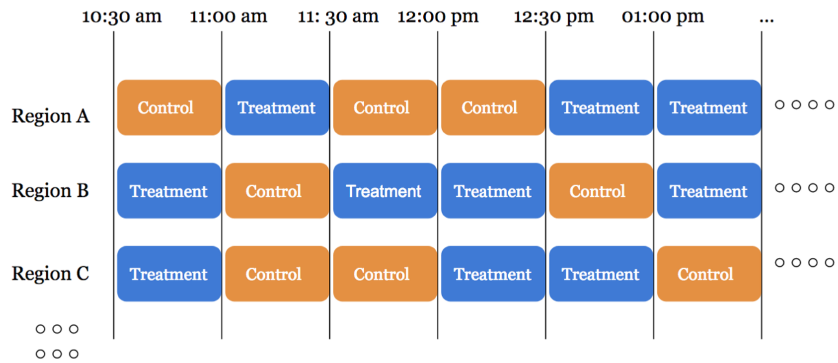

Figure 1: Switchback experiments for pricing at DoorDash (Kastelman and Ramesh, 2018).

OLS yields $\hat{\beta} = (X^\top X)^{-1} X^\top y$, where $\hat{\beta}_0$ is the control mean and $\hat{\beta}_1$ equals the standard difference-in-means (DM) estimator for the ATE β_1 (Wager, 2024). Variance is typically estimated via the Huber–White sandwich estimator (White, 1980) (with $n = N_1 + N_0$):

$$\widehat{\mathrm{Var}}(\hat{\beta}) = \frac{1}{n}\left(\frac{1}{n}X^\top X\right)^{-1}\frac{1}{n}X^\top \Omega X\left(\frac{1}{n}X^\top X\right)^{-1},$$

where Ω is a diagonal matrix with entries $(y - X\hat{\beta})^2$. Similar switchback experiments have been reported by Lyft (Chamandy, 2016), Tubi (Silbert, 2022), and Bolt (O'Connell and Bentes, 2022).

Main Contributions. In this paper, we tackle the following question: *when and why does the standard switchback experiment produce overly conservative inference?* Our analysis provides several meaningful insights on this question.

- **Markovian Framework.** In a continuous Markov chain (CTMC), we show that stochastic monotonicity, stochastically ordered policy kernels, and state-monotone rewards induce a *monotone bias* during transient periods—conditions often satisfied in real-world applications such as ride-sharing and online marketplaces.
- **Estimand Discrepancy.** We show that transient dynamics cause the DM estimator to target an average *transient treatment effect* rather than the ATE. We show that under stochastic monotonicity and state-monotone rewards, the difference-in-means estimator exhibits systematic underestimation, resulting in conservative inference.
- **A Stylized Model.** In an Ornstein–Uhlenbeck (OU) process setting, we verify the preceding monotonicity conditions and establish asymptotic underestimation bias, even with washout periods. We further characterize both the bias and the variance and show that the Huber–White variance estimator can be overly conservative. These findings are illustrated through a ride-sharing simulation using real-world data (TLC, 2024).
- **Practical Guidance.** Our results explain the conservative bias observed in practice and highlight design trade-offs (interval length, washout) that mitigate bias and reduce the DM estimator MSE.

Related Work. We review three areas: (i) switchback experiments, (ii) Markovian models for dynamic treatment effects, and (iii) industrial practice.

Switchback experiments, originating in dairy studies (Brandt, 1938), are now common for marketplace-level A/B tests (Chamandy, 2016). Related work addresses carryover via design and estimation—optimal designs with finite-horizon carryover (Bojinov et al., 2023; Ni and Bojinov, 2025), Bernoulli switchbacks in nonstationary MDPs with geometric mixing (Hu and Wager, 2022), and bias–variance trade-offs for multiple simultaneous switchbacks (Xiong et al., 2023). Recent work also develops robust switchback methodologies for large-scale industrial settings (Missault et al., 2025). In contrast, we focus on the *post-experiment* estimand–estimator gap—its direction and decision impact—under structural interference that induces systematic bias.

Markovian frameworks have been used to model treatment/control dynamics in experiments (Glynn et al., 2020; Hu and Wager, 2022; Farias et al., 2022, 2023; Li et al., 2023). Most related are policy-specific stationary chains for optimal design (Glynn et al., 2020) and a discrete-time MDP with an estimator balancing bias and variance (Farias et al., 2022). We extend this line by modeling with a continuous-state CTMC, broadening applicability to continuous-time experimental processes.

Large platforms now deploy switchbacks at scale: DoorDash provides implementation guidance (Kastelman and Ramesh, 2018); Lyft pairs switchbacks with simulation to study bias-variance and two-stage designs

(Chamandy, 2016; Quan, 2021); Tubi addresses seasonality, carryover, and power (Silbert, 2022); and Bolt discusses unit interference, temporal imbalance, and concurrent tests (O'Connell and Bentes, 2022). We aim to complement these operational deployments with theoretical insight that characterizes the validity of inference from switchback experiments.

2 SWITCHBACK EXPERIMENTS

In this section, we introduce experiments in dynamic systems modeled as continuous-state CTMCs. Section 2.1 introduces the CTMC model under a general experimental setting, presents the conditions that lead to structural interference and monotone bias, and defines the average treatment effect (ATE). Section 2.2 then describes the standard switchback design used in practice along with the DM estimator.

2.1 A Monotone Markov Chain Model

We consider two distinct CTMCs (treatment and control) on a common real-valued state space ($\mathbb{R}$). We consider dynamic experiments on *a single unit*, e.g., an entire ridesharing platform. The state captures the unit's current condition (e.g., driver availability). Control and treatment are modeled as actions taken in that state, and outcomes as rewards associated with it. Similar models have been studied in prior work for discrete-time, discrete-state settings (Glynn et al., 2020; Hu and Wager, 2022; Farias et al., 2022).

Notation. Let $f(x) = \mathcal{O}(g(x))$ if there exists $C > 0$, independent of x, with $|f(x)| \le C|g(x)|$ for all large x. For $N \in \mathbb{N}$, let $[N] = \{0, 1, \ldots, N-1\}$.

Time and State Space. We consider a continuous time space $t \in [0, \infty)$ and a real state space $\mathcal{X} \subseteq \mathbb{R}$. Throughout for simplicity we assume the Borel measure space on $\mathbb{R}$; we assume that $\mathcal{X}$ is Borel measurable.

Markov Chains. A stochastic process $X = \{X_t : t \ge 0\}$ is a continuous-time Markov chain (CTMC) if, for any time $s, t \ge 0$, any state $x \in \mathcal{X}$, and any measurable set $A \subseteq \mathcal{X}$, the following holds

$$\mathbb{P}(X_{t+s} \in A \mid \{X_u\}_{0 \le u \le t}) = \mathbb{P}(X_{t+s} \in A \mid X_t).$$

Transition Kernel. The dynamics of a CTMC for any $t \ge 0$ are characterized by the transition kernel

$$P(x, A; t) := \mathbb{P}(X_t \in A \mid X_0 = x), \forall x \in \mathcal{X}, A \subseteq \mathcal{X}.$$

We assume the standard Feller property (see, e.g., Ch 4.2 in Ethier and Kurtz (2009)) on $P(x, A; t)$ to ensure continuity for our theoretical analysis. In addition, the Feller property enforces time homogeneity of the CTMC.

Two Chains. We consider two CTMCs, indexed by $\kappa \in \{0, 1\}$, on the common state space $\mathcal{X}$, with transition kernels P_κ. Specifically, P_0 and P_1 correspond to the control and treatment policies. Both kernels are assumed to be *stochastically monotone*, and we assume the treatment kernel stochastically dominates the control kernel, denoted $P_0 \preceq_{\text{st}} P_1$, as defined below.

Assumption 1 (Stochastic monotonicity)**.** *The treatment and control chains, with transition kernels P_κ for $\kappa \in \{0, 1\}$, satisfy stochastic monotonicity, i.e., for any $t \ge 0$ and $s \in \mathcal{X}$, $P_\kappa(x, (s, \infty); t)$ is non-decreasing in the initial state $x \in \mathcal{X}$.*

Assumption 2 (Stochastic dominance)**.** *For any $t \ge 0$, the control kernel P_0 is stochastically dominated by the treatment kernel P_1, denoted $P_0 \preceq_{\text{st}} P_1$. That is, for every bounded measurable non-decreasing $f : \mathcal{X} \to \mathbb{R}$ and all $x \in \mathcal{X}$,*

$$\int_{\mathcal{X}} f(y)\, P_0(x, dy; t) \;\le\; \int_{\mathcal{X}} f(y)\, P_1(x, dy; t).$$

Remark 1. *Assumptions 1 and 2 are well motivated, e.g., stochastic monotonicity arises in queueing and inventory systems (Stoyan and Daley, 1983; Lund and Tweedie, 1996; Li et al., 2023), while stochastic dominance between treatment and control has been previously assumed in online marketplace pricing (Farias et al., 2023; Johari et al., 2024). These two conditions induce structural interference that creates a* monotone transient effect *in states, the key driver of systematic underestimation in switchback designs.*

We also impose a *mixing-time* condition on the transition kernels, akin to geometric ergodicity assumptions in experiments with Markovian dynamics (Farias et al., 2022; Hu and Wager, 2022).

Assumption 3 (Mixing time)**.** *For each $\kappa \in \{0, 1\}$, the CTMC is geometrically ergodic: there exist $C \ge 1$ and $\lambda_\kappa > 0$ such that for all $x \in \mathcal{X}$ and $t \ge 0$,*

$$\|P_\kappa(x, \cdot; t) - \pi_\kappa\|_{\text{TV}} \le Ce^{-\lambda_\kappa t},$$

where $\|\cdot\|_{\text{TV}}$ denotes the total variation distance.

Moreover, we denote $\lambda_{\min} = \min\{\lambda_0, \lambda_1\}$ as the minimum mixing rate between the two policies.

Rewards. Reward is accrued at rate $r_\kappa(x)$ in state x, i.e., the expected reward in $[t, t+dt]$ is $\int_t^{t+dt} r_\kappa(X_s) ds$.

Assumption 4 (Monotone rewards)**.** *The reward rate functions r_κ satisfy $|r_\kappa(x)| \le M$, are non-decreasing in x, and obey $r_0(x) \le r_1(x)$ for all $x \in \mathcal{X}$ and $\kappa \in \{0, 1\}$.*

Remark 2. *Assumption 4 is reasonable in many practical experiments where the treatment has a valid effect. Combined with Assumptions 1 and 2, it implies* *that the treatment policy yields higher expected immediate* and *future outcomes than the control policy. (In Johari et al. (2024) these are referred to as "positive" treatments; see also Holtz et al. (2020); Johari et al. (2022); Li et al. (2022); Dhaouadi et al. (2023); Bright et al. (2025).) We later show in the stylized OU model (Section 4) that the bounded reward assumption can be relaxed.*

Treatment Effect. We define the causal estimand τ, the average treatment effect (ATE), as the difference in stationary rewards between treatment and control:

$$\tau := \pi_1 r_1 - \pi_0 r_0, \text{ with } \pi_\kappa r_\kappa := \int_{\mathcal{X}} \pi_\kappa(dx) r_\kappa(x).$$

2.2 Switchback Experiments and Estimation

We adopt a standard switchback design alternating between two policies. Each policy is applied for n periods of length l, giving a total duration $T = 2nl$. A formal definition of the switchback process X is given below.

Definition 1 (Switchback process)**.** *The process of a switchback experiment $X = \{X_t\}_{0\le t\le T}$ evolves under a time-varying transition kernel*

$$Q = \begin{cases} P_1, & 2il \le t < (2i+1)l, \quad i \in [n], \\ P_0, & (2i+1)l \le t < 2(i+1)l, \quad i \in [n]. \end{cases}$$

The process X is also referred to as the *data-generating process* of the experiment. We assume the process is initialized from the stationary distribution of the control policy, $X_0 \sim \pi_0$, i.e., the system is in steady state under control prior to the experiment (though our results hold for any initial distribution).

Definition 2 (Difference-in-means)**.** *The difference-in-means estimator $\hat{\tau}_n$ for the switchback process X is*

$$\hat{\tau}_n = \frac{1}{n}\sum_{i=0}^{n-1}\left(I_i^{(1)} - I_i^{(0)}\right),$$

where $I_i^{(1)}$ and $I_i^{(0)}$ are the average rewards in the i-th treatment and control periods, i.e.,

$$I_i^{(j)} = \frac{1}{l}\int_{(2i-j+1)l}^{(2i-j+2)l} r_j(X_t)\, dt, \quad j \in \{0,1\}.$$

Interference and Bias. In switchback experiments, outcomes in the previous interval can impact outcomes in the next interval, through stateful intertemporal dependence; this is known as a "carryover" effect (Bojinov et al., 2023), and constitutes a form of intertemporal interference that violates the Stable Unit Treatment Value Assumption (Imbens and Rubin, 2015). As a consequence of this interference, naive estimators such as the difference-in-means will be biased.

For example, in the CTMC model, outcomes depend on evolving states, inducing temporal correlation. This creates carryover effects during *transient periods*, when the system has not reached steady state and past treatments affect future states. In the next section, we show that under Assumptions 1, 2, 3, and 4, the bias is systematically downward.

3 BIAS DECOMPOSITION

In this section, we decompose the bias of the DM estimator. Before comparing $\hat{\tau}_n$ with τ, we introduce a related estimand, the average *transient treatment effect* $\tilde{\tau}$, which accounts for transient dynamics within each switchback period. In Section 3.1, we formally define this estimand and decompose the bias, and in Section 3.2, we explain the source of underestimation. A general variance analysis for the CTMC setting is provided in Appendix C.

3.1 Transient Treatment Effect

We consider the following approximation of the expected reward rate in a single switchback period under both treatment and control policies, denoted by $\tilde{I}^{(\kappa)}$:

$$\tilde{I}^{(1)} = \pi_1 r_1 + \frac{1}{l}\pi_0 g_1, \quad \tilde{I}^{(0)} = \pi_0 r_0 + \frac{1}{l}\pi_1 g_0, \qquad (1)$$

where the function $g_\kappa : \mathcal{X} \to \mathbb{R}$ is defined as

$$g_\kappa(x) = \int_0^\infty \left(\mathbb{E}_\kappa[r_\kappa(X_t) \mid X_0 = x] - \pi_\kappa r_\kappa\right) dt. \qquad (2)$$

Here, $\mathbb{E}_\kappa$ denotes the expectation under the transition kernel P_κ. Therefore, $g_\kappa(x)$ represents the *total deviation* of the expected reward $r_\kappa(X_t)$ from its stationary average reward $\pi_\kappa r_\kappa$, accumulated over time from initial state $X_0 = x$. We then define the average *transient treatment effect* $\tilde{\tau}$ as

$$\tilde{\tau} := \tilde{I}^{(1)} - \tilde{I}^{(0)} = \pi_1 r_1 - \pi_0 r_0 + \frac{1}{l}(\pi_0 g_1 - \pi_1 g_0). \qquad (3)$$

Note that $\tilde{\tau}$ is an estimand defined for a given switchback interval length l. In contrast to the ATE τ, the average transient treatment effect $\tilde{\tau}$ accounts for transient dynamics within each switchback interval via a *deviation correction term* depending on g. This correction is evaluated under the stationary distribution of the alternative policy; e.g., $\pi_0 g_1$ captures the deviation when the system starts in control steady state but evolves under the treatment policy.

Under the geometric ergodicity condition in Assumption 3, the function g is well-defined and bounded. We summarize the result in the lemma below.

Lemma 1. *Under Assumptions 3 & 4, the function g_κ in (2) is well-defined and uniformly bounded for each $\kappa \in \{0,1\}$. In particular, for all $x \in \mathcal{X}$,*

$$|g_\kappa(x)| \leq \frac{2MC}{\lambda_\kappa}.$$

Remark 3. *The function g_κ is the unique solution to* Poisson's equation *for the CTMC with kernel P_κ:*

$$\mathcal{A}_\kappa g_\kappa = -\bar{r}_\kappa, \tag{4}$$

where $\mathcal{A}_\kappa$ is the infinitesimal generator,

$$\mathcal{A}_\kappa g_\kappa(x) = \lim_{t \to 0^+} \frac{\mathbb{E}_{x,\kappa}[g_\kappa(X_t)] - g_\kappa(x)}{t},$$

and $\bar{r}_\kappa(x) = r_\kappa(x) - \pi_\kappa r_\kappa$ is the centered reward.

To analyze the bias between the DM estimator $\hat{\tau}_n$ and the transient effect $\tilde{\tau}$, note that $\tilde{I}^{(\kappa)}$ is a mean-field approximation of the average reward in a single switchback period, i.e., it approximates $\mathbb{E}[I_i^{(\kappa)}]$ under P_κ for any $i \in [n]$ and $\kappa \in \{0,1\}$. The approximation error is bounded in the following lemma.

Lemma 2. *Under the conditions of Lemma 1, for any $i \in [n]$ and $\kappa \in \{0,1\}$, the following bounds hold*

$$\left|\mathbb{E}[I_i^{(\kappa)}] - \tilde{I}^{(\kappa)}\right| \leq \frac{4MC^2}{\lambda_\kappa l}\,(e^{-\lambda_0 l} + e^{-\lambda_1 l}).$$

Lemma 2 shows that the approximation error between $\tilde{I}^{(\kappa)}$ and $\mathbb{E}[I_i^{(\kappa)}]$ decays exponentially in the interval length l. Applying this result, we can further bound the bias between the DM estimator $\hat{\tau}_n$ and the transient effect $\tilde{\tau}$ in the following corollary.

Corollary 1. *Under Lemma 2, for any $n \geq 1$, the bias between the DM estimator $\hat{\tau}_n$ and the transient effect $\tilde{\tau}$, denoted by $B(\hat{\tau}_n, \tilde{\tau})$, is bounded by*

$$|B(\hat{\tau}_n, \tilde{\tau})| = |\mathbb{E}[\hat{\tau}_n] - \tilde{\tau}| \leq \frac{16MC^2}{l\lambda_{\min}} e^{-\lambda_{\min} l}.$$

Remark 4. *Corollary 1 shows that the bias between $\hat{\tau}_n$ and $\tilde{\tau}$ decays exponentially in the interval length l. Moreover, in switchback experiments, $\hat{\tau}_n$ aligns more closely with $\tilde{\tau}$ than with τ, since transient dynamics are intrinsic to the switchback process in Definition 1.*

Analyzing the gap between $\tilde{\tau}$ and τ reveals the source of bias between the DM estimator $\hat{\tau}_n$ and the ATE τ. We call this gap the *estimand discrepancy*, denoted δ_τ. Formally, we have

$$\delta_\tau := \tilde{\tau} - \tau = \frac{1}{l}(\pi_0 g_1 - \pi_1 g_0). \tag{5}$$

We note that the estimand discrepancy δ_τ decays at rate $\mathcal{O}(l^{-1})$. Applying Lemma 1, we further obtain:

$$|\delta_\tau| = \frac{1}{l}\left|\pi_0 g_1 - \pi_1 g_0\right| \leq \frac{2MC}{l}\Big(\frac{1}{\lambda_0} + \frac{1}{\lambda_1}\Big) \leq \frac{4MC}{l\lambda_{\min}}.$$

In contrast to the bias between $\hat{\tau}_n$ and $\tilde{\tau}$, which decays at rate $\mathcal{O}(l^{-1}e^{-\lambda_{\min} l})$ (Corollary 1), the estimand discrepancy δ_τ decreases more slowly at rate $\mathcal{O}(l^{-1})$. Hence, for large l, the bias between $\hat{\tau}_n$ and τ, denoted as $B(\hat{\tau}_n, \tau)$, is dominated by δ_τ. Using Lemma 1 and Corollary 1, we obtain the following bound.

Corollary 2. *Under Lemma 1, for any $n \geq 1$,*

$$\begin{aligned}|B(\hat{\tau}_n, \tau)| = \left|\mathbb{E}[\hat{\tau}_n] - \tau\right| &\leq \underbrace{\left|\mathbb{E}[\hat{\tau}_n] - \tilde{\tau}\right|}_{|B(\hat{\tau}_n,\tilde{\tau})|} + \underbrace{\left|\tilde{\tau} - \tau\right|}_{|\delta_\tau|} \\ &\leq \frac{16MC^2}{l\lambda_{\min}} e^{-\lambda_{\min} l} + \frac{4MC}{l\lambda_{\min}}.\end{aligned}$$

Thus, the bias $B(\hat{\tau}_n, \tau)$ decomposes into two terms: $B(\hat{\tau}_n, \tilde{\tau})$, which decays at rate $\mathcal{O}(l^{-1}e^{-\lambda_{\min} l})$, and δ_τ, which decays at rate $\mathcal{O}(l^{-1})$. For large l, δ_τ determines the direction of the bias $B(\hat{\tau}_n, \tau)$.

3.2 Systematic Underestimation

We now analyze the sign of the bias term $B(\hat{\tau}_n, \tau)$ and show it can systematically underestimate the ATE. By Corollary 2, for large l, the bias is dominated by the estimand discrepancy δ_τ, which reflects the difference between $\pi_0 g_1$ and $\pi_1 g_0$. To characterize its sign, we establish monotonicity of g_κ and a stochastic ordering of the stationary distributions π_κ for $\kappa \in \{0,1\}$, following from Assumptions 1–2.

Theorem 1. *Let $\{X_t : t \geq 0\}$ be an ergodic, stochastically monotone CTMC on $\mathcal{X} \subseteq \mathbb{R}$ with stationary distribution π. Let $P_t f(x) := \mathbb{E}_x[f(X_t)]$, and let*

$$\bar{r}(x) := r(x) - \pi r, \qquad \pi r := \int_{\mathcal{X}} r(y)\,\pi(dy).$$

Suppose $r : \mathcal{X} \to \mathbb{R}$ is non-decreasing and $\forall x \in \mathcal{X}$,

$$\int_0^\infty P_t|\bar{r}|(x)\,dt = \int_0^\infty \mathbb{E}_x[\,|r(X_t) - \pi r|\,]\,dt < \infty,$$

and assume further that $P_t\bar{r}(x) \to \bar{r}(x)$ as $t \downarrow 0$ (e.g., this holds if X is Feller and $r \in C_b(\mathcal{X})$). Then

$$g(x) := \int_0^\infty P_t\bar{r}(x)\,dt = \int_0^\infty \big(\mathbb{E}_x[r(X_t)] - \pi r\big)\,dt$$

is well defined and non-decreasing, and g solves Poisson's equation $\mathcal{A}g(x) = -\bar{r}(x) = \pi r - r(x)$.

Remark 5. *Theorem 1 is a continuous-time analogue of Proposition 1 in Glynn and Infanger (2022).*

Under Assumptions 1 and 4, Theorem 1 applies to the Markov chains under both control (P_0) and treatment (P_1). Hence, g_0 and g_1 are non-decreasing and satisfy Poisson's equation (4). We next establish a stochastic ordering between the stationary distributions π_0 and π_1 in the following lemma.

Lemma 3. *Under Assumptions 2 & 3, the stationary distributions of P_0 and P_1 satisfy $\pi_0 \preceq_{\text{st}} \pi_1$. That is, for any bounded non-decreasing measurable $f : \mathcal{X} \to \mathbb{R}$,*

$$\pi_0 f \leq \pi_1 f.$$

Remark 6. *Under Assumptions 2, 3, and 4, Lemma 3 implies that $\pi_0 r_0 \leq \pi_1 r_0 \leq \pi_1 r_1$, i.e., $\tau \geq 0$.*

Finally, by combining Theorem 1 and Lemma 3, we arrive at the following main result.

Theorem 2. *Under the conditions of Theorem 1 and Lemma 3, the estimand discrepancy δ_τ defined in (5) is non-positive in switchback experiments, i.e.,*

$$\delta_\tau = \frac{1}{l}(\pi_0 g_1 - \pi_1 g_0) \leq 0. \tag{6}$$

Remark 7. *Theorem 2 reflects a monotone bias in switchback experiments, i.e., when the treatment has higher outcomes than the control, $\pi_0 g_1 \leq 0$ in treatment intervals, while $\pi_1 g_0 \geq 0$ in control intervals. We note that equality in (6) holds iff $\pi_0 g_1 = \pi_1 g_0 = 0$, i.e., when there is no transient effect in either the treatment or control intervals. As a result, treatment outcomes are underestimated, and control outcomes are overestimated, so the average transient effect $\tilde{\tau}$ is smaller than the ATE τ under the conditions of Theorem 2.*

3.3 Regime of Underestimation

We now characterize a threshold interval length $\bar{l}$ such that for all $l > \bar{l}$, the estimand discrepancy δ_τ dominates $B(\hat{\tau}_n, \tau)$, resulting in an underestimation bias, i.e., $B(\hat{\tau}_n, \tau) < 0$. Throughout this section, we assume nontrivial transient effects, i.e., $\pi_0 g_1 < 0$ and $\pi_1 g_0 > 0$, which implies $\delta_\tau < 0$. By Corollary 1 and Theorem 2, there exists $\bar{l} > 0$ such that for all $l > \bar{l}$,

$$\frac{8MC^2 e^{-\lambda_{\min} l}}{l\lambda_{\min}} < \left| \frac{\pi_0 g_1 - \pi_1 g_0}{l} \right| = \frac{\pi_1 g_0 - \pi_0 g_1}{l},$$

which indicates that the magnitude of the bias $B(\hat{\tau}_n, \tilde{\tau})$ is smaller than the estimand discrepancy δ_τ, i.e., δ_τ determines the sign of the overall bias $B(\hat{\tau}_n, \tau)$. Specifically, one may set

$$\bar{l} := \frac{1}{\lambda_{\min}} \ln \left(\frac{8MC^2}{\lambda_{\min}(\pi_1 g_0 - \pi_0 g_1)} \right). \tag{7}$$

Then for all $l > \bar{l}$, the DM estimator $\hat{\tau}_n$ exhibits a systematic underestimation bias.

Equation (7) shows that faster mixing yields a smaller threshold $\bar{l}$, while slower mixing yields a larger one. For $l < \bar{l}$, the bias depends on both $B(\hat{\tau}_n, \tilde{\tau})$ and δ_τ, requiring further specification of the CTMC model. To gain insight, we next analyze a stylized mean-reverting process, deriving asymptotic bias and variance of the DM estimator under switchback designs with *washout periods* (Senn, 2002), and examining their dependence on interval length and washout proportion.

4 A STYLIZED MODEL

In this section, we focus on a specific instance of the CTMC model: the Ornstein–Uhlenbeck (OU) process. In Section 4.1, we introduce the OU process and discuss its connection to the CTMC framework. In Section 4.2, we describe the estimation procedure under switchback experiments with washout periods. In Section 4.3, we analyze the asymptotic bias and variance of the DM estimator. In Section 4.4, we study the bias–variance trade-offs with respect to the washout proportion and the equilibrium convergence speed. Additional robustness analysis under nonstationarity is provided in Appendix B.

4.1 The OU Process as a Continuous-State CTMC under Monotonicity

For simplicity, we omit model settings that are identical to those described in Section 2.1.

Process Dynamics. Under each policy $\kappa \in \{0, 1\}$, the state follows an OU process with mean μ_κ and common mean-reversion rate θ and volatility σ:

$$dX_t = -\theta(X_t - \mu_\kappa)\,dt + \sigma\,dW_t, \tag{8}$$

where $\{W_t\}_{t\geq 0}$ is a standard Brownian motion. Assume X_0 is Gaussian with bounded moments. To simplify analysis, we express any OU process from (8) in terms of a *canonical OU process* $\{Z_t\}_{t\geq 0}$, i.e.,

$$dZ_t = -Z_t\,dt + dW_t. \tag{9}$$

Thus Z is an OU process with $\theta = 1$, $\mu = 0$, $\sigma = 1$, and $Z_0 = 0$. Its connection to a general OU process is given below.

Lemma 4. *Given Z_t as in (9), and for $\kappa \in \{0, 1\}$ and $t \geq 0$, define:*

$$X_t = (X_0 - \mu_\kappa)e^{-\theta t} + \mu_\kappa + \tfrac{\sigma}{\sqrt{\theta}} Z_{\theta t}. \tag{10}$$

Then X_t is an OU process with mean μ_κ, mean-reversion rate θ, and volatility σ, and obeys (8).

Transition Kernel. The transition kernel $P_\kappa(x, y; t)$ of the OU process is given by the Gaussian density

$$P_\kappa(x, y; t) := \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(y - m_\kappa(x; t))^2}{2\sigma_t^2} \right),$$

where

$$m_\kappa(x; t) := xe^{-\theta t} + \mu_\kappa(1 - e^{-\theta t}), \quad \sigma_t^2 := \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t}).$$

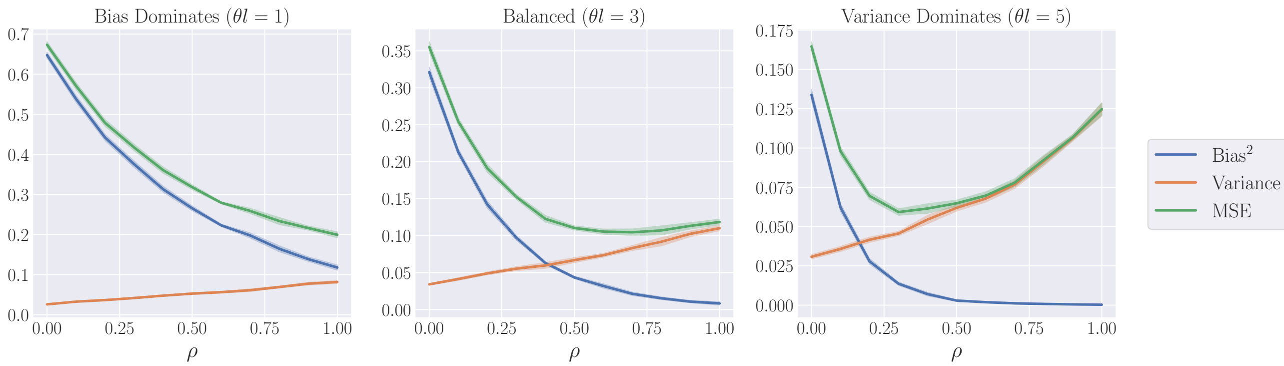

Figure 2: Mean squared error (MSE) decomposition for varying choices of θl and ρ (with 95% CI). Results are based on 10 runs with $n = 8$ periods and a synthetic OU process with parameters $\mu_1 = 1$, $\mu_0 = 0$, and $\sigma = 1$.

Remark 8. *The OU processes under both policies satisfy stochastic monotonicity (Assumption 1); for any $t \geq 0$ and $s \in \mathcal{X}$, $P_\kappa(x, (s, \infty); t)$ is non-decreasing in x. Moreover, the treatment kernel strictly stochastically dominates the control kernel if $\mu_1 > \mu_0$, consistent with Assumption 2. Finally, the OU process is also geometrically ergodic, satisfying Assumption 3 with $\lambda_0 = \lambda_1 = \theta$ and some constant $C \geq 1$.*

Reward. The reward at time t when the process is in state x is given by the value of the state itself, i.e.,

$$r_\kappa(x) = x, \quad \forall x \in \mathcal{X}, \kappa \in \{0, 1\}.$$

The reward function is measurable and non-decreasing with respect to the state $x \in \mathcal{X}$. Therefore, the estimand (ATE) is $\tau = \mu_1 - \mu_0$, and we assume a positive treatment effect, i.e., $\mu_1 > \mu_0$. While Assumption 4 requires boundedness in the general CTMC model, in the OU process, it can be shown that underestimation bias persists even without the boundedness assumption (see, e.g., the proof of Theorem 3).

4.2 Estimation with Washout Periods

To mitigate carryover effects, recent work has proposed introducing *washout periods* that discard part of each interval's initial data (Hu and Wager, 2022). We adopt this approach by introducing a *washout proportion* $\rho \in [0, 1)$. In each interval $[il, (i+1)l)$, the first ρl units are discarded, and the remaining $(1 - \rho)l$ units are used to compute outcomes. For $i \in [n]$,

$$I_{i,\rho}^{(j)} = \frac{1}{l(1-\rho)} \int_{(2i+\rho-j+1)l}^{(2i-j+2)l} X_t \, dt, \quad j \in \{0, 1\}, \quad (11)$$

and the washout-adjusted DM estimator is

$$\hat{\tau}_{n,\rho} = \frac{1}{n} \sum_{i=0}^{n-1} \big(I_{i,\rho}^{(1)} - I_{i,\rho}^{(0)}\big). \quad (12)$$

Variance Estimation. In practice, variance is often estimated using the Huber–White (sandwich) estimator (White, 1980), which in the OU model reduces to

$$\hat{s}_{n,\rho}^2 = \frac{1}{n} \sum_{i=0}^{n-1} \big(I_{i,\rho}^{(1)} - \bar{I}_{n,\rho}^{(1)}\big)^2 + \big(I_{i,\rho}^{(0)} - \bar{I}_{n,\rho}^{(0)}\big)^2, \quad (13)$$

where

$$\bar{I}_{n,\rho}^{(1)} = \frac{1}{n} \sum_{i=0}^{n-1} I_{i,\rho}^{(1)}, \qquad \bar{I}_{n,\rho}^{(0)} = \frac{1}{n} \sum_{i=0}^{n-1} I_{i,\rho}^{(0)}.$$

This estimator is widely used in industry (e.g., DoorDash; Kastelman and Ramesh (2018)) for its simplicity, but it assumes independence across intervals–an assumption violated in the presence of carryover effects. In our ride-sharing case study (Section 5), we find that $\hat{s}_{n,\rho}^2$ tends to overestimate the true variance when $\rho = 0$ and n is large.

4.3 Asymptotic Bias-Variance Trade-offs

In Section 4.1, we showed that the OU process satisfies the conditions of Theorem 2, making it a specific instance of the CTMC model. Moreover, leveraging this stylized model, we conduct a more in-depth analysis of the bias and variance of the DM estimator $\hat{\tau}_{n,\rho}$, and further examine how they are influenced by the washout proportion ρ and period length l. In the following Theorem 3, we characterize the asymptotic bias (B) and variance (V) of the DM estimator $\hat{\tau}_{n,\rho}$ in the limit as the number of switches $n \to \infty$.

Theorem 3. *Under the switchback design (Def. 1), for any $\rho \in [0, 1)$, the asymptotic bias B and variance V of the DM estimator $\hat{\tau}_{n,\rho}$ in (12) as $n \to \infty$ are*

$$B := \lim_{n\to\infty} \mathbb{E}[\hat{\tau}_{n,\rho} - \tau] = -\frac{2\tau(e^{-\rho\theta l} - e^{-\theta l})}{\theta l(1-\rho)(1 + e^{-\theta l})},$$

$$V := \lim_{n\to\infty} n \, Var(\hat{\tau}_{n,\rho}) = V_{Interval} + V_{Intra} - V_{Inter},$$

with asymptotic variance components given by:

$$V_{Interval} := \frac{2\sigma^2}{\theta^3 l^2 (1-\rho)^2}\left[\theta l(1-\rho) - 1 + e^{\theta l(\rho-1)}\right],$$
$$V_{Intra} := \frac{2\sigma^2}{\theta^3 l^2 (1-\rho)^2}\frac{(e^{\theta l} - e^{\rho\theta l})^2}{e^{2\theta l}-1}e^{-\theta l(1+\rho)},$$
$$V_{Inter} := \frac{2\sigma^2}{\theta^3 l^2 (1-\rho)^2}\frac{(e^{\theta l} - e^{\rho\theta l})^2}{e^{2\theta l}-1}e^{-\rho\theta l}.$$

Moreover, $\lim_{n\to\infty}\mathbb{E}\left[\hat{s}^2_{n,\rho}\right] = V_{Interval}$ *while* $V_{Intra} < V_{Inter}$*, so the asymptotic bias of* $\hat{s}^2_{n\rho}$ *is positive.*

Corollary 3. *In the switchback design, the DM estimator* $\hat{\tau}_{n,\rho}$ *asymptotically underestimates the ATE* τ*. For all* $\rho \in [0,1)$*,*

$$\lim_{n\to\infty}\mathbb{E}[\hat{\tau}_{n,\rho}] < \tau.$$

The above results yield two key insights about the DM estimator with washout periods: (i) Corollary 3 shows that the treatment effect is asymptotically *underestimated* for any $\rho \in [0,1)$; (ii) Theorem 3 shows that the variance estimator $\hat{s}^2_{n,\rho}$ in (13) ignores cross-interval covariance and thus asymptotically *overestimates* variance. Moreover, Theorem 3 decomposes the variance into *interval variance* V_{Interval}, *intra-policy covariance* V_{Intra}, and *inter-policy covariance* V_{Inter}, and shows the bias of $\hat{s}^2_{n,\rho}$ is

$$\lim_{n\to\infty}\mathbb{E}[\hat{s}^2_{n,\rho}] - V = V_{\text{Inter}} - V_{\text{Intra}} > 0,$$

confirming that the Huber–White variance estimator inflates variance by ignoring inter-policy dependence.

Asymptotic Normality. Combining Theorem 3 with the Gaussian properties of the OU process, we obtain the following asymptotic normality of the DM estimator under washout periods. The corresponding statistical inferences are discussed in Appendix A.

Theorem 4. *In the switchback design, as* $n \to \infty$*, the DM estimator* $\hat{\tau}_{n,\rho}$ *satisfies*

$$\sqrt{n}\,(\hat{\tau}_{n,\rho} - \tau - B) \;\Rightarrow\; \mathcal{N}(0, V),$$

where B *and* V *are the asymptotic bias and variance characterized in Theorem 3.*

4.4 Selection of Washout Proportion and Equilibrium Convergence Speed

Lemma 4 shows that a switchback design with interval length l in an OU process with mean-reversion speed θ is equivalent to one with interval length θl in a canonical OU process. Hence, θl serves as a measure of *equilibrium convergence speed*, which we treat as a single parameter. Moreover, Theorem 3 shows that both the washout proportion ρ and convergence speed θl critically affect bias and variance. Our goal is to identify suitable choices of these parameters for practical use of the DM estimator. Corollary 4 shows that increasing ρ (discarding more data) reduces bias, while Corollary 5 shows it raises the variance of the estimator, and decreases the bias of its variance estimate.

Corollary 4. *For fixed interval length* l *and number of switches* n*, the bias magnitude of* $\hat{\tau}_{n,\rho}$ *decreases monotonically with the washout proportion* ρ*. Moreover, the asymptotic bias is bounded by*

$$\frac{2\tau e^{-\theta l}}{1+e^{-\theta l}} \;\le\; |B| \;\le\; \frac{2(1-e^{-\theta l})\tau}{\theta l(1+e^{-\theta l})}.$$

Corollary 5. *For fixed* l *and* n*, the asymptotic variance* V *increases monotonically with* ρ*, with asymptotic bounds*

$$\frac{2\sigma^2\left(2+e^{\theta l}(\theta l - 2)+\theta l\right)}{\theta^3 l^2 (e^{\theta l}+1)} \;\le\; V \;\le\; \frac{\sigma^2(e^{\theta l}-1)}{\theta(e^{\theta l}+1)}.$$

Moreover, $V_{Interval}$ *grows with* ρ*, while* $V_{Inter} - V_{Intra}$ *remains positive but decreases with* ρ*.*

The bias-variance trade-off of washout periods is intuitive: discarding more data reduces sample size, raising both the interval variance V_{Interval} and total variance V, while shifting observations away from transients and closer to equilibrium, thereby reducing bias.

Figure 2 shows how equilibrium convergence speed θl and washout proportion ρ affect squared bias, variance, and MSE in OU simulations with $\tau = 1$, $\sigma = 1$, and $n = 8$, with confidence intervals based on 10 independent runs. In particular, we observe that slower convergence (smaller θl) lengthens transients and amplifies bias, while faster convergence (larger θl) shortens transients and shifts the trade-off toward variance.

5 CASE STUDY: A RIDE-SHARING SIMULATION

We present a case study on ride-sharing in New York City using Lyft trip records from the Taxi and Limousine Commission (TLC, 2024). The outcome is defined as the number of trips initiated per 15-minute interval. We apply an ARIMA(2,0,2) model[1], decomposing each observation into a treatment-independent trend and a residual. Since the data are observational, we study hypothetical interventions, e.g., offering ride discounts, by treating the residual as a mean-reverting OU process and modeling treatment as a shift in its

[1]Implemented via the Python `statsmodels` package (Box et al., 2015). All experiments were run on a PC with a 12-core CPU and 18 GB of RAM.

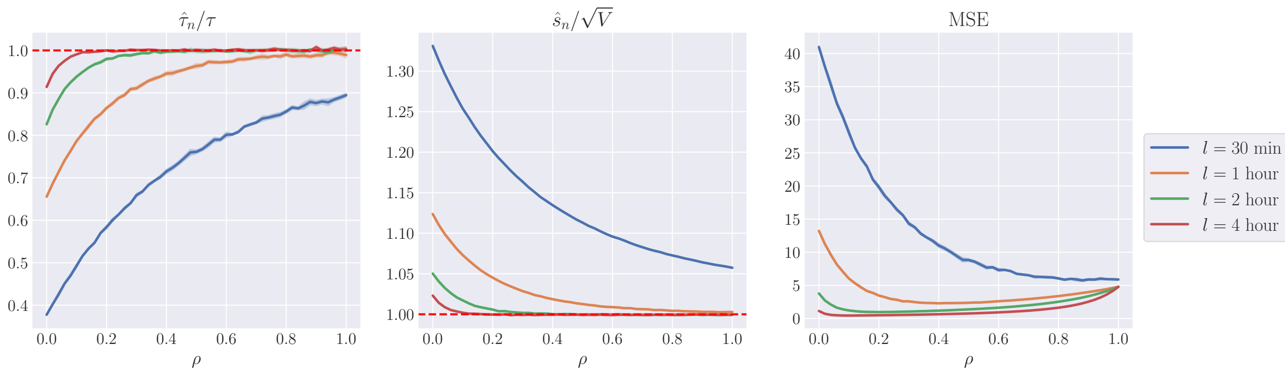

Figure 3: Ratios of $\hat{\tau}_n/\tau$, $\hat{s}_n/\sqrt{V}$, and MSE for varying choices of l and ρ (with 95% CI). Results are based on 10 runs with $n = 1{,}000$ periods under a calibrated OU process with a hypothetical treatment effect $\tau = 10$.

mean. Parameters (μ, θ, σ) are estimated via maximum likelihood. Let X_t be the residual at time t, with m observations at time $\{t_i\}_{i=0}^m$. The log-likelihood is

$$L_m(\mu, \theta, \sigma^2) = -\frac{1}{2}\sum_{i=1}^{m} \log\left(\frac{\sigma^2\pi}{\theta}(1 - e^{-2\theta(t_i - t_{i-1})})\right) - \frac{1}{2}\sum_{i=1}^{m} \frac{(X_{t_i} - \mu + e^{-\theta(t_i - t_{i-1})}(\mu - X_{t_{i-1}}))^2}{\frac{\sigma^2}{2\theta}(1 - e^{-2\theta(t_i - t_{i-1})})}.$$

The MLE is $(\hat{\mu}_m, \hat{\theta}_m, \hat{\sigma}_m^2) \in \arg\max_{\mu,\theta,\sigma^2} L_m(\mu, \theta, \sigma^2)$. Calibrating the model using two weeks of Lyft data from January 2022 (measured in hours) yields the following parameter estimates for the residual process: $\hat{\theta}_m = 5.8$, $\hat{\sigma}_m = 165.7$, and $\hat{\mu}_m = 0.8$, where $\hat{\mu}_m$ denotes the control-policy mean. In the simulations, we set a hypothetical treatment effect $\tau = 10$ (i.e., $\mu_1 = 10$ and $\mu_0 = 0$), vary the interval length $l \in \{1/2, 1, 2, 4\}$ hours, and consider washout proportions $\rho \in [0, 1)$. Each configuration is simulated 10 times with $n = 1{,}000$ to estimate means and 95% confidence intervals (CI).

In Figure 3, the left panel shows the ratio of the estimated treatment effect $\hat{\tau}_n$ from the DM estimator to the underlying ATE τ; the middle panel shows the ratio of the estimated standard deviation $\hat{s}_n$—computed using the Huber–White estimator in (13)—to the true $\sqrt{V}$ estimated via Monte Carlo simulation with 500 samples; and the right panel reports the corresponding mean squared error (MSE). The results reveal a pronounced *downward bias* in $\hat{\tau}_n$ and an *upward bias* in $\hat{s}_n$, both of which diminish as ρ and l increase, consistent with Theorem 3. These patterns indicate that inference based on the DM estimator with the Huber-White variance estimator is conservative under such conditions. The right panel further illustrates the MSE trade-off in selecting ρ and l, i.e., longer switchback intervals l tend to favor smaller washout proportions ρ for minimizing the MSE.

6 CONCLUSION

In this paper, we analyze switchback experiments as commonly implemented in industry within Markovian environments, where treatment effects evolve with the system state. In a CTMC framework with stochastic monotonicity and state-monotone rewards, we show that the widely used difference-in-means estimator is systematically biased downward relative to the stationary estimand, the average treatment effect (ATE). The primary source of bias is a persistent discrepancy between the average transient treatment effect and the stationary ATE, which remains even after within-interval transients have decayed. In a stylized case study based on the OU process, we derive closed-form asymptotics with washout periods and characterize a clear bias–variance trade-off. We further provide asymptotic approximations for key inferential metrics. Finally, a simulation study of NYC ride-sharing data shows that standard switchback pipelines can yield overly conservative inference in practice, underscoring the need for care in design and interpretation.

Our analysis also suggests practical directions for improving switchback experimentation. We identify negative autocorrelation between successive periods as a driver of the estimator's downward bias and inflated variance, which points to a natural direction for future work: developing debiasing procedures that explicitly estimate and adjust for this dependence. The results also motivate *a practical two-stage workflow*, if practitioners expect the treatment to obey the conditions of this paper. Practitioners can first run a standard switchback experiment and assess the estimate using conventional inference. If the ATE is positive, no further adjustment is needed; if results are inconclusive, more advanced designs or debiasing methods can be applied. Because baseline inference is typically conservative, a positive result is likely to remain so after debiasing, supporting an efficient pipeline while allowing more rigorous analysis when needed.

Acknowledgment

This work is partially supported by the Stanford Institute for Human-Centered AI. This work is also partially supported by the National Science Foundation under Grant 2205084. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Bojinov, I., Simchi-Levi, D., and Zhao, J. (2023). Design and analysis of switchback experiments. *Management Science*, 69(7):3759–3777.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Brandt, A. (1938). Tests of significance in reversal or switchback trials, iowa agric. exp. stn. *Bull*, 234.

Bright, I., Delarue, A., and Lobel, I. (2025). Reducing marketplace interference bias via shadow prices. *Management Science*, 71(8):7094–7112.

Chamandy, N. (2016). Experimentation in a ridesharing marketplace lyft engineering.

Cooprider, J. and Nassiri, S. (2023). The science of price experiments in the amazon store.

Dhaouadi, W., Johari, R., Page, O. B., and Weintraub, G. Y. (2023). Price experimentation and interference. *arXiv preprint arXiv:2310.17165*.

Ethier, S. N. and Kurtz, T. G. (2009). *Markov processes: characterization and convergence*. John Wiley & Sons.

Farias, V., Li, A., Peng, T., and Zheng, A. (2022). Markovian interference in experiments. *Advances in Neural Information Processing Systems*, 35:535–549.

Farias, V., Li, H., Peng, T., Ren, X., Zhang, H., and Zheng, A. (2023). Correcting for interference in experiments: A case study at douyin. In *Proceedings of the 17th ACM Conference on Recommender Systems*, pages 455–466.

Glynn, P., Johari, R., and Rasouli, M. (2020). Adaptive experimental design with temporal interference: A maximum likelihood approach. *arXiv preprint arXiv:2006.05591*.

Glynn, P. W. and Infanger, A. (2022). Solutions of poisson's equation for stochastically monotone markov chains. *arXiv preprint arXiv:2202.10578*.

Holtz, D., Lobel, R., Liskovich, I., and Aral, S. (2020). Reducing interference bias in online marketplace pricing experiments. *arXiv preprint arXiv:2004.12489*.

Hu, Y. and Wager, S. (2022). Switchback experiments under geometric mixing. *arXiv preprint arXiv:2209.00197*.

Imbens, G. W. and Rubin, D. B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge university press.

Johari, R., Li, H., Liskovich, I., and Weintraub, G. Y. (2022). Experimental design in two-sided platforms: An analysis of bias. *Management Science*, 68(10):7069–7089.

Johari, R., Li, H., Murthy, A., and Weintraub, G. Y. (2024). When does interference matter? decision-making in platform experiments. *arXiv preprint arXiv:2410.06580*.

Kastelman, D. and Ramesh, R. (2018). Switchback tests and randomized experimentation under network effects at doordash.

Li, H., Zhao, G., Johari, R., and Weintraub, G. Y. (2022). Interference, bias, and variance in two-sided marketplace experimentation: Guidance for platforms. In *Proceedings of the ACM Web Conference 2022*, pages 182–192.

Li, S., Johari, R., Kuang, X., and Wager, S. (2023). Experimenting under stochastic congestion. *arXiv preprint arXiv:2302.12093*.

Lund, R. B. and Tweedie, R. L. (1996). Geometric convergence rates for stochastically ordered markov chains. *Mathematics of operations research*, 21(1):182–194.

Missault, P., Masoero, L., Delbé, C., Richardson, T., and Imbens, G. (2025). Robust and efficient multiple-unit switchback experimentation. *arXiv preprint arXiv:2506.12654*.

Nassiri, S. (2025). Using marketplace marginal values to address interference bias.

Ni, T. and Bojinov, I. (2025). Enhancing efficiency and robustness for switchback experiments: A practical model-assisted framework. *Available at SSRN 5229804*.

O'Connell, G. and Bentes, C. (2022). Tips and considerations for switchback test designs.

Quan, H. (2021). Parameter exploration at lyft.

Senn, S. S. (2002). *Cross-over trials in clinical research*. John Wiley & Sons.

Silbert, N. (2022). Switchback experimentation at tubi.

Stoyan, D. and Daley, D. J. (1983). Comparison methods for queues and other stochastic models. *Journal of the Operational Research Society*.

TLC (2024). TLC Trip Record Data. Accessed: 2024-03-16.

Wager, S. (2024). Causal inference: A statistical learning approach.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, pages 817–838.

Xiong, R., Chin, A., and Taylor, S. (2023). Bias-variance tradeoffs for designing simultaneous temporal experiments. In *The KDD'23 Workshop on Causal Discovery, Prediction and Decision*, pages 115–131. PMLR.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes]
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes]
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Yes]

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes]
   (b) Complete proofs of all theoretical results. [Yes]
   (c) Clear explanations of any assumptions. [Yes]

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Yes]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Yes]
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Yes]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Yes]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:
   (a) Citations of the creator If your work uses existing assets. [Not Applicable]
   (b) The license information of the assets, if applicable. [Not Applicable]
   (c) New assets either in the supplemental material or as a URL, if applicable. [Yes]
   (d) Information about consent from data providers/curators. [Not Applicable]
   (e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]

5. If you used crowdsourcing or conducted research with human subjects, check if you include:
   (a) The full text of instructions given to participants and screenshots. [Not Applicable]
   (b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]
   (c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

Appendix: Conservative Inference in Switchback Experiments

Contents

A Statistical Inference

In the standard industry pipeline for switchback experiments, practitioners often construct statistical inference directly using the DM estimator $\hat{\tau}_{n,\rho}$ and the variance estimator $\hat{s}^2_{n,\rho}$ (Kastelman and Ramesh, 2018). However, due to bias underestimation and variance overestimation, such inference may lead to misleading conclusions. Building on the asymptotic normality of the DM estimator in Theorem 4, we conduct a numerical analysis of the resulting inference procedures in this section. Specifically, we derive the asymptotic type I error, statistical power, and coverage probability and examine how these inferential quantities are affected by different parameter settings, e.g., the washout proportion ρ and the equilibrium convergence speed θl.

A.1 Hypothesis Testing

We begin by considering hypothesis testing for the average treatment effect τ. To test the null hypothesis that the treatment effect is zero, i.e., $H_0 : \tau = 0$, against the two-sided alternative $H_1 : \tau \neq 0$, we construct the following z-statistic:

$$z = \frac{\hat{\tau}_{n,\rho}}{\hat{s}_{n,\rho}/\sqrt{n}}.$$

Decision rule: We reject the null hypothesis H_0 if $|z| \geq \Phi_{\alpha/2}$, and fail to reject H_0 if $|z| < \Phi_{\alpha/2}$. Here, α denotes the significance level, and $\Phi_{\alpha/2}$ is the critical value corresponding to the upper $\alpha/2$ quantile of the standard normal distribution, i.e., the unique value satisfying $\mathbb{P}(W \geq \Phi_{\alpha/2}) = \alpha/2$, where $W \sim \mathcal{N}(0, 1)$.

A.2 Type I Error, Statistical Power and Coverage Probability

For the remainder of the analysis, let $\mathbb{P}_\tau$ denote the probability under the assumption that the treatment effect is τ. The type I error is defined as the probability of rejecting the null hypothesis when there is no treatment effect, i.e.,

$$\text{Type I Error} = \mathbb{P}_0\left(\left|\frac{\hat{\tau}_{n,\rho}}{\sqrt{\hat{s}^2_{n,\rho}/n}}\right| \geq \Phi_{\alpha/2}\right). \tag{14}$$

The statistical power is defined as the probability of rejecting the null hypothesis when the alternative hypothesis is true, i.e.,

$$\text{Power} = \mathbb{P}_\tau\left(\left|\frac{\hat{\tau}_{n,\rho}}{\sqrt{\hat{s}^2_{n,\rho}/n}}\right| \geq \Phi_{\alpha/2}\right). \tag{15}$$

The coverage probability is defined as the probability that the confidence interval includes the true treatment effect, i.e.,

$$\text{Coverage Prob} = \mathbb{P}_\tau\left(\tau \in \left(\hat{\tau}_{n,\rho} - \Phi_{\alpha/2}\frac{\hat{s}_{n,\rho}}{\sqrt{n}}, \hat{\tau}_{n,\rho} + \Phi_{\alpha/2}\frac{\hat{s}_{n,\rho}}{\sqrt{n}}\right)\right). \tag{16}$$

Based on Theorem 4, $\hat{\tau}_{n,\rho} - \tau$ is asymptotically normally distributed as $\mathcal{N}\left(B, \frac{V}{n}\right)$, and $\hat{s}^2_{n,\rho} \approx V_{\text{Interval}}$. Using these two approximations, the type I error in (14), the statistical power in (15), and the coverage probability in (16) can be approximated as follows:

$$\text{Type I Error} \approx \mathbb{P}\left(W \notin \left(-\Phi_{\alpha/2}\sqrt{\frac{V_{\text{Interval}}}{V}} - \frac{B}{\sqrt{V/n}}, \Phi_{\alpha/2}\sqrt{\frac{V_{\text{Interval}}}{V}} - \frac{B}{\sqrt{V/n}}\right)\right), \tag{17}$$

$$\text{Power} \approx \mathbb{P}\left(W \notin \left(-\Phi_{\alpha/2}\sqrt{\frac{V_{\text{Interval}}}{V}} - \frac{\tau + B}{\sqrt{V/n}}, \Phi_{\alpha/2}\sqrt{\frac{V_{\text{Interval}}}{V}} - \frac{\tau + B}{\sqrt{V/n}}\right)\right), \tag{18}$$

$$\text{Coverage Probability} \approx \mathbb{P}\left(W \in \left(-\Phi_{\alpha/2}\sqrt{\frac{V_{\text{Interval}}}{V}} - \frac{B}{\sqrt{V/n}}, \Phi_{\alpha/2}\sqrt{\frac{V_{\text{Interval}}}{V}} - \frac{B}{\sqrt{V/n}}\right)\right), \tag{19}$$

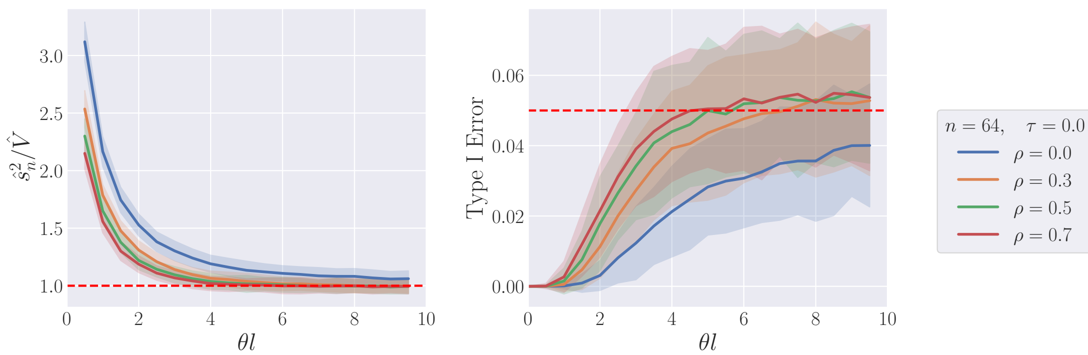

Figure 4: Simulated variance ratio and type I error as functions of θl (with 95% CI). Results are based on 10 runs with $n = 64$ periods under a synthetic OU process with $\sigma = 1$.

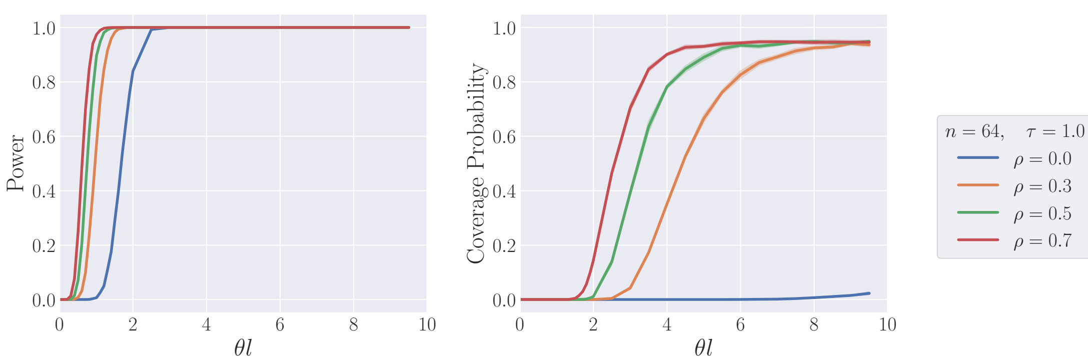

Figure 5: Simulated power and coverage probability as functions of θl (with 95% CI). Results are based on 10 runs with $n = 64$ periods under a synthetic OU process with $\sigma = 1$.

where $W \sim \mathcal{N}(0, 1)$. As shown in (18) and (19), statistical power and coverage probability are governed by two key ratios: (i) the ratio of interval variance to total variance, i.e., V_{Interval}/V, and (ii) the ratio of bias in the estimated treatment effect to the finite-sample standard deviation, i.e., $B/(V/\sqrt{n})$ or $(\tau + B)/(V/\sqrt{n})$. Both ratios are influenced by the equilibrium convergence speed parameter θl and the washout period parameter ρ, while the latter ratio also depends on the number of switches n. To illustrate the impact of these parameters $(\theta l, \rho, n)$ on power and coverage probability, we conduct several numerical simulations.

We simulate switchback experiments with $\sigma = 1$, number of switches $n = 64$, and varying equilibrium convergence speeds θl under different treatment effects (e.g., $\tau = 0, 1/2, 1$). Each parameter setting is evaluated using 500 independent simulations, from which the type I error, statistical power, and coverage probability are estimated. Additionally, we repeat the experiments 10 times (each consisting of 500 simulations) to construct the 95% confidence intervals for these statistics.

As shown in Figure 4, the bias of variance estimation is significantly large for small θl. The variance estimation improves and type I error increases as convergence gets faster and washout period ρ gets longer. Additionally, the sampling variance of type I error increases with faster convergence speed. We also notice that type I error exceeds the nominal level of 0.05 for larger values of the washout proportion ρ and convergence speed θl. Figure 5 shows that statistical power increases with θl under positive treatment effect (e.g., $\tau = 1$). Moreover, a larger washout period proportion ρ results in an earlier rise in statistical power. Similarly, coverage probability also improves with θl in both cases, with a larger ρ leading to an earlier increase, suggesting that longer washout periods contribute to more reliable inference.

In Figure 6, we use the same simulation settings but fix the treatment effect at $\tau = 1$ and plot statistical power and coverage probability as functions of ρ for different choices of n. Additionally, we include the approximated statistical power and coverage probability obtained from the asymptotic expressions in Equations (18) and (19) (via plug-in estimates from Theorem 3). We observe that for larger sample sizes (e.g., $n = 64$), the simulated

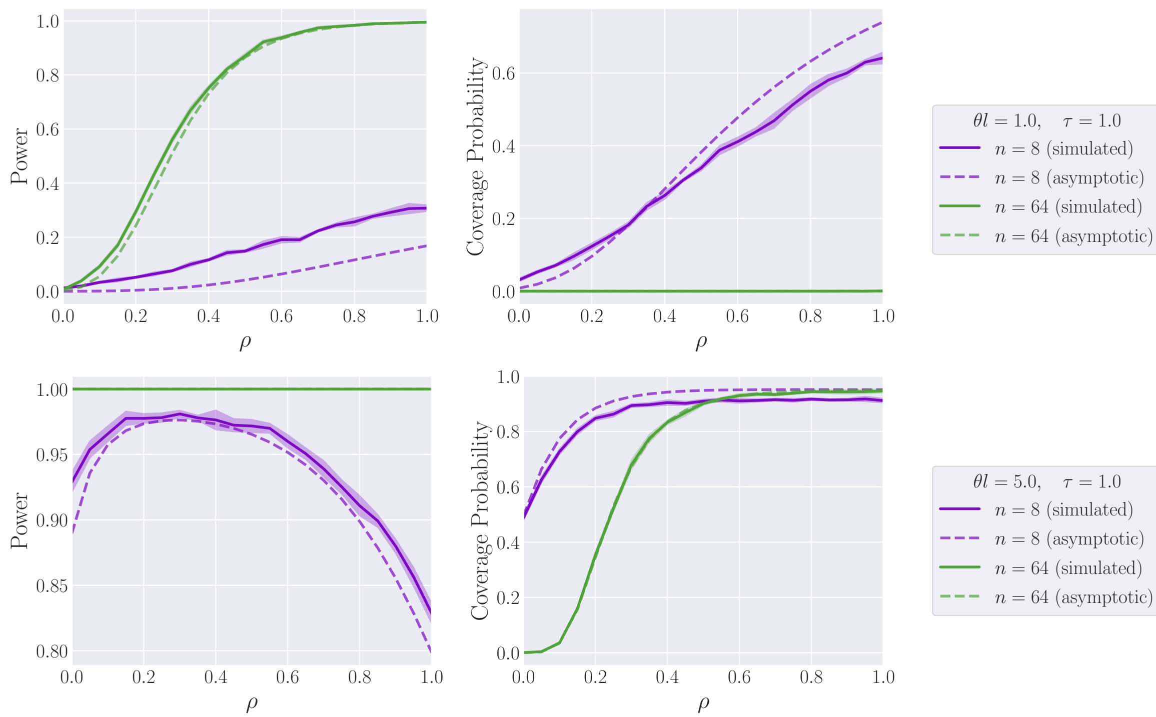

Figure 6: Asymptotic and simulated power and coverage probability as functions of ρ (with 95% CI). Results are based on 10 runs with $n = 8$ and 64 periods under a synthetic OU process with $\sigma = 1$.

statistical power and coverage probability closely align with their asymptotic approximations. Moreover, in the asymptotic regime where n is large, the effect of ρ varies depending on the equilibrium convergence speed. Specifically, when the convergence speed is slow (e.g., $\theta l = 1$), increasing ρ does not necessarily improve coverage probability but can enhance statistical power. Conversely, when the equilibrium convergence speed is fast (e.g., $\theta l = 5$), statistical power remains close to 1 regardless of ρ. In this case, increasing ρ may only improve coverage probability.

B Robustness Analysis under Nonstationarity

An interesting artifact of the standard switchback design is that it is not optimal in a stationary environment, as the treatment effect estimator remains biased. In fact, a much simpler alternative, referred to as the *single-switchback design*, which runs treatment policy for the first half of the experiment and switches to control policy for the second half, yields a consistent estimator as the experiment duration tends to infinity. However, in practice, the *multiple-switchback design* (i.e., the standard approach involving frequent policy switches) is often preferred, as it produces results that are more robust to non-stationarity.

To evaluate robustness, we consider a class of non-stationary environments where the mean process of the underlying OU process is a piecewise constant function, i.e., for $\kappa \in \{0, 1\}$,

$$\mu_{\kappa,u} = \mu_\kappa(i), u \in [i\Delta, (i+1)\Delta).$$

For simplicity, we set $\Delta = 2l$. We consider two types of piecewise constant mean processes: a linear mean process and a periodic mean process.

For the linear mean process, we assume the increments of the mean process are identical, i.e., for $\kappa \in \{0, 1\}$,

$$\mu_\kappa(i) - \mu_\kappa(i-1) = \delta_{\kappa,n}, \quad \forall i \in [n].$$

We impose the following assumption about $\delta_{\kappa,n}$ to make sure the treatment effect is well-defined.

Assumption 5 (Asymptotic linear trend)**.** *For policies $\kappa \in \{0,1\}$, there exists a finite $\gamma_\kappa < \infty$ such that*

$$\lim_{n\to\infty} n\delta_{\kappa,n} \to \gamma_\kappa.$$

The long-run average treatment effect (ATE) τ is defined as

$$\tau = \lim_n \frac{1}{n}\sum_{i=0}^{n-1}(\mu_1(i) - \mu_0(i)) = \mu_1(0) - \mu_0(0) + \frac{\gamma_1 - \gamma_0}{2}$$

For the periodic mean process, we assume that

$$\mu_\kappa(i) = \mu_\kappa^{\mathrm{H}}, \text{ if } i \text{ is even,}$$
$$\mu_\kappa(i) = \mu_\kappa^{\mathrm{L}}, \text{ if } i \text{ is odd.}$$

The treatment effect τ is defined as

$$\tau = \frac{\mu_1^{\mathrm{H}} + \mu_1^{\mathrm{L}}}{2} - \frac{\mu_0^{\mathrm{H}} + \mu_0^{\mathrm{L}}}{2}.$$

With slight abuse of notation, denote by $\hat{\tau}_{n,\rho}$ the treatment effect estimator under the multiple-switchback design. The definition of $\hat{\tau}_{n,\rho}$ follows (12). Let $\hat{\tau}_{n,\rho}^{\mathrm{S}}$ be the treatment effect estimator under the single-switchback design,

$$\hat{\tau}_{n,\rho}^{\mathrm{S}} = \frac{1}{nl(1-\rho)}\left(\int_{\rho nl}^{nl} X_t dt - \int_{nl(1+\rho)}^{2nl} X_t dt\right).$$

Proposition 1 characterizes the asymptotic bias and variance of the estimators $\hat{\tau}_{n,\rho}$ and $\hat{\tau}_{n,\rho}^{\mathrm{S}}$.

Proposition 1 (Asymptotic bias and variance under nonstationarity)**.** *Let $\hat{\tau}_{n,\rho}$ denote the estimator under a multiple-switchback design and $\hat{\tau}_{n,\rho}^{\mathrm{S}}$ the estimator under a single-switchback design, with interval length $l > 0$ and washout ratio $\rho \in (0,1)$. Under Assumption 5, as $n \to \infty$, the limits*

$$\lim_{n\to\infty} \mathbb{E}[\hat{\tau}_{n,\rho} - \tau], \quad \lim_{n\to\infty} n\,\mathrm{Var}(\hat{\tau}_{n,\rho}), \quad \lim_{n\to\infty} \mathbb{E}\left[\hat{\tau}_{n,\rho}^{\mathrm{S}} - \tau\right], \quad \lim_{n\to\infty} n\,\mathrm{Var}\left(\hat{\tau}_{n,\rho}^{\mathrm{S}}\right)$$

exist. They are summarized below for two non-stationary environments.

	Multiple-switchback ($\hat{\tau}_{n,\rho}$)		**Single-switchback** ($\hat{\tau}_{n,\rho}^{\mathrm{S}}$)	
	Asymptotic Bias	Asymptotic Var	Asymptotic Bias	Asymptotic Var
Linear environment	B	V	$\frac{\rho}{4}(\gamma_1 - \gamma_0) - \frac{1}{4}(\gamma_1 + \gamma_0)$	$\frac{2\sigma^2}{\theta^2 l(1-\rho)}$
Periodic environment	B	V	0	$\frac{2\sigma^2}{\theta^2 l(1-\rho)}$

Table 1: Asymptotic bias and variance of treatment-effect estimators under non-stationary environments.

As established in Lemma 4, nonstationarity affects only the deterministic drift of the process and does not alter its stochastic component. Consequently, the asymptotic variance of both estimators remains unchanged under nonstationary environments. Moreover, a direct calculation shows that the multiple-switchback design achieves a strictly smaller variance than the single-switchback design:

$$\lim_{n\to\infty} n\mathrm{Var}(\hat{\tau}_{n,\rho}) < V_{\mathrm{Interval}} < \lim_{n\to\infty} n\mathrm{Var}(\hat{\tau}_{n,\rho}^{\mathrm{S}}).$$

The bias behavior differs markedly across designs. For the multiple-switchback estimator, the asymptotic bias is invariant to changes in the underlying environment and remains equal to B. In contrast, the single-switchback estimator exhibits environment-dependent bias: it is nonzero in the linear environment and vanishes in the periodic one. This distinction arises because, in the multiple-switchback design, bias primarily stems from limited time within each treatment interval for the system to approach equilibrium; hence, it depends mainly on the magnitude of the true treatment effect rather than on long-term trends. The single-switchback estimator, however, relies on a single extended segment of data, making its bias sensitive to persistent drifts or trends in the environment.

Due to this robustness property, inferential procedures derived under stationary assumptions, e.g., bias correction, variance estimation, type-I error control, statistical-power analysis, and coverage guarantees, continue to hold in the nonstationary setting.

C Variance Analysis under a Single-Switchback Design

We investigate the systematic overestimation of variance that can arise when applying the Huber–White variance estimator (White, 1980) to a single-switchback design within the CTMC model. Consider a single-switchback design with total horizon $T = 2l$, where the experimental process X starts from the stationary distribution of the control policy, i.e., $X_0 \sim \pi_0$, then evolves under the treatment policy P_1 for l time units, followed by the control policy P_0 for another l time units. Recall that the average rewards during the first two intervals are $I_0^{(1)}$ and $I_0^{(0)}$. Under stochastic monotonicity and monotone rewards, the covariance between the average outcomes of the two periods is non-negative, i.e., $\operatorname{Cov}(I_0^{(0)}, I_0^{(1)}) \geq 0$. We formalize a more general result in the following lemma. The proof relies on the Markov property to show that, under these assumptions, average rewards across any two periods are positively correlated.

Proposition 2. *Under Assumptions 1 and 4, and the standard switchback design in Definition 1, the covariance of average rewards between any two periods (whether under the same policy or different policies) is non-negative, i.e.,* $\operatorname{Cov}(I_i^{(u)}, I_j^{(v)}) \geq 0$ *for any* $i, j \in [n]$ *and* $u, v \in \{0, 1\}$.

To enable variance estimation for such a single-switchback design, we assume K i.i.d. replications of the single-switchback period, each yielding averages $\{(I_0^{(0)}, I_0^{(1)})\}_k$ for $k \in [K]$. By stacking these $2K$ average outcomes into a regression framework, we can apply the DM estimator, denoted $\hat{\tau}$, using linear regression with the Huber–White variance estimator. From Proposition 2, we observe that the *asymptotic variance*

$$\operatorname{Var}(\hat{\tau}) = \operatorname{Var}(I_0^{(0)}) + \operatorname{Var}(I_0^{(1)}) - 2\operatorname{Cov}(I_0^{(0)}, I_0^{(1)}).$$

In this setting, the HW variance estimator, denoted $\hat{s}^2$, satisfies

$$\lim_{k \to \infty} \mathbb{E}\big[\hat{s}^2\big] = \operatorname{Var}(I_0^{(0)}) + \operatorname{Var}(I_0^{(1)}) \geq \operatorname{Var}(\hat{\tau}),$$

thus asymptotically overestimating the true variance of $\hat{\tau}$.

As observed, the asymptotic overestimation of the HW variance estimator arises from the positive covariance established in Proposition 2 and from the fact that it ignores covariance across intervals (both inter- and intra-policies). In a multiple-switchback design, however, the situation is more nuanced: covariances of the same policy across periods are positive and contribute to the asymptotic variance, whereas covariances between different policies are negative and reduce the asymptotic variance. A similar variance decomposition appears in Theorem 3 for the OU process. In such settings, additional model specification is required to determine whether the asymptotic variance is ultimately overestimated or underestimated.

D Proofs for Section 3

D.1 Proof of Lemma 1

Proof of Lemma 1. For each policy $\kappa \in \{0, 1\}$, the function $g_\kappa : \mathcal{X} \to \mathbb{R}$ is defined as

$$g_\kappa(x) = \int_0^\infty \left(\mathbb{E}_\kappa[r_\kappa(X_t) \mid X_0 = x] - \pi_\kappa r_\kappa\right) dt,$$

which represents the total deviation of the expected reward $r_\kappa(X_t)$ from its stationary average $\pi_\kappa r_\kappa$ over time, starting from state x. Define $h_\kappa(x, s) = \mathbb{E}_\kappa[r_\kappa(X_s) \mid X_0 = x] - \pi_\kappa r_\kappa$ as the deviation of the expected reward at time s. For any $x \in \mathcal{X}$, we bound $h_\kappa(x,s)$ using the uniform bound M on $r_\kappa(x)$ and the total variation distance as follows

$$|h_\kappa(x,s)| = \left| \int_{\mathcal{X}} r_\kappa(y) \left[P_\kappa(x,dy;s) - \pi_\kappa(dy) \right] \right| \leq M \int_{\mathcal{X}} |P_\kappa(x,dy;s) - \pi_\kappa(dy)| = 2M \|P_\kappa(x,\cdot;s) - \pi_\kappa\|_{\mathrm{TV}}.$$

By Assumption 3, the total variation distance decays exponentially as

$$|h_\kappa(x,s)| \leq 2MCe^{-\lambda_\kappa s}.$$

Substituting this bound into the integral defining $g_\kappa(x)$, we obtain

$$|g_\kappa(x)| = \left| \int_0^\infty h_\kappa(x,s) ds \right| \leq \int_0^\infty 2MCe^{-\lambda_\kappa s} ds = \frac{2MC}{\lambda_\kappa}.$$

Thus, $g_\kappa(x)$ is well-defined and bounded as stated. □

D.2 Proof of Lemma 2

Proof of Lemma 2. Define $\mu_i^{(1)}$ and $\mu_i^{(0)}$ as the distributions of the states X_{2il} and $X_{(2i+1)l}$, corresponding to the initial states of the switchback intervals associated with $I_i^{(1)}$ and $I_i^{(0)}$, where the preceding interval (if any) was under the alternative policies P_0 and P_1, respectively. Observe that for any period $i \geq 1$ and state $x' \in \mathcal{X}$, it follows that

$$\mu_i^{(1)}(x') = \int_{\mathcal{X}} \mu_{i-1}^{(0)}(x) P_0(x,x';l)\, dx, \quad \mu_i^{(0)}(x') = \int_{\mathcal{X}} \mu_i^{(1)}(x) P_1(x,x';l)\, dx.$$

For simplicity, we denote this as $\mu_i^{(1)} = \mu_{i-1}^{(0)} P_0(l)$ and $\mu_i^{(0)} = \mu_i^{(1)} P_1(l)$. Additionally, for the initial state X_0, we have $\mu_0^{(1)} = \pi_0$, and for the state X_l at the end of the first treatment interval, we have $\mu_0^{(0)} = \pi_0 P_1(l)$.

We aim to show that for any $i \in [n]$, the following bounds hold

$$\left| \mathbb{E}\left[I_i^{(0)} \right] - \tilde{I}^{(0)} \right| \leq \frac{4MC^2}{\lambda_0 l} \left(e^{-\lambda_0 l} + e^{-\lambda_1 l} \right), \quad \left| \mathbb{E}\left[I_i^{(1)} \right] - \tilde{I}^{(1)} \right| \leq \frac{4MC^2}{\lambda_1 l} \left(e^{-\lambda_0 l} + e^{-\lambda_1 l} \right).$$

First, we can write down the expected average reward of the i-th control interval $I_i^{(0)}$ as

$$\mathbb{E}\left[I_i^{(0)} \right] = \mathbb{E}\left[\frac{1}{l} \int_{(2i+1)l}^{2(i+1)l} r_0(X_t)\, dt \right],$$

where X_s evolves under P_0 from the initial state $X_{(2i+1)l}$. For any $x \in \mathcal{X}$, conditioning on the initial state $X_{(2i+1)l} = x$, the expected cumulative reward over the period of length l evolving under P_0 can be written as (recall that $\mathbb{E}_\kappa$ denotes expectation under policy $\kappa \in \{0,1\}$)

$$\mathbb{E}_0\left[\int_0^l r_0(X_s)\, ds \,\middle|\, X_0 = x \right] = \int_0^l \mathbb{E}_0\left[r_0(X_s) \mid X_0 = x \right] ds,$$

where the expected reward at time s can be written as

$$\mathbb{E}_0\left[r_0(X_s) \mid X_0 = x \right] = \pi_0 r_0 + \left(\mathbb{E}_0\left[r_0(X_s) \mid X_0 = x \right] - \pi_0 r_0 \right).$$

Define $h_\kappa(x,s)$ as the deviation of the expected reward from its stationary value at time s, starting from the initial state x under the policy $\kappa \in \{0,1\}$. Specifically, we have

$$h_\kappa(x,s) = \mathbb{E}_\kappa\left[r_\kappa(X_s) \mid X_0 = x \right] - \pi_\kappa r_\kappa, \quad g_\kappa(x) = \int_0^\infty h_\kappa(x,s) ds.$$

It follows that

$$\mathbb{E}_0\left[\int_0^l r_0(X_s)\, ds \,\middle|\, X_0 = x \right] = \int_0^l \left(\pi_0 r_0 + h_0(x,s) \right) ds = l \pi_0 r_0 + \int_0^l h_0(x,s) ds.$$

Denote the cumulative deviation of the expected reward over the period as

$$G_\kappa(x,l) = \int_0^l h_\kappa(x,s)ds.$$

Then, we can write

$$\begin{aligned}
\mathbb{E}\left[I_i^{(0)}\right] &= \int_{\mathcal{X}} \mathbb{E}_0\left[\int_0^l \frac{1}{l} r_0\left(X_s\right) ds \,\middle|\, X_0 = x\right] \mu_i^{(0)}(dx) \\
&= \int_{\mathcal{X}} \left(\pi_0 r_0 + \frac{1}{l} G_0(x,l)\right) \mu_i^{(0)}(dx) \\
&= \pi_0 r_0 + \frac{1}{l}\int_{\mathcal{X}} G_0(x,l)\mu_i^{(0)}(dx),
\end{aligned}$$

where $G_\kappa(x,l)$ is related to $g_\kappa(x)$ by

$$G_\kappa(x,l) = g_\kappa(x) - \int_l^\infty h_\kappa(x,s)ds.$$

Therefore, we have

$$\begin{aligned}
\mathbb{E}\left[I_i^{(0)}\right] &= \pi_0 r_0 + \frac{1}{l}\int_{\mathcal{X}} \left(g_0(x) - \int_l^\infty h_0(x,s)ds\right) \mu_i^{(0)}(dx) \\
&= \pi_0 r_0 + \frac{1}{l}\int_{\mathcal{X}} g_0(x)\mu_i^{(0)}(dx) - \frac{1}{l}\int_{\mathcal{X}}\int_l^\infty h_0(x,s)ds\mu_i^{(0)}(dx),
\end{aligned}$$

and the corresponding bias can be expressed by

$$\mathbb{E}\left[I_i^{(0)}\right] - \tilde{I}^{(0)} = \frac{1}{l}\int_{\mathcal{X}} g_0(x)\mu_i^{(0)}(dx) - \frac{1}{l}\pi_1 g_0 - \frac{1}{l}\int_{\mathcal{X}}\int_l^\infty h_0(x,s)ds\mu_i^{(0)}(dx).$$

Since r_κ is bounded by a sufficiently large M, and by the definition of the total variation bounds, for any initial state $x \in \mathcal{X}$ and policy $\kappa \in \{0,1\}$, we have

$$\begin{aligned}
|h_\kappa(x,s)| &= \left|\mathbb{E}_\kappa\left[r_\kappa\left(X_s\right) \mid X_0 = x\right] - \pi_\kappa r_\kappa\right| \\
&= \left|\int_{\mathcal{X}} r_\kappa(y)\left[P_\kappa(x,dy;s) - \pi_\kappa(dy)\right]\right| \\
&\le \int_{\mathcal{X}} |r_\kappa(y)|\left|P_\kappa(x,dy;s) - \pi_\kappa(dy)\right| \\
&\le M\int_{\mathcal{X}} \left|P_\kappa(x,dy;s) - \pi_\kappa(dy)\right| \\
&= 2M\left\|P_\kappa(x,\cdot;s) - \pi_\kappa\right\|_{\mathrm{TV}}.
\end{aligned}$$

By Assumption 3, we can derive the upper bound for any $x \in \mathcal{X}$ as

$$|h_\kappa(x,s)| \le 2MCe^{-\lambda_\kappa s}.$$

Then, it follows that

$$\begin{aligned}
\left|\int_{\mathcal{X}}\int_l^\infty h_\kappa(x,s)ds\mu_i^{(\kappa)}(dx)\right| &\le \int_{\mathcal{X}}\int_l^\infty |h_\kappa(x,s)|\,ds\mu_i^{(\kappa)}(dx) \\
&\le \int_{\mathcal{X}}\int_l^\infty 2MCe^{-\lambda_\kappa s}ds\mu_i^{(\kappa)}(dx) \\
&= \int_{\mathcal{X}} \mu_i^{(\kappa)}(dx)\frac{2MC}{\lambda_\kappa}e^{-\lambda_\kappa l} \\
&= \frac{2MC}{\lambda_\kappa}e^{-\lambda_\kappa l}.
\end{aligned}$$

Additionally, Lemma 1 implies that $\|g_\kappa\|_\infty \leq \frac{2MC}{\lambda_\kappa}$. Consequently, for $i \in [n]$, we have

$$\begin{aligned}\left|\int_{\mathcal{X}} g_0(x)\mu_i^{(0)}(dx) - \pi_1 g_0\right| &= \left|\int_{\mathcal{X}} g_0(x)\mu_i^{(0)}(dx) - \int_{\mathcal{X}} g_0(x)\pi_1(dx)\right| \\ &= \left|\int_{\mathcal{X}} g_0(x)\left(\mu_i^{(0)}(dx) - \pi_1(dx)\right)\right| \\ &\leq 2\|g_0\|_\infty \left\|\mu_i^{(0)} - \pi_1\right\|_{\mathrm{TV}} \\ &= 2\|g_0\|_\infty \left\|\mu_i^{(1)} P_1(l) - \pi_1\right\|_{\mathrm{TV}} \\ &\leq 2\|g_0\|_\infty \int_{\mathcal{X}} \mu_i^{(1)}(dx)\left\|P_1(x,\cdot;l) - \pi_1\right\|_{\mathrm{TV}} \\ &\leq \frac{4MC^2}{\lambda_0} e^{-\lambda_1 l}.\end{aligned}$$

Combining those results, we have

$$\left|\mathbb{E}\left[I_i^{(0)}\right] - \tilde{I}^{(0)}\right| = \frac{1}{l}\left(\left|\int_{\mathcal{X}} g_0(x)\mu_i^{(0)}(dx) - \pi_1 g_0\right| + \left|\int_{\mathcal{X}}\int_l^\infty h_0(x,s)ds\mu_i^{(0)}(dx)\right|\right) \leq \frac{4MC^2}{\lambda_0 l}\left(e^{-\lambda_0 l} + e^{-\lambda_1 l}\right).$$

Similarly, one can show that

$$\mathbb{E}\left[I_i^{(1)}\right] - \tilde{I}^{(1)} = \frac{1}{l}\int_{\mathcal{X}} g_1(x)\mu_i^{(1)}(dx) - \frac{1}{l}\pi_0 g_1 - \frac{1}{l}\int_{\mathcal{X}}\int_l^\infty h_1(x,s)ds\mu_i^{(1)}(dx),$$

where for all $i \geq 1$ (and for $i = 0$, since $\mu_0^{(1)} = \pi_0$, the argument still applies), we have

$$\begin{aligned}\left|\int_{\mathcal{X}} g_1(x)\mu_i^{(1)}(dx) - \pi_0 g_1\right| &= \left|\int_{\mathcal{X}} g_1(x)\mu_i^{(1)}(dx) - \int_{\mathcal{X}} g_1(x)\pi_0(dx)\right| \\ &= \left|\int_{\mathcal{X}} g_1(x)\left(\mu_i^{(1)}(dx) - \pi_0(dx)\right)\right| \\ &\leq 2\|g_1\|_\infty \left\|\mu_i^{(1)} - \pi_0\right\|_{\mathrm{TV}} \\ &= 2\|g_1\|_\infty \left\|\mu_{i-1}^{(0)} P_0(l) - \pi_0\right\|_{\mathrm{TV}} \\ &\leq 2\|g_1\|_\infty \int_{\mathcal{X}} \mu_{i-1}^{(0)}(dx)\left\|P_0(x,\cdot;l) - \pi_0\right\|_{\mathrm{TV}} \\ &\leq \frac{4MC^2}{\lambda_1} e^{-\lambda_0 l}.\end{aligned}$$

Finally, we conclude that

$$\left|\mathbb{E}\left[I_i^{(1)}\right] - \tilde{I}^{(1)}\right| = \frac{1}{l}\left(\left|\int_{\mathcal{X}} g_1(x)\mu_i^{(1)}(dx) - \pi_0 g_1\right| + \left|\int_{\mathcal{X}}\int_l^\infty h_1(x,s)ds\mu_i^{(1)}(dx)\right|\right) \leq \frac{4MC^2}{\lambda_1 l}\left(e^{-\lambda_0 l} + e^{-\lambda_1 l}\right).$$

□

D.3 Proof of Lemma 3

Proof of Lemma 3. Under Assumption 2, for any bounded, measurable, non-decreasing function $f : \mathcal{X} \to \mathbb{R}$ and all $t \geq 0$, it follows that

$$\int_{\mathcal{X}} \mathbb{E}_0[f(X_t) \mid X_0 = x]\pi_0(dx) \leq \int_{\mathcal{X}} \mathbb{E}_1[f(X_t) \mid X_0 = x]\pi_0(dx),$$

Given π_0 is the stationary distribution under P_0, for all $t \geq 0$, we have

$$\int_{\mathcal{X}} \mathbb{E}_0[f(X_t) \mid X_0 = x]\pi_0(dx) = \int_{\mathcal{X}}\int_{\mathcal{X}} f(y)\, P_0(x, dy; t)\, \pi_0(dx) = \pi_0 f.$$

Since P_1 is also ergodic, by the mixing condition in Assumption 3, for any initial state $x \in \mathcal{X}$ and any bounded measurable function f, we have

$$\lim_{t\to\infty} \mathbb{E}_1 \left[f\left(X_t\right) \mid X_0 = x\right] = \int_{\mathcal{X}} f(y)\pi_1(dy) = \pi_1 f.$$

Additionally, since f is bounded, there exists $N > 0$ such that $|f(x)| \leq N$ for all $x \in \mathcal{X}$. Then, one can apply the Dominated Convergence Theorem and derive

$$\lim_{t\to\infty} \int_{\mathcal{X}} \mathbb{E}_1 \left[f\left(X_t\right) \mid X_0 = x\right] \pi_0(dx) = \int_{\mathcal{X}} \left(\lim_{t\to\infty} \mathbb{E}_1 \left[f\left(X_t\right) \mid X_0 = x\right]\right) \pi_0(dx) = \int_{\mathcal{X}} \pi_1 f \pi_0(dx) = \pi_1 f.$$

Putting everything together, we conclude that

$$\pi_0 f = \lim_{t\to\infty} \int_{\mathcal{X}} \mathbb{E}_0[f(X_t) \mid X_0 = x]\pi_0(dx) \leq \lim_{t\to\infty} \int_{\mathcal{X}} \mathbb{E}_1 \left[f\left(X_t\right) \mid X_0 = x\right] \pi_0(dx) = \pi_1 f.$$

□

D.4 Proof of Theorem 1

Proof. For $T > 0$, define

$$g_T(x) := \int_0^T P_t\bar{r}(x)\, dt.$$

Since X is stochastically monotone, for every $t \geq 0$ and every $x \leq y$,

$$P_t(x, \cdot) \preceq_{\mathrm{st}} P_t(y, \cdot).$$

Because $\bar{r}$ is non-decreasing, it follows that

$$P_t\bar{r}(x) \leq P_t\bar{r}(y), \qquad t \geq 0,\ x \leq y,$$

Hence

$$g_T(x) \leq g_T(y), \qquad x \leq y.$$

By the assumed absolute integrability,

$$g_T(x) \to g(x) := \int_0^\infty P_t\bar{r}(x)\, dt \qquad \text{as } T \to \infty,$$

for every $x \in \mathcal{X}$. Therefore g is non-decreasing.

Next, for $t \geq 0$,

$$P_t|g|(x) \leq \int_0^\infty P_tP_s|\bar{r}|(x)\, ds = \int_0^\infty P_{t+s}|\bar{r}|(x)\, ds = \int_t^\infty P_u|\bar{r}|(x)\, du < \infty,$$

so $P_tg(x) = \mathbb{E}_x[g(X_t)]$ is well defined. By Fubini–Tonelli's theorem and the semigroup property,

$$P_tg(x) = \int_0^\infty P_tP_s\bar{r}(x)\, ds = \int_0^\infty P_{t+s}\bar{r}(x)\, ds = \int_t^\infty P_u\bar{r}(x)\, du.$$

Therefore,

$$P_tg(x) - g(x) = -\int_0^t P_s\bar{r}(x)\, ds.$$

Dividing by t gives

$$\frac{P_tg(x) - g(x)}{t} = -\frac{1}{t}\int_0^t P_s\bar{r}(x)\, ds.$$

Since $P_s\bar{r}(x) \to \bar{r}(x)$ as $s \downarrow 0$, Cesàro's mean theorem implies

$$\frac{1}{t}\int_0^t P_s\bar{r}(x)\,ds \to \bar{r}(x) \qquad \text{as } t \downarrow 0.$$

Hence

$$\lim_{t\downarrow 0} \frac{P_t g(x) - g(x)}{t} = -\bar{r}(x) = \pi r - r(x).$$

Thus $g \in \mathcal{D}(\mathcal{A})$ (i.e., g belongs to the domain of the generator $\mathcal{A}$), and since $\mathcal{A}$ is the infinitesimal generator,

$$\mathcal{A}g(x) = \pi r - r(x).$$

i.e. g solves Poisson's equation. □

D.5 Proof of Theorem 2

Proof of Theorem 2. By Lemma 1 and Theorem 1, the function g_κ is bounded, measurable, and non-decreasing for $\kappa \in \{0, 1\}$. It follows that

$$\pi_\kappa g_\kappa = \int_0^\infty \int_{\mathcal{X}} \left(\mathbb{E}_\kappa[r_\kappa(X_t) \mid X_0 = x] - \pi_\kappa r_\kappa\right) \pi_\kappa(dx)dt = \int_0^\infty (\pi_\kappa r_\kappa - \pi_\kappa r_\kappa)dt = 0.$$

Then, applying Lemma 3, we conclude

$$\pi_0 g_1 \le \pi_1 g_1 = 0 = \pi_0 g_0 \le \pi_1 g_0.$$

Therefore, the estimand discrepancy δ_τ is non-positive as $\pi_0 g_1 - \pi_1 g_0 \le 0$. □

D.6 Proof of Corollaries 1 and 2

The proofs of Corollaries 1 and 2 follow directly from Lemmas 1 and 2.

E Proofs for Section 4

E.1 Proof of Lemma 4

Proof of Lemma 4. Define the rescaled process

$$Z'_t := \frac{\sigma}{\sqrt{\theta}} Z_{\theta t}.$$

Recall that the canonical OU process Z_t satisfies

$$Z_t = -\int_0^t Z_s\,ds + W_t.$$

Thus,

$$Z_{\theta t} = -\int_0^{\theta t} Z_v\,dv + W_{\theta t}.$$

Multiplying both sides by $\frac{\sigma}{\sqrt{\theta}}$, we obtain

$$Z'_t = -\frac{\sigma}{\sqrt{\theta}}\int_0^{\theta t} Z_v\,dv + \frac{\sigma}{\sqrt{\theta}} W_{\theta t}.$$

Now perform a change of variables $v = \theta u$:

$$\int_0^{\theta t} Z_v\,dv = \int_0^t \theta Z_{\theta u}\,du.$$

Hence,

$$Z'_t = -\frac{\sigma}{\sqrt{\theta}}\int_0^t \theta Z_{\theta u}\, du + \frac{\sigma}{\sqrt{\theta}} W_{\theta t} = -\theta \int_0^t Z'_u\, du + \frac{\sigma}{\sqrt{\theta}} W_{\theta t}.$$

Define a rescaled Brownian motion

$$\widetilde{W}_t := \frac{1}{\sqrt{\theta}} W_{\theta t}.$$

Then $\widetilde{W}_t$ is again a standard Brownian motion, and we obtain

$$Z'_t = -\theta \int_0^t Z'_u\, du + \sigma \widetilde{W}_t.$$

Thus, Z'_t solves the SDE

$$dZ'_t = -\theta Z'_t\, dt + \sigma\, d\widetilde{W}_t,$$

whose solution is

$$Z'_t = \sigma \int_0^t e^{-\theta(t-u)}\, d\widetilde{W}_u.$$

Now consider the OU process X_t defined in (8). Its solution is

$$X_t = (X_0 - \mu)e^{-\theta t} + \mu + Z'_t.$$

Substituting $Z'_t = \frac{\sigma}{\sqrt{\theta}} Z_{\theta t}$ yields

$$X_t = (X_0 - \mu)e^{-\theta t} + \mu + \frac{\sigma}{\sqrt{\theta}} Z_{\theta t}.$$

This establishes the canonical representation of X_t, completing the proof. □

E.2 Proof of Theorem 3

Proof of Theorem 3. We begin by analyzing the asymptotic bias. The expectation of the integral $I_{i,\rho}^{(\kappa)}$ for both policies $\kappa \in \{0, 1\}$ can be expressed as

$$\mathbb{E}[I_{i,\rho}^{(1)}] - \mu_1 = \frac{e^{-\rho\theta l} - e^{-\theta l}}{\theta l(1-\rho)}(\mathbb{E}[X_{2il}] - \mu_1),$$
$$\mathbb{E}[I_{i,\rho}^{(0)}] - \mu_0 = \frac{e^{-\rho\theta l} - e^{-\theta l}}{\theta l(1-\rho)}(\mathbb{E}[X_{(2i+1)l}] - \mu_0).$$

Then the asymptotic bias in (12) follows

$$\lim_{n\to\infty} \mathbb{E}\left[\hat{\tau}_{n,\rho} - \tau\right] = \lim_{n\to\infty} \frac{e^{-\rho\theta l} - e^{-\theta l}}{n\theta l(1-\rho)} \sum_{i=0}^{n-1} \big((\mathbb{E}[X_{2il}] - \mu_1) - \big(\mathbb{E}[X_{(2i+1)l}] - \mu_0\big)\big).$$

Notice that under the switchback design, we have the following recursive formula,

$$\mathbb{E}[X_{2il}] = e^{-\theta l}\mathbb{E}[X_{(2i-1)l}] + (1 - e^{-\theta l})\mu_0,$$
$$\mathbb{E}[X_{(2i+1)l}] = e^{-\theta l}\mathbb{E}[X_{2il}] + (1 - e^{-\theta l})\mu_1.$$

To analyze the limiting behavior of $\mathbb{E}[X_{2il}]$ and $\mathbb{E}[X_{(2i+1)l}]$ as $i \to \infty$, we consider the recursive system formed by the two equations. Iterating this process and leveraging the contraction property of the coefficients (as $e^{-\theta l} \in (0, 1)$), the system converges to a steady state, where

$$\lim_{i\to\infty} \mathbb{E}[X_{2il}] = \frac{e^{-\theta l}\mu_1 + \mu_0}{1 + e^{-\theta l}},$$
$$\lim_{i\to\infty} \mathbb{E}[X_{(2i+1)l}] = \frac{e^{-\theta l}\mu_0 + \mu_1}{1 + e^{-\theta l}}.$$

These steady-state expectations substitute into the expression for the asymptotic bias, yielding

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\left[\hat{\tau}_{n,\rho} - \tau\right] &= \frac{e^{-\rho\theta l} - e^{-\theta l}}{\theta l(1-\rho)} \left(\frac{e^{-\theta l}\mu_1 + \mu_0}{1+e^{-\theta l}} - \mu_1 - \frac{e^{-\theta l}\mu_0 + \mu_1}{1+e^{-\theta l}} + \mu_0 \right) \\ &= -\frac{e^{-\rho\theta l} - e^{-\theta l}}{\theta l(1-\rho)} \frac{2\tau}{1+e^{-\theta l}}.\end{aligned}$$

The variance of the estimator $\hat{\tau}_{n,\rho}$ can be expressed as

$$\begin{aligned}\mathrm{Var}\left(\hat{\tau}_{n,\rho}\right) &= \frac{1}{n^2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \mathrm{Cov}\left(I_{i,\rho}^{(1)} - I_{i,\rho}^{(0)}, I_{j,\rho}^{(1)} - I_{j,\rho}^{(0)}\right) \\ &= \frac{1}{n^2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \left(\mathrm{Cov}\left(I_{i,\rho}^{(1)}, I_{j,\rho}^{(1)}\right) + \mathrm{Cov}\left(I_{i,\rho}^{(0)}, I_{j,\rho}^{(0)}\right) - \mathrm{Cov}\left(I_{i,\rho}^{(1)}, I_{j,\rho}^{(0)}\right) - \mathrm{Cov}\left(I_{i,\rho}^{(0)}, I_{j,\rho}^{(1)}\right)\right) \\ &= \frac{1}{n^2}\left[\sum_{i=0}^{n-1} \mathrm{Var}\Big(I_{i,\rho}^{(1)}\Big) + \mathrm{Var}\Big(I_{i,\rho}^{(0)}\Big) + \sum_{i\neq j} \mathrm{Cov}\Big(I_{i,\rho}^{(1)}, I_{j,\rho}^{(1)}\Big) + \mathrm{Cov}\Big(I_{i,\rho}^{(0)}, I_{j,\rho}^{(0)}\Big) - \sum_{i,j} 2\mathrm{Cov}\Big(I_{i,\rho}^{(1)}, I_{j,\rho}^{(0)}\Big)\right],\end{aligned}$$

where $\mathrm{Var}(I_{i,\rho}^{(\kappa)})$ are the variance of each interval $\kappa \in \{0,1\}$, $\mathrm{Cov}(I_{i,\rho}^{(1)}, I_{j,\rho}^{(1)})$, $\mathrm{Cov}(I_{i,\rho}^{(0)}, I_{j,\rho}^{(0)})$ are the within-policy covariance and $\mathrm{Cov}(I_{i,\rho}^{(1)}, I_{j,\rho}^{(0)})$ are the inter-policy covariance. It therefore suffices to analyze the variance structure of the canonical OU process Z in Lemma 4. We omit the variance contribution from the initial condition X_0, since its influence decays exponentially in the asymptotic regime $n \to \infty$. Specifically, the coefficient of $X_0 - \mathbb{E}[X_0]$ in the interval average is of order $O\big(e^{-\theta il}\big)$. Consequently, its contribution to $n\,\mathrm{Var}(\hat{\tau}_{n,\rho})$ can be bounded as

$$n \cdot \frac{1}{n^2}\left(\sum_{i=0}^{n-1} O\big(e^{-\theta il}\big)\right)^2 = O\big(n^{-1}\big) \;\to\; 0,$$

and is therefore asymptotically negligible.

To derive the variance term, we first consider the variance of a general interval I_i, i.e.,

$$I_i = \frac{1}{l(1-\rho)}\int_{(i+\rho)l}^{(i+1)l} X_t dt, \quad i \in [2n],$$

where one can observe $I_{i,\rho}^{(1)} = I_{2i}$ and $I_{i,\rho}^{(0)} = I_{2i+1}$ for all $i \in [n]$.

Then, the variance of a general interval I_i is given by

$$\begin{aligned}\mathrm{Var}(I_i) &= \frac{\sigma^2}{\theta}\mathrm{Var}\left(\frac{1}{l(1-\rho)}\int_{(i+\rho)l}^{(i+1)l} Z(\theta u)du\right) \\ &= \frac{\sigma^2}{\theta^3 l^2(1-\rho)^2}\left(\theta l(1-\rho) + e^{-\theta l(2i+1+\rho)} - \frac{1}{2}e^{-2\theta l(i+\rho)} - \frac{1}{2}e^{-2\theta l(i+1)} + e^{-\theta l(1-\rho)} - 1\right).\end{aligned}$$

Moreover, the covariance term between any two intervals I_i and I_j $(i<j)$ can be derived as

$$\begin{aligned}\mathrm{Cov}(I_i, I_j) &= \frac{\sigma^2}{\theta l^2(1-\rho)^2}\mathrm{Cov}\left(\int_{(i+\rho)l}^{(i+1)l} Z(\theta u)du, \int_{(j+\rho)l}^{(j+1)l} Z(\theta u)du\right) \\ &= \frac{\sigma^2}{2\theta^3 l^2(1-\rho)^2}\Big(e^{-\theta(2+i+j+2\rho)l}(e^{\theta l\rho} - e^{\theta l})^2(e^{\theta(2i+1+\rho)l} - 1)\Big).\end{aligned}$$

Therefore, the asymptotic variance follows

$$\lim_{n\to\infty} n\,\mathrm{Var}(\hat{\tau}_{n,\rho}) = \underbrace{\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\mathrm{Var}\Big(I_{i,\rho}^{(1)}\Big)+\mathrm{Var}\Big(I_{i,\rho}^{(0)}\Big)}_{V_{\text{Interval}}} + \underbrace{\lim_{n\to\infty}\frac{1}{n}\sum_{i\neq j}\mathrm{Cov}\Big(I_{i,\rho}^{(1)},I_{j,\rho}^{(1)}\Big)+\mathrm{Cov}\Big(I_{i,\rho}^{(0)},I_{j,\rho}^{(0)}\Big)}_{V_{\text{Intra}}}$$
$$-\underbrace{\lim_{n\to\infty}\frac{1}{n}\sum_{i,j}2\mathrm{Cov}\Big(I_{i,\rho}^{(1)},I_{j,\rho}^{(0)}\Big)}_{V_{\text{Inter}}},$$

where

$$\begin{aligned}
V_{\text{Interval}} &= \lim_{n\to\infty}\frac{1}{n}\left(\sum_{i=0}^{n-1}\mathrm{Var}(I_{i,\rho}^{(1)})+\mathrm{Var}(I_{i,\rho}^{(0)})\right)\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{2n-1}\mathrm{Var}(I_i)\\
&= \lim_{n\to\infty}\frac{\sigma^2}{n\theta^3l^2(1-\rho)^2}\sum_{i=0}^{2n-1}\left\{\theta l(1-\rho)+e^{-\theta l(2i+1+\rho)}-\frac{1}{2}e^{-2\theta l(i+\rho)}-\frac{1}{2}e^{-2\theta l(i+1)}+e^{-\theta l(1-\rho)}-1\right\}\\
&= \lim_{n\to\infty}\frac{\sigma^2}{n\theta^3l^2(1-\rho)^2}\left\{\sum_{i=0}^{2n-1}\theta l(1-\rho)+\sum_{i=0}^{2n-1}e^{-\theta l(2i+1+\rho)}-\frac{1}{2}\sum_{i=0}^{2n-1}e^{-2\theta l(i+\rho)}\right.\\
&\quad\left.-\frac{1}{2}\sum_{i=0}^{2n-1}e^{-2\theta l(i+1)}+\sum_{i=0}^{2n-1}e^{-\theta l(1-\rho)}-\sum_{i=0}^{2n-1}1\right\}\\
&= \lim_{n\to\infty}\frac{\sigma^2}{n\theta^3l^2(1-\rho)^2}\left\{2n\theta l(1-\rho)+\frac{e^{-\theta l(1+\rho)}\big(1-e^{-4\theta ln}\big)}{1-e^{-2\theta l}}-\frac{1}{2}\cdot\frac{e^{-2\theta l\rho}\big(1-e^{-4\theta ln}\big)}{1-e^{-2\theta l}}\right.\\
&\quad\left.-\frac{1}{2}\cdot\frac{e^{-2\theta l}\big(1-e^{-4\theta ln}\big)}{1-e^{-2\theta l}}+2n\Big(e^{-\theta l(1-\rho)}-1\Big)\right\}\\
&= \lim_{n\to\infty}\frac{\sigma^2}{\theta^3l^2(1-\rho)^2}\left\{2\Big(\theta l(1-\rho)-1+e^{-\theta l(1-\rho)}\Big)+\frac{1}{n}\cdot\frac{1-e^{-4\theta ln}}{1-e^{-2\theta l}}\Big(e^{-\theta l(1+\rho)}-\frac{1}{2}e^{-2\theta l\rho}-\frac{1}{2}e^{-2\theta l}\Big)\right\}\\
&= \frac{2\sigma^2}{\theta^3l^2(1-\rho)^2}(\theta l(1-\rho)-1+e^{\theta(\rho-1)l}).
\end{aligned}$$

Similarly, it can be shown that

$$\begin{aligned}
V_{\text{Intra}} &= \lim_{n\to\infty}\frac{1}{n}\sum_{i\neq j}\mathrm{Cov}(I_{i,\rho}^{(1)},I_{j,\rho}^{(1)})+\mathrm{Cov}(I_{i,\rho}^{(0)},I_{j,\rho}^{(0)})\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{i\neq j}\mathrm{Cov}\left(I_{2i},I_{2j}\right)+\mathrm{Cov}\left(I_{2i+1},I_{2j+1}\right)\\
&= \lim_{n\to\infty}\frac{\sigma^2e^{-\theta l(1+4n+2\rho)}(e^{\theta l\rho}-e^{\theta l})^2}{n(e^{2\theta l}-1)^2\theta^3l^2(1-\rho)^2}\Big(e^{\theta l(2n+1)}-e^{3\theta l}+e^{\theta l(2n+3)}-e^{\theta l(4n+1)}+2e^{\theta l(4n+2+\rho)}(n-1)\\
&\quad+2e^{\theta l(2+2n+\rho)}-2e^{\theta l(4n+\rho)}n\Big)\\
&= \frac{2\sigma^2}{\theta^3l^2(1-\rho)^2}\frac{(e^{\theta l}-e^{\rho\theta l})^2}{(e^{2\theta l}-1)}e^{-\theta(1+\rho)l},
\end{aligned}$$

and

$$
\begin{aligned}
V_{\text{Inter}} &= \lim_{n\to\infty} \frac{1}{n} \sum_{i,j} 2\,\mathrm{Cov}(I^{(1)}_{i,\rho}, I^{(0)}_{j,\rho}) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{i,j} 2\,\mathrm{Cov}\left(I_{2i}, I_{2j+1}\right) \\
&= \lim_{n\to\infty} -\frac{\sigma^2 e^{-2\theta l(2n+\rho)}(e^{\theta l\rho}-e^{\theta l})^2}{n(e^{2\theta l}-1)^2\theta^3 l^2(1-\rho)^2}\Big(e^{\theta l} - 2e^{(1+2n)\theta l} + e^{(1+4n)\theta l} - e^{(2n+\rho)\theta l} - e^{(2+2n+\rho)\theta l} \\
&\quad + e^{\theta l(2+4n+\rho)}(1-2n) + e^{(4n+\rho)\theta l}(2n+1)\Big) \\
&= \frac{2\sigma^2}{\theta^3 l^2(1-\rho)^2}\frac{(e^{\theta l}-e^{\rho\theta l})^2}{(e^{2\theta l}-1)} e^{-\rho\theta l}.
\end{aligned}
$$

We next show $\lim_{n\to\infty}\mathbb{E}[\hat{s}^2_{n,\rho}] = V_{\text{Interval}}$. Denote

$$
\mu^{(1)}_i := \mathbb{E}[I^{(1)}_{i,\rho}], \qquad \mu^{(0)}_i := \mathbb{E}[I^{(0)}_{i,\rho}],
$$

and

$$
\bar{\mu}^{(1)}_{n,\rho} := \frac{1}{n}\sum_{i=0}^{n-1}\mu^{(1)}_i, \qquad \bar{\mu}^{(0)}_{n,\rho} := \frac{1}{n}\sum_{i=0}^{n-1}\mu^{(0)}_i.
$$

Then, we obtain

$$
\begin{aligned}
\mathbb{E}[\hat{s}^2_{n,\rho}] &= \mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1}\big(I^{(1)}_{i,\rho} - \bar{I}^{(1)}_{n,\rho}\big)^2\right] + \mathbb{E}\left[\frac{1}{n}\sum_{i=0}^{n-1}\big(I^{(0)}_{i,\rho} - \bar{I}^{(0)}_{n,\rho}\big)^2\right] \\
&= \frac{1}{n}\sum_{i=0}^{n-1}\mathrm{Var}(I^{(1)}_{i,\rho}) - \mathrm{Var}(\bar{I}^{(1)}_{n,\rho}) + \frac{1}{n}\sum_{i=0}^{n-1}\big(\mu^{(1)}_i - \bar{\mu}^{(1)}_{n,\rho}\big)^2 \\
&\quad + \frac{1}{n}\sum_{i=0}^{n-1}\mathrm{Var}(I^{(0)}_{i,\rho}) - \mathrm{Var}(\bar{I}^{(0)}_{n,\rho}) + \frac{1}{n}\sum_{i=0}^{n-1}\big(\mu^{(0)}_i - \bar{\mu}^{(0)}_{n,\rho}\big)^2.
\end{aligned}
$$

It remains to show that the additional terms vanish as $n\to\infty$. Since

$$
\lim_{i\to\infty}\mathbb{E}[X_{2il}] = \frac{e^{-\theta l}\mu_1+\mu_0}{1+e^{-\theta l}}, \qquad \lim_{i\to\infty}\mathbb{E}[X_{(2i+1)l}] = \frac{e^{-\theta l}\mu_0+\mu_1}{1+e^{-\theta l}},
$$

and the convergence is exponential at rate $e^{-2\theta l i}$, we have

$$
\mathbb{E}[X_{2il}] = \frac{e^{-\theta l}\mu_1+\mu_0}{1+e^{-\theta l}} + O(e^{-2\theta l i}), \qquad \mathbb{E}[X_{(2i+1)l}] = \frac{e^{-\theta l}\mu_0+\mu_1}{1+e^{-\theta l}} + O(e^{-2\theta l i}).
$$

Substituting these into

$$
\begin{aligned}
\mathbb{E}[I^{(1)}_{i,\rho}] - \mu_1 &= \frac{e^{-\rho\theta l}-e^{-\theta l}}{\theta l(1-\rho)}\big(\mathbb{E}[X_{2il}]-\mu_1\big), \\
\mathbb{E}[I^{(0)}_{i,\rho}] - \mu_0 &= \frac{e^{-\rho\theta l}-e^{-\theta l}}{\theta l(1-\rho)}\big(\mathbb{E}[X_{(2i+1)l}]-\mu_0\big),
\end{aligned}
$$

yields

$$
\begin{aligned}
\mu^{(1)}_i &= \mu_1 + \frac{e^{-\rho\theta l}-e^{-\theta l}}{\theta l(1-\rho)}\left(\frac{e^{-\theta l}\mu_1+\mu_0}{1+e^{-\theta l}} - \mu_1\right) + O(e^{-2\theta l i}), \\
\mu^{(0)}_i &= \mu_0 + \frac{e^{-\rho\theta l}-e^{-\theta l}}{\theta l(1-\rho)}\left(\frac{e^{-\theta l}\mu_0+\mu_1}{1+e^{-\theta l}} - \mu_0\right) + O(e^{-2\theta l i}).
\end{aligned}
$$

Therefore, defining

$$\mu_\infty^{(1)} := \mu_1 + \frac{e^{-\rho\theta l} - e^{-\theta l}}{\theta l(1-\rho)} \left(\frac{e^{-\theta l}\mu_1 + \mu_0}{1 + e^{-\theta l}} - \mu_1 \right),$$
$$\mu_\infty^{(0)} := \mu_0 + \frac{e^{-\rho\theta l} - e^{-\theta l}}{\theta l(1-\rho)} \left(\frac{e^{-\theta l}\mu_0 + \mu_1}{1 + e^{-\theta l}} - \mu_0 \right),$$

we obtain

$$\mu_i^{(1)} - \mu_\infty^{(1)} = O(e^{-2\theta l i}), \qquad \mu_i^{(0)} - \mu_\infty^{(0)} = O(e^{-2\theta l i}).$$

In particular, $\mu_i^{(k)} \to \mu_\infty^{(k)}$ for $k \in \{0,1\}$, and thus by Cesàro's lemma,

$$\bar{\mu}_{n,\rho}^{(1)} \to \mu_\infty^{(1)}, \qquad \bar{\mu}_{n,\rho}^{(0)} \to \mu_\infty^{(0)}.$$

Moreover, the exponential decay implies

$$\frac{1}{n}\sum_{i=0}^{n-1} \big(\mu_i^{(1)} - \bar{\mu}_{n,\rho}^{(1)}\big)^2 \to 0, \qquad \frac{1}{n}\sum_{i=0}^{n-1} \big(\mu_i^{(0)} - \bar{\mu}_{n,\rho}^{(0)}\big)^2 \to 0.$$

Next, using the covariance decomposition,

$$\mathrm{Var}(\bar{I}_{n,\rho}^{(\kappa)}) = \frac{1}{n^2}\sum_{i=0}^{n-1} \mathrm{Var}(I_{i,\rho}^{(\kappa)}) + \frac{2}{n^2}\sum_{0 \le i < j \le n-1} \mathrm{Cov}(I_{i,\rho}^{(\kappa)}, I_{j,\rho}^{(\kappa)}), \qquad \kappa \in \{0,1\},$$

it suffices to control the two terms on the right-hand side. Since $\mathrm{Var}(I_i)$ converges to a finite limit, we have $\sup_i \mathrm{Var}(I_i) < \infty$, so

$$\frac{1}{n^2}\sum_{i=0}^{n-1} \mathrm{Var}(I_{i,\rho}^{(\kappa)}) \le \frac{1}{n}\sup_i \mathrm{Var}(I_i) \to 0.$$

Moreover, from the explicit covariance formula for I_i and I_j, for $i < j$,

$$|\mathrm{Cov}(I_i, I_j)| \le C e^{-c(j-i)}$$

for some constants $C, c > 0$. Therefore,

$$\begin{aligned} \frac{2}{n^2}\sum_{0 \le i < j \le n-1} \big|\mathrm{Cov}(I_{i,\rho}^{(\kappa)}, I_{j,\rho}^{(\kappa)})\big| &\le \frac{2}{n^2}\sum_{h=1}^{n-1} (n-h) C e^{-ch} \\ &\le \frac{2C}{n}\sum_{h=1}^{\infty} e^{-ch} \to 0. \end{aligned}$$

Hence

$$\mathrm{Var}(\bar{I}_{n,\rho}^{(1)}) \to 0, \qquad \mathrm{Var}(\bar{I}_{n,\rho}^{(0)}) \to 0.$$

Combining the above estimates yields

$$\mathbb{E}[\hat{s}_{n,\rho}^2] = \frac{1}{n}\sum_{i=0}^{n-1} \Big(\mathrm{Var}(I_{i,\rho}^{(1)}) + \mathrm{Var}(I_{i,\rho}^{(0)}) \Big) + o(1).$$

Therefore, by the definition of V_{Interval},

$$\lim_{n\to\infty} \mathbb{E}[\hat{s}_{n,\rho}^2] = \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \Big(\mathrm{Var}(I_{i,\rho}^{(1)}) + \mathrm{Var}(I_{i,\rho}^{(0)}) \Big) = V_{\text{Interval}}.$$

□

E.3 Proof of Theorem 4

Proof of Theorem 4. Recall that the estimator $\hat{\tau}_{n,\rho}$ is defined as

$$\hat{\tau}_{n,\rho} = \frac{1}{n}\sum_{i=0}^{n-1}\Big(I_{i,\rho}^{(1)} - I_{i,\rho}^{(0)}\Big),$$

where

$$I_{i,\rho}^{(1)} = \frac{1}{(1-\rho)l}\int_{(2i+\rho)l}^{(2i+1)l} X_t\,dt, \quad I_{i,\rho}^{(0)} = \frac{1}{(1-\rho)l}\int_{(2i+1+\rho)l}^{2(i+1)l} X_t\,dt.$$

Since X_t is an OU process, it is Gaussian. Each $I_{i,\rho}^{(\kappa)}$ is a linear functional of X_t, hence Gaussian. Therefore,

$$Y_i := I_{i,\rho}^{(1)} - I_{i,\rho}^{(0)}$$

is Gaussian, and $\hat{\tau}_{n,\rho} = \frac{1}{n}\sum_{i=0}^{n-1} Y_i$ is Gaussian for each finite n.

From Theorem 3, we have

$$\mathbb{E}[\hat{\tau}_{n,\rho}] - \tau = B + O\left(\frac{1}{n}\right).$$

This follows from the geometric convergence of the two-cycle recursion for $\mathbb{E}[X_{2il}]$ and $\mathbb{E}[X_{(2i+1)l}]$, which implies that the transient bias decays exponentially fast in i, and hence the averaged bias is $O(1/n)$.

Consequently,

$$\sqrt{n}\big(\mathbb{E}[\hat{\tau}_{n,\rho}] - \tau - B\big) \to 0. \tag{20}$$

Again by Theorem 3,

$$n\,\mathrm{Var}(\hat{\tau}_{n,\rho}) \to V.$$

Hence,

$$\mathrm{Var}\big(\sqrt{n}(\hat{\tau}_{n,\rho} - \mathbb{E}[\hat{\tau}_{n,\rho}])\big) = n\,\mathrm{Var}(\hat{\tau}_{n,\rho}) \to V.$$

Since $\hat{\tau}_{n,\rho}$ is Gaussian for each n, the centered and scaled estimator

$$\sqrt{n}\big(\hat{\tau}_{n,\rho} - \mathbb{E}[\hat{\tau}_{n,\rho}]\big)$$

is Gaussian with variance converging to V. Combining this with (20), we obtain

$$\sqrt{n}(\hat{\tau}_{n,\rho} - \tau - B) = \sqrt{n}(\hat{\tau}_{n,\rho} - \mathbb{E}[\hat{\tau}_{n,\rho}]) + \sqrt{n}(\mathbb{E}[\hat{\tau}_{n,\rho}] - \tau - B) \Rightarrow \mathcal{N}(0,V).$$

□

E.4 Proof of Corollary 3

The proof of Corollary 3 follows directly from Theorem 3.

E.5 Proof of Corollary 4

Proof of Corollary 4. Recall the absolute asymptotic bias of $\hat{\tau}_{n,\rho}$ is given by

$$|B| = \frac{2\tau(e^{-\rho\theta l} - e^{-\theta l})}{\theta l(1-\rho)(1+e^{-\theta l})}.$$

Taking the derivative with respect to ρ, we have

$$\frac{d|B|}{d\rho} = \frac{2e^{-l\theta\rho}\left(-e^{l\theta\rho} + e^{l\theta}(1 + l\theta(-1+\rho))\right)\tau}{(1+e^{l\theta})\,l\theta(-1+\rho)^2}.$$

Since we assume a positive treatment effect $\tau > 0$, to demonstrate that $|B|$ is monotonically decreasing with ρ, it is sufficient to show that

$$\Delta(\rho) := -e^{l\theta\rho} + e^{l\theta}(1 + l\theta(-1+\rho)) < 0 \quad \text{for all } \rho \in [0,1).$$

Note that for all $\rho \in [0,1)$, the derivative of $\Delta(\rho)$ is

$$\frac{d\Delta(\rho)}{d\rho} = l\theta\left(e^{l\theta} - e^{l\theta\rho}\right) > 0.$$

Therefore, $\Delta(\rho)$ is increasing with ρ. Next, evaluate $\Delta(\rho)$ at $\rho = 1$, we have

$$\Delta(1) = -e^{l\theta} + e^{l\theta} = 0$$

Thus, as $\Delta(\rho)$ is increasing with ρ, we conclude that $\Delta(\rho) < 0$ for all $\rho \in [0,1)$.

To derive the lower bound, we apply L'Hôpital's rule for $\rho \to 1^-$ (approaching from the left) and obtain

$$\lim_{\rho\to 1}\left|\lim_{n\to\infty}\mathbb{E}[\hat{\tau}_{n,\rho} - \tau]\right| = \frac{2\tau e^{-\theta l}}{1+e^{-\theta l}}.$$

Finally, the upper and lower bounds are summarized as follows

$$\frac{2\tau e^{-\theta l}}{1+e^{-\theta l}} \le \left|\lim_{n\to\infty}\mathbb{E}[\hat{\tau}_{n,\rho} - \tau]\right| \le \frac{2(1-e^{-\theta l})\tau}{\theta l(1+e^{-\theta l})}.$$

□

E.6 Proof of Corollary 5

Proof of Corollary 5. For simplicity, let us denote $x = \rho\theta l$ and $y = \theta l$. Note that for $\rho \in [0,1)$, we have $x \in [0,y)$. Then, the asymptotic variance can be written as

$$V = \frac{2\sigma^2}{\theta(y-x)^2(e^{2y}-1)}\left(-e^{x-y} + e^{x+y} + 1 + x - y + e^{2y}(y-x-1) + (e^y - e^x)^2 e^{-x}(e^{-y}-1)\right).$$

The derivative of V with respect to x is

$$\frac{dV}{dx} = \frac{2\sigma^2}{\theta}\frac{e^{-x}(e^x+1)(e^x(-2+x-y) + e^y(2+x-y))}{(x-y)^3(1+e^y)}.$$

To show V is monotonically increasing with ρ, we only need to show that for all $x \in [0,y)$,

$$\Delta(x) := e^x(-2+x-y) + e^y(2+x-y) < 0.$$

To show this, we observe that the derivative of $\Delta(x)$ is

$$\frac{d\Delta(x)}{dx} = e^y + e^x(-1+x-y),$$

and its second derivative is

$$\frac{d^2\Delta(x)}{dx^2} = e^x(x-y).$$

Then, one can observe that $\frac{d^2\Delta(x)}{dx^2} < 0$ for all $x \in [0,y)$, indicating that $\frac{d\Delta(x)}{dx}$ is decreasing. Furthermore, it follows that $\frac{d\Delta(x)}{dx} > \frac{d\Delta(x)}{dx}\big|_{x=y} = 0$ for all $x \in [0,y)$, indicating that $\Delta(x)$ is increasing. Finally, we have

$$\Delta(x) < \Delta(y) = 0, \quad \text{for all } x \in [0,y).$$

Therefore, the asymptotic variance of $\hat{\tau}_{n,\rho}$ is monotonically increasing with respect to the washout period proportion ρ. Consequently, we can establish both a lower bound and an upper bound as follows

$$\frac{2\sigma^2(2 + e^{\theta l}(\theta l - 2) + \theta l)}{\theta^3 l^2(e^{\theta l}+1)} \le V \le \frac{\sigma^2\left(e^{\theta l}-1\right)}{\theta(e^{\theta l}+1)}.$$

Properties for the asymptotic variance components. It can be shown that

$$\frac{d(V_{\text{Interval}})}{dx} = \frac{2e^{-y}\sigma^2\Delta(x)}{(x-y)^3\theta} > 0,$$
$$\frac{d(V_{\text{Inter}} - V_{\text{Intra}})}{dx} = -\frac{2e^{-x-y}(-e^x + e^y)\sigma^2\Delta(x)}{(1+e^y)(x-y)^3\theta} < 0.$$

These derivatives indicate that the asymptotic variance within the switchback intervals, denoted as V_{Interval}, increases with ρ. Conversely, the difference between the asymptotic variance under inter-policy and intra-policy conditions, $V_{\text{Inter}} - V_{\text{Intra}}$, decreases as ρ increases. Therefore, we can derive the upper and lower bounds as follows

$$\frac{2\sigma^2(\theta l - 1 + e^{-\theta l})}{\theta^3 l^2} \leq V_{\text{Interval}} \leq \frac{\sigma^2}{\theta},$$
$$\frac{2\sigma^2}{(1+e^{\theta l})\theta} \leq V_{\text{Inter}} - V_{\text{Intra}} \leq \frac{2\sigma^2(e^{\theta l}-1)^2}{\theta^3 l^2 (e^{\theta l}+1)e^{\theta l}}.$$

□

F Proofs for Appendix

F.1 Proof of Proposition 1

Proof of Proposition 1. For notational simplicity, assume throughout that n is even; the odd case only changes $O(n^{-1})$ boundary terms and does not affect any limit. Set

$$a := e^{-\theta l}, \qquad c_\rho := \frac{e^{-\rho\theta l} - a}{\theta l(1-\rho)}, \qquad b := \frac{1-a^2}{2\theta l}.$$

We first note a common fact that will be used several times. Let

$$\widetilde{X}_t := X_t - \mathbb{E}[X_t].$$

Since the nonstationarity enters only through the deterministic mean process, $\widetilde{X}$ always satisfies the same centered OU equation

$$d\widetilde{X}_t = -\theta \widetilde{X}_t \, dt + \sigma \, dW_t.$$

Hence the covariance structure is independent of the nonstationary drift, and only the mean changes. Therefore the asymptotic variance of the multiple-switchback estimator is the same as in the stationary case:

$$\lim_{n\to\infty} n \operatorname{Var}(\hat{\tau}_{n,\rho}) = V.$$

Multiple-switchback bias: linear environment. For each switchback block i, define

$$u_i := \mathbb{E}[X_{2il}], \qquad v_i := \mathbb{E}[X_{(2i+1)l}], \qquad \tau(i) := \mu_1(i) - \mu_0(i).$$

Over the treatment sub-interval $[2il, (2i+1)l)$, the mean is $\mu_1(i)$, so

$$\mathbb{E}[I_{i,\rho}^{(1)}] - \mu_1(i) = \frac{1}{l(1-\rho)} \int_{\rho l}^{l} e^{-\theta t} \, dt \, \big(u_i - \mu_1(i)\big) = c_\rho \big(u_i - \mu_1(i)\big).$$

Similarly,

$$\mathbb{E}[I_{i,\rho}^{(0)}] - \mu_0(i) = c_\rho \big(v_i - \mu_0(i)\big).$$

Therefore

$$\mathbb{E}[\hat{\tau}_{n,\rho} - \tau] = \frac{c_\rho}{n} \sum_{i=0}^{n-1} \big(u_i - v_i - \tau(i)\big) + \underbrace{\frac{1}{n}\sum_{i=0}^{n-1} \tau^{(i)} - \tau}_{=o(1)}. \tag{21}$$

Now write

$$D_i := u_i - v_i.$$

The endpoint means satisfy, for $i \geq 1$,

$$u_i = a\, v_{i-1} + (1-a)\mu_0(i-1), \qquad v_i = a\, u_i + (1-a)\mu_1(i).$$

Hence

$$\begin{aligned} D_i &= a^2 D_{i-1} + a(1-a)\tau(i-1) + (1-a)\big(\mu_0(i-1) - \mu_1(i)\big) \\ &= a^2 D_{i-1} - (1-a)^2\tau(i-1) + (1-a)\big(\mu_1(i-1) - \mu_1(i)\big). \end{aligned} \tag{22}$$

Under the linear environment,

$$\mu_\kappa(i) = \mu_\kappa(0) + i\,\delta_{\kappa,n}, \qquad \tau(i) = \tau(0) + i(\delta_{1,n} - \delta_{0,n}).$$

To isolate the stationary two-cycle term, define

$$\bar{D}_i := -\frac{1-a}{1+a}\,\tau(i), \qquad \varepsilon_i := D_i - \bar{D}_i.$$

Substituting (22) into this definition yields

$$\varepsilon_i = a^2\varepsilon_{i-1} + \underbrace{\left[\frac{1-a}{1+a}\big(\delta_{1,n} - \delta_{0,n}\big) - (1-a)\delta_{1,n}\right]}_{=:r_n}.$$

By Assumption 5, $n\delta_{\kappa,n} \to \gamma_\kappa$, so $r_n = O(n^{-1})$. Iterating,

$$\varepsilon_i = a^{2i}\varepsilon_0 + r_n \sum_{k=0}^{i-1} a^{2k},$$

and therefore

$$\frac{1}{n}\sum_{i=0}^{n-1} |\varepsilon_i| \leq \frac{|\varepsilon_0|}{n}\sum_{i=0}^{n-1} a^{2i} + \frac{|r_n|}{n}\sum_{i=0}^{n-1}\sum_{k=0}^{i-1} a^{2k} = O(n^{-1}) \to 0.$$

Thus

$$\frac{1}{n}\sum_{i=0}^{n-1}\big(D_i - \tau(i)\big) = -\frac{2}{1+a}\cdot\frac{1}{n}\sum_{i=0}^{n-1}\tau(i) + o(1).$$

Since $\tau = \lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1}\tau(i)$, (21) gives

$$\lim_{n\to\infty} \mathbb{E}[\hat{\tau}_{n,\rho} - \tau] = -\frac{2c_\rho}{1+a}\,\tau = -\frac{2\tau(e^{-\rho\theta l} - e^{-\theta l})}{\theta l(1-\rho)(1+e^{-\theta l})} = B.$$

Multiple-switchback bias: periodic environment. Let

$$x_0 := \lim_{m\to\infty} \mathbb{E}[X_{4ml}], \quad x_1 := \lim_{m\to\infty} \mathbb{E}[X_{(4m+1)l}], \quad x_2 := \lim_{m\to\infty} \mathbb{E}[X_{(4m+2)l}], \quad x_3 := \lim_{m\to\infty} \mathbb{E}[X_{(4m+3)l}].$$

Because one full four-step cycle contracts by the factor $a^4 < 1$, this fixed point is unique. It solves the explicit four-state system

$$x_1 = ax_0 + (1-a)\mu_1^{\mathrm{H}}, \qquad x_2 = ax_1 + (1-a)\mu_0^{\mathrm{H}}, \qquad x_3 = ax_2 + (1-a)\mu_1^{\mathrm{L}}, \qquad x_0 = ax_3 + (1-a)\mu_0^{\mathrm{L}}.$$

Solving gives

$$x_0 = \frac{a^3\mu_1^{\mathrm{H}} + a^2\mu_0^{\mathrm{H}} + a\mu_1^{\mathrm{L}} + \mu_0^{\mathrm{L}}}{1 + a + a^2 + a^3},$$

$$x_1 = \frac{\mu_1^{\mathrm{H}} + a^3\mu_0^{\mathrm{H}} + a^2\mu_1^{\mathrm{L}} + a\mu_0^{\mathrm{L}}}{1+a+a^2+a^3},$$
$$x_2 = \frac{\mu_0^{\mathrm{H}} + a\mu_1^{\mathrm{H}} + a^3\mu_1^{\mathrm{L}} + a^2\mu_0^{\mathrm{L}}}{1+a+a^2+a^3},$$
$$x_3 = \frac{\mu_1^{\mathrm{L}} + a^2\mu_1^{\mathrm{H}} + a\mu_0^{\mathrm{H}} + a^3\mu_0^{\mathrm{L}}}{1+a+a^2+a^3}.$$

Since the sequence of endpoint means converges geometrically to this four-cycle, the Cesàro average in the bias converges to the cycle average. Hence

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}[\hat{\tau}_{n,\rho} - \tau] &= c_\rho \, \frac{(x_0 - \mu_1^{\mathrm{H}}) - (x_1 - \mu_0^{\mathrm{H}}) + (x_2 - \mu_1^{\mathrm{L}}) - (x_3 - \mu_0^{\mathrm{L}})}{2} \\ &= c_\rho \, \frac{\mu_0^{\mathrm{H}} + \mu_0^{\mathrm{L}} - \mu_1^{\mathrm{H}} - \mu_1^{\mathrm{L}}}{1+a}.\end{aligned}$$

Using

$$\tau = \frac{\mu_1^{\mathrm{H}} + \mu_1^{\mathrm{L}}}{2} - \frac{\mu_0^{\mathrm{H}} + \mu_0^{\mathrm{L}}}{2},$$

we obtain

$$\lim_{n\to\infty} \mathbb{E}[\hat{\tau}_{n,\rho} - \tau] = -\frac{2c_\rho}{1+a}\,\tau = -\frac{2\tau(e^{-\rho\theta l} - e^{-\theta l})}{\theta l(1-\rho)(1+e^{-\theta l})} = B.$$

Single-switchback variance. Let

$$T_n := nl(1-\rho), \qquad G_n := \rho n l.$$

Because only the centered process contributes to the variance, it suffices to study

$$A_n := \frac{1}{T_n}\int_{\rho n l}^{nl} \widetilde{X}_t \, dt, \qquad B_n := \frac{1}{T_n}\int_{(1+\rho)nl}^{2nl} \widetilde{X}_t \, dt,$$

so that

$$\hat{\tau}^{\mathrm{S}}_{n,\rho} - \mathbb{E}[\hat{\tau}^{\mathrm{S}}_{n,\rho}] = A_n - B_n.$$

The centered OU covariance is

$$\mathrm{Cov}(\widetilde{X}_s, \widetilde{X}_t) = \frac{\sigma^2}{2\theta} e^{-\theta|t-s|} + O(e^{-\theta(s+t)}).$$

The $O(e^{-\theta(s+t)})$ term contributes only $O(T_n^{-2})$ to each averaged covariance, hence is negligible after multiplication by n. Using only the leading kernel $\frac{\sigma^2}{2\theta}e^{-\theta|t-s|}$, we obtain

$$\begin{aligned}\mathrm{Var}(A_n) &= \frac{\sigma^2}{2\theta T_n^2}\int_0^{T_n}\int_0^{T_n} e^{-\theta|u-v|}\,du\,dv + O(T_n^{-2}) \\ &= \frac{\sigma^2}{2\theta T_n^2}\left(2\int_0^{T_n}\int_0^{v} e^{-\theta(v-u)}\,du\,dv\right) + O(T_n^{-2}) \\ &= \frac{\sigma^2}{\theta^2 T_n} - \frac{\sigma^2(1-e^{-\theta T_n})}{\theta^3 T_n^2} + O(T_n^{-2}).\end{aligned}$$

Likewise,

$$\begin{aligned}\mathrm{Cov}(A_n, B_n) &= \frac{\sigma^2}{2\theta T_n^2}\int_0^{T_n}\int_{T_n+G_n}^{2T_n+G_n} e^{-\theta(v-u)}\,dv\,du + O(T_n^{-2}e^{-\theta G_n}) \\ &= \frac{\sigma^2(1-e^{-\theta T_n})^2 e^{-\theta G_n}}{2\theta^3 T_n^2} + O(T_n^{-2}e^{-\theta G_n}).\end{aligned}$$

Therefore

$$n\,\mathrm{Var}(\hat{\tau}^{\mathrm{S}}_{n,\rho}) = n\big(\mathrm{Var}(A_n) + \mathrm{Var}(B_n) - 2\mathrm{Cov}(A_n, B_n)\big) = 2n\,\mathrm{Var}(A_n) + o(1),$$

because $n\mathrm{Cov}(A_n, B_n) = O(n^{-1}e^{-\theta G_n}) \to 0$ (including the case $\rho = 0$). Since $T_n = nl(1-\rho)$,

$$\lim_{n\to\infty} n\,\mathrm{Var}(\hat{\tau}^{\mathrm{S}}_{n,\rho}) = \frac{2\sigma^2}{\theta^2 l(1-\rho)}.$$

Single-switchback bias: linear environment. Let

$$m_n := \left\lceil \frac{\rho n}{2} \right\rceil, \qquad N_n := \frac{n}{2} - m_n.$$

Replacing the retained intervals by the union of the complete $2l$-blocks with indices $i = m_n, \ldots, \frac{n}{2} - 1$ changes the expectation by only $O(n^{-1})$, because at most one partial block is lost at each boundary. Hence

$$\mathbb{E}[\hat{\tau}^{\mathrm{S}}_{n,\rho}] = \frac{1}{N_n} \sum_{i=m_n}^{n/2-1} \Big(\mathbb{E}[J_i^{(1)}] - \mathbb{E}[J_i^{(0)}] \Big) + O(n^{-1}), \tag{23}$$

where

$$J_i^{(1)} := \frac{1}{2l} \int_{2il}^{2(i+1)l} X_t\,dt, \qquad J_i^{(0)} := \frac{1}{2l} \int_{(n+2i)l}^{(n+2i+2)l} X_t\,dt.$$

For the treatment half, define $u_i := \mathbb{E}[X_{2il}]$. Since the process stays under treatment on $[2il, 2(i+1)l)$,

$$u_{i+1} = a^2 u_i + (1-a^2)\mu_1(i), \qquad \mathbb{E}[J_i^{(1)}] = \mu_1(i) + b\big(u_i - \mu_1(i)\big).$$

Writing $e_i := u_i - \mu_1(i)$, we obtain

$$e_{i+1} = a^2 e_i - \delta_{1,n}, \qquad e_i = a^{2i} e_0 - \delta_{1,n} \sum_{k=0}^{i-1} a^{2k}.$$

Thus

$$\frac{1}{N_n} \sum_{i=m_n}^{n/2-1} e_i \to 0.$$

For the control half, define $w_i := \mathbb{E}[X_{(n+2i)l}]$. Then

$$w_{i+1} = a^2 w_i + (1-a^2)\mu_0\Big(i + \frac{n}{2}\Big), \qquad \mathbb{E}[J_i^{(0)}] = \mu_0\Big(i+\frac{n}{2}\Big) + b\Big(w_i - \mu_0\Big(i + \frac{n}{2}\Big)\Big).$$

With $f_i := w_i - \mu_0(i + n/2)$, we have

$$f_{i+1} = a^2 f_i - \delta_{0,n}, \qquad f_i = a^{2i} f_0 - \delta_{0,n} \sum_{k=0}^{i-1} a^{2k},$$

so again

$$\frac{1}{N_n} \sum_{i=m_n}^{n/2-1} f_i \to 0.$$

Substituting these estimates into (23),

$$\lim_{n\to\infty} \mathbb{E}[\hat{\tau}^{\mathrm{S}}_{n,\rho} - \tau] = \lim_{n\to\infty} \left[\frac{1}{N_n} \sum_{i=m_n}^{n/2-1} \mu_1(i) - \frac{1}{N_n} \sum_{i=m_n}^{n/2-1} \mu_0\Big(i + \frac{n}{2}\Big) - \tau \right].$$

Since

$$\mu_\kappa(i) = \mu_\kappa(0) + i\,\delta_{\kappa,n},$$

the two arithmetic means are

$$\frac{1}{N_n}\sum_{i=m_n}^{n/2-1}\mu_1(i)=\mu_1(0)+\frac{m_n+n/2-1}{2}\,\delta_{1,n},$$

$$\frac{1}{N_n}\sum_{i=m_n}^{n/2-1}\mu_0\Big(i+\frac{n}{2}\Big)=\mu_0(0)+\frac{m_n+3n/2-1}{2}\,\delta_{0,n}.$$

Moreover, under the linear environment,

$$\tau=\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\big(\mu_1(i)-\mu_0(i)\big)=\mu_1(0)-\mu_0(0)+\frac{\gamma_1-\gamma_0}{2}.$$

Using $m_n/n\to\rho/2$ and $n\delta_{\kappa,n}\to\gamma_\kappa$, we conclude

$$\lim_{n\to\infty}\mathbb{E}[\hat{\tau}^{\mathrm{S}}_{n,\rho}-\tau]=\frac{\rho}{4}(\gamma_1-\gamma_0)-\frac{1}{4}(\gamma_1+\gamma_0).$$

Single-switchback bias: periodic environment. As above, it suffices up to $O(n^{-1})$ to work with complete $2l$-blocks. For the treatment half, the block-start means $u_i:=\mathbb{E}[X_{2il}]$ satisfy

$$u_{i+1}=a^2u_i+(1-a^2)\mu_1(i),\qquad \mu_1(i)=\begin{cases}\mu_1^{\mathrm{H}}, & i \text{ even},\\ \mu_1^{\mathrm{L}}, & i \text{ odd}.\end{cases}$$

Hence (u_{2m},u_{2m+1}) converges geometrically to the unique two-cycle $(x_{\mathrm{H}},x_{\mathrm{L}})$ solving

$$x_{\mathrm{L}}=a^2x_{\mathrm{H}}+(1-a^2)\mu_1^{\mathrm{H}},\qquad x_{\mathrm{H}}=a^2x_{\mathrm{L}}+(1-a^2)\mu_1^{\mathrm{L}}.$$

Therefore

$$x_{\mathrm{H}}=\frac{a^2\mu_1^{\mathrm{H}}+\mu_1^{\mathrm{L}}}{1+a^2},\qquad x_{\mathrm{L}}=\frac{\mu_1^{\mathrm{H}}+a^2\mu_1^{\mathrm{L}}}{1+a^2},$$

and in particular

$$x_{\mathrm{H}}+x_{\mathrm{L}}=\mu_1^{\mathrm{H}}+\mu_1^{\mathrm{L}}.$$

This is the intermediate computation behind the endpoint limit:

$$\lim_{n\to\infty}\frac{\mathbb{E}[X_{(n-2)l}]+\mathbb{E}[X_{nl}]}{2}=\frac{x_{\mathrm{H}}+x_{\mathrm{L}}}{2}=\frac{\mu_1^{\mathrm{H}}+\mu_1^{\mathrm{L}}}{2}.$$

Since

$$\mathbb{E}[J_i^{(1)}]=\mu_1(i)+b\big(u_i-\mu_1(i)\big),$$

the two-cycle average of the treatment-block expectations is exactly

$$\frac{1}{2}\big(\mathbb{E}[J_{\mathrm{H}}^{(1)}]+\mathbb{E}[J_{\mathrm{L}}^{(1)}]\big)=\frac{\mu_1^{\mathrm{H}}+\mu_1^{\mathrm{L}}}{2}+\frac{b}{2}\Big[(x_{\mathrm{H}}+x_{\mathrm{L}})-(\mu_1^{\mathrm{H}}+\mu_1^{\mathrm{L}})\Big]=\frac{\mu_1^{\mathrm{H}}+\mu_1^{\mathrm{L}}}{2}.$$

The control half is identical. If $y_{\mathrm{H}},y_{\mathrm{L}}$ denote the control-side two-cycle, then

$$y_{\mathrm{H}}=\frac{a^2\mu_0^{\mathrm{H}}+\mu_0^{\mathrm{L}}}{1+a^2},\qquad y_{\mathrm{L}}=\frac{\mu_0^{\mathrm{H}}+a^2\mu_0^{\mathrm{L}}}{1+a^2},\qquad y_{\mathrm{H}}+y_{\mathrm{L}}=\mu_0^{\mathrm{H}}+\mu_0^{\mathrm{L}},$$

so

$$\lim_{n\to\infty}\frac{\mathbb{E}[X_{(2n-2)l}]+\mathbb{E}[X_{2nl}]}{2}=\frac{\mu_0^{\mathrm{H}}+\mu_0^{\mathrm{L}}}{2},$$

and

$$\frac{1}{2}\big(\mathbb{E}[J_{\mathrm{H}}^{(0)}]+\mathbb{E}[J_{\mathrm{L}}^{(0)}]\big)=\frac{\mu_0^{\mathrm{H}}+\mu_0^{\mathrm{L}}}{2}.$$

Thus the Cesàro averages of the retained treatment and control blocks converge to

$$\frac{\mu_1^{\mathrm{H}} + \mu_1^{\mathrm{L}}}{2}, \qquad \frac{\mu_0^{\mathrm{H}} + \mu_0^{\mathrm{L}}}{2},$$

respectively, and therefore

$$\lim_{n\to\infty} \mathbb{E}[\hat{\tau}_{n,\rho}^{\mathrm{S}} - \tau] = \frac{\mu_1^{\mathrm{H}} + \mu_1^{\mathrm{L}}}{2} - \frac{\mu_0^{\mathrm{H}} + \mu_0^{\mathrm{L}}}{2} - \tau = 0.$$

This completes the proof. □

F.2 Proof of Proposition 2

Proof of Proposition 2. Without loss of generality, suppose $0 \le i \le j \le n-1$. Consider the case $u = 1$ and $v = 0$, where the two periods are assigned to different policies (treatment and control), and we examine their covariance. Recall that

$$I_i^{(1)} = \frac{1}{l}\int_{2il}^{(2i+1)l} r_1(X_s)\,ds, \quad I_j^{(0)} = \frac{1}{l}\int_{(2j+1)l}^{2(j+1)l} r_0(X_t)\,dt.$$

Then, the covariance can be written as (by Fubini)

$$\begin{aligned}
\mathrm{Cov}\big(I_i^{(1)}, I_j^{(0)}\big) &= \frac{1}{l^2}\mathrm{Cov}\left(\int_{2il}^{(2i+1)l} r_1(X_s)\,ds,\ \int_{(2j+1)l}^{2(j+1)l} r_0(X_t)\,dt\right) \\
&= \frac{1}{l^2}\left(\mathbb{E}\left[\int_{2il}^{(2i+1)l}\int_{(2j+1)l}^{2(j+1)l} r_1(X_s) r_0(X_t)\,dt\,ds\right] - \mathbb{E}\left[\int_{2il}^{(2i+1)l} r_1(X_s) ds\right]\mathbb{E}\left[\int_{(2j+1)l}^{2(j+1)l} r_0(X_t) dt\right]\right) \\
&= \frac{1}{l^2}\int_{2il}^{(2i+1)l}\int_{(2j+1)l}^{2(j+1)l}\Big(\mathbb{E}\big[r_1(X_s)\,r_0(X_t)\big] - \mathbb{E}[r_1(X_s)]\,\mathbb{E}[r_0(X_t)]\Big)\,dt\,ds.
\end{aligned}$$

By the Markov property and tower property,

$$\mathbb{E}\big[r_1(X_s)\,r_0(X_t)\big] = \mathbb{E}\left[r_1(X_s)\int_{\mathcal{X}} r_0(y)\,P\big(X_s, dy; s \to t\big)\right],$$

where $P\big(x, A; s \to t\big) := \mathbb{P}(X_t \in A \mid X_s = x)$ is the compound transition kernel of the experimental process under the switchback design in Definition 1, obtained by composing the policy-specific kernels (i.e., P_0 and P_1) between times s and t. Since both the treatment and control kernels are stochastically monotone, their composition $P(\cdot,\cdot; s \to t)$ is also stochastically monotone.

Given both r_0 and r_1 are nondecreasing and the compound kernel $P(\cdot,\cdot;\, s \to t)$ is stochastically monotone, the mapping $x \mapsto \int_{\mathcal{X}} r_0(y)\,P\big(x, dy; s \to t\big)$ is itself nondecreasing in x. Hence, by Chebyshev's sum inequality,

$$\mathbb{E}\left[r_1(X_s)\int_{\mathcal{X}} r_0(y)\,P\big(X_s, dy; s \to t\big)\right] \ \ge\ \mathbb{E}[r_1(X_s)]\,\mathbb{E}\left[\int_{\mathcal{X}} r_0(y)\,P\big(X_s, dy; s \to t\big)\right].$$

Then, the integrand is nonnegative and $\mathrm{Cov}\big(I_i^{(1)}, I_j^{(0)}\big) \ge 0$. By a similar argument, the covariance is nonnegative as well when both periods are assigned to the same policy (either treatment or control). □