This is the authors' version of the conference work. The definitive version will be published in KDD 2022 at https://doi.org/10.1145/3534678.3539325. Extended long version of this work, with additional theory results, can be found at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4077638.

Non-stationary A/B Tests

Yuhang Wu
wuyh@berkeley.edu
University of California, Berkeley
USA

Zeyu Zheng*
zyzheng@berkeley.edu
University of California, Berkeley & Amazon
USA

Guangyu Zhang
gyzhang@amazon.com
Amazon
USA

Zuohua Zhang
zhzhang@amazon.com
Amazon
USA

Chu Wang*
chuwang@amazon.com
Amazon
USA

ABSTRACT

A/B tests, also known as online controlled experiments, have been used at scale by data-driven enterprises to guide decisions and test innovative ideas. Meanwhile, nonstationarity, such as the time-of-day effect, can commonly arise in various business metrics. We show that inadequately addressing nonstationarity can cause A/B tests to be statistically inefficient or invalid, leading to wrong conclusions. To address these issues, we develop a new framework that provides appropriate modeling and adequate statistical analysis for nonstationary A/B tests. Without changing the infrastructure for any existing A/B test procedure, we propose a new estimator that views time as a continuous covariate to perform post stratification with a sample-dependent number of stratification levels. We prove central limit theorem in a natural limiting regime under nonstationarity, so that valid large-sample statistical inference is available. We show that the proposed estimator achieves the optimal asymptotic variance among all estimators. When the experiment design phase of an A/B test allows, we propose a new time-grouped randomization approach to make a better balance on treatment and control assignments in presence of time nonstationarity. A brief account of numerical experiments are conducted to illustrate the theoretical analysis.

CCS CONCEPTS

• **Computing methodologies** → *Simulation theory*; • **Mathematics of computing** → *Hypothesis testing and confidence interval computation.*

KEYWORDS

A/B test, non-stationarity, statistical inference, bias correction, variance reduction, central limit theorem

ACM Reference Format:
Yuhang Wu, Zeyu Zheng, Guangyu Zhang, Zuohua Zhang, and Chu Wang. 2022. Non-stationary A/B Tests. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22), August 14–18, 2022, Washington, DC, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3534678.3539325

*Corresponding author. This work was done during Zeyu Zheng's role as an Amazon Visiting Academic.

KDD '22, August 14–18, 2022, Washington, DC, USA

ACM ISBN 978-1-4503-9385-0/22/08.
https://doi.org/10.1145/3534678.3539325

1 INTRODUCTION

A/B tests, also known as A/B testing or online controlled experiments, have been massively used by data-driven companies to evaluate an update in service and product, test innovative ideas, and guide decision making. The simplest A/B tests randomly assign each customer to either a control (A) or a treatment (B), collect outcomes, and conduct statistical inference on the difference between treatment and control. We refer to a seminal book [11] by Kohavi, Tang, and Xu for an accessible and comprehensive discussion on the state of art for A/B tests. Gupta et.al. [4] document that A/B tests are heavily used at Airbnb, Amazon, Booking.com, eBay, Facebook, Google, LinkedIn, Lyft, Microsoft, Netflix, Twitter, Uber, among others, and that tens of thousands of experiments/tests are run by these enterprises, involving millions of customers. A/B tests have been documented to provide strong support for data-driven decisions; see [4, 7, 9–11, 13].

Crucially, valid or large-sample asymptotically valid statistical inference is fundamental to trustworthy design and analysis for A/B tests. Valid statistical inference requires accurate understanding of the bias and variance in analysis for an A/B test, while a failure to do so can hurt the reliability and power of A/B tests, causing wrong decisions [7, 11]. In A/B tests, *bias* occurs when the expectation of the estimator has a systemic difference from the true value; such difference cannot be eliminated by simply increasing sample size and may come from platform error, interference, and flawed design. The existence of unaware bias can cause opposite and wrong decisions. Meanwhile, *variance*, usually referred to as the variance of the estimator, is the most critical element for computing confidence intervals and p-values [6, 8, 11]. Higher variance for an A/B test estimator indicates wider confidence interval, more insignificant p-value, and therefore more likely a failure to timely detect a significant improvement or a significant loss on certain metrics. Different from bias, variance can always be reduced by increasing the sample size or extending the length of the experiment, but this comes at undesirable costs of missing business opportunities or extended losses because of failure to timely identify them. Variance reduction, sometimes referred to as sensitivity improvement, have been a keen focus in the A/B tests literature [2, 3, 12, 14, 15].

In this work, we show that *non-stationarity* can cause invalid and inefficient statistical inference for A/B tests and propose solutions to mitigate the impact of non-stationarity. Here, non-stationarity refers to the time-of-day effect and the day-of-week effect, which may reflect the time-varying nature of customer patterns regarding certain metrics. For example, Figure 1 plots the expected treatment outcome and control outcome as a function of time over a week, for two important metrics on purchase and conversion, using real data from a leading e-commerce company. The exact magnitudes are removed for data privacy reasons. The non-stationarity on the time-of-day effect and the day-of-week effect is clearly demonstrated. Similar non-stationarities can commonly arise in many A/B tests metrics in the aforementioned enterprises. We note that such non-stationarity differentiates from *seasonality*, a term that is often used to capture time-varying pattern that arise at a much longer time horizon, e.g., from season to season. When A/B tests are run over the the horizon of several weeks, they may be free from the potential impact from seasonality, but not from the time-of-day and day-of-week non-stationarity.

First, ignoring or inadequately addressing non-stationarity may cause A/B test estimators to be inefficient and possess higher than necessary variance. In the extreme case, if the estimator completely ignores the non-stationarity and views the data as i.i.d. (independent and identically distributed), a large portion of the estimator variance may be due to the mean fluctuations across time. Standard post stratification method by viewing each day of week as a strata can partially reduce the estimator variance [3, 11, 15], but can still miss significant time-varying fluctuations at more granular level of time. In fact, a key challenge is, "time" is by nature a continuous strata variable, different from other common strata variable such as browser types (Firefox, Edge, and Chrome) that are naturally discrete. Valid statistical inference involving the optimal use of a continuous strata variable is under-explored in the literature. On one hand, a coarse use of the continuous strata variable (e.g., viewing each day as a strata) leads to suboptimality, lost efficiency and high variance. On the other hand, a too granular use (e.g., viewing each minute as a strata) can lead to invalid and un-guaranteed statistical inference. We fill this gap in the literature by providing in Section 3 a statistically principled way to use a continuous strata that achieves asymptotically optimal efficiency and allows the proof of valid statistical inference results.

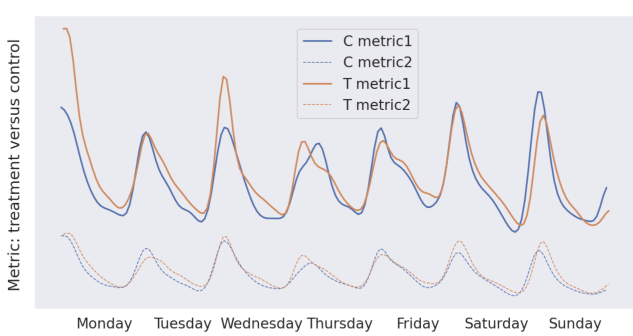


Figure 1: Non-stationarity in the expected treatment (T) and control (C) outcomes for two different metrics related to purchases and conversion

Second, ignoring or inadequately addressing non-stationarity may cause A/B test estimators to have non-negligible bias, which can cause flipped wrong decisions. This bias issue especially arises when the A/B test experimentation uses a *ramping-up* process (see Chapter 15 of [11]) or other reinforcement learning-assisted experimentation algorithms, causing the treatment/control assignment probability to frequently change over the experiment time horizon. Such dynamic change of assignment probability may unintentionally over-weigh the time of the day during when the treatment is much worse than the control, while under-weigh the time of day when treatment is better than the control, leading the overall estimate of the treatment effect to have a wrong sign. This particular bias issue would not exist when there is no non-stationarity, an issue that is similar to the Simpson's paradox but amplified by the frequently changing assignment probabilities. In Section 4, to address this issue, we provide a method *post stratification with continuous strata* that can effectively eliminate this bias issue regardless of how frequent the changes in assignment probabilities are over the experiment time horizon.

Third, when the infrastructure allows changes in the experiment design phase for A/B tests, we propose a simple experiment design approach that uses *time-grouped randomization* in Section 5, which can be an add-on to standard experiment designs. For example, if the current design decides to assign each customer randomly to treatment or control with equal probability, the time-grouped randomization "groups" sequentially arriving customers into consecutive pairs in real time, and then randomly picks one customer in each pair to receive treatment and the other customer in the same pair to receive control. We show that such time-grouped randomization creates no bias and permit simple estimators to enjoy asymptotically optimal efficiency and meanwhile superior finite-sample efficiency. We then show that this time-grouped randomization can naturally be extended to accommodate arbitrary treatment assignment probability. This randomization approach does not need adaptive adjustments over the horizon that the A/B tests are run, so that it does not need timely collection of treatment/control outcomes and has a controllable exposure on latency.

We formulate the non-stationary A/B tests problems in Section 2. In correspond to the issues is the aforementioned first and second items, the core solutions are provided in Section 3 and Section 4; they are developed under the presumption that the existing A/B tests design and infrastructure do not need to be changed and that the techniques can serve as a data analysis tool for all A/B tests experiments that have already been finished. Section 5, in addition provides a time-grouped randomization tool to enhance the experiment design phase when changes in A/B tests design are allowed by the infrastructure.

2 BACKGROUND AND FRAMEWORK

In this section, we describe the typical approach for analyzing standard A/B tests in Section 2.1, without the consideration of non-stationarity. In Section 2.2, we build a framework for analyzing non-stationary A/B tests.

2.1 Standard A/B tests

We describe a very basic setting for analyzing a standard A/B test that has one treatment and one control, without the consideration of non-stationarity. Suppose that the test has been run over a time window $[0, T]$, e.g. two weeks. The data observed from an A/B test is given by $\{I_i, y_i\}_{i=1,2,\cdots,n}$, where I_i takes value 1/0 representing the label of of treatment/control and y_i represents the outcome for the i-th sample. The sample size n is a realization of the total number of samples assigned to this test over $[0, T]$, which can sometimes be the realization of a random variable itself. For example, a specific A/B test may have been decided to take 10% of the total customer flow over the two weeks; in this way, n is a realization of a random variable instead of a pre-specified constant.

The standard analysis of an A/B test is to provide an estimator and evaluate its variance for the average treatment effect given by

$$\alpha \triangleq \mathbb{E}(R_1 - R_0)$$

where R_1 is the random outcome for a random sample if it is assigned to treatment and R_0 is the random outcome for a random sample if it is assigned to control. The variance is then used to construct confidence interval and evaluate p-value. Among the collected data samples $\{I_i, y_i\}_{i=1,2,\cdots,n}$, the $\{y_i\}_{i\in\{j:I_j=1,1\le j\le n\}}$ are typically assumed to be independent and identically distributed (i.i.d.) realizations of R_1, and $\{y_i\}_{i\in\{j:I_j=0,1\le j\le n\}}$ to be i.i.d. realizations of R_0. A naive and standard estimator in this case is the difference of the sample means for the treatment group and the control group (assuming that $\sum_{i=1}^n 1(I_i = 1) \ge 2$ and $\sum_{i=1}^n 1(I_i = 0) \ge 2$):

$$\widehat{\alpha}_{\text{naive}} = \widehat{Y}_1 - \widehat{Y}_0, \tag{1}$$

where

$$\widehat{Y}_1 = \frac{\sum_{i=1}^n 1(I_i = 1)y_i}{\sum_{i=1}^n 1(I_i = 1)},$$

and

$$\widehat{Y}_0 = \frac{\sum_{i=1}^n 1(I_i = 0)y_i}{\sum_{i=1}^n 1(I_i = 0)}.$$

In order to do further statistical inference, we also need to estimate the variance of $\widehat{\alpha}_{\text{naive}}$. It is evident that

$$\text{Var}(\widehat{Y}_1 - \widehat{Y}_0) = \text{Var}(\widehat{Y}_1) + \text{Var}(\widehat{Y}_0),$$

and we can estimate $\text{Var}(\widehat{Y}_1)$ and $\text{Var}(\widehat{Y}_0)$ through sample variances respectively, which gives a variance estimator for $\widehat{\alpha}_{\text{naive}}$ as follows:

$$\begin{aligned}\text{Var}(\widehat{\alpha}_{\text{naive}}) = &\frac{\sum_{i=1}^n 1(I_i = 1)(y_i - \widehat{Y}_1)^2}{\left(\sum_{i=1}^n 1(I_i = 1)\right)\left(\sum_{i=1}^n 1(I_i = 1) - 1\right)} \\ &+ \frac{\sum_{i=1}^n 1(I_i = 0)(y_i - \widehat{Y}_0)^2}{\left(\sum_{i=1}^n 1(I_i = 0)\right)\left(\sum_{i=1}^n 1(I_i = 0) - 1\right)}.\end{aligned} \tag{2}$$

Note that $\text{Var}(\widehat{\alpha}_{\text{naive}})$ is at the order of $O(1/n)$. We denote

$$\widehat{s}^2_{\text{naive}} = n\text{Var}(\widehat{\alpha}_{\text{naive}})$$

to be a re-scaled variance estimator at the order of $O(1)$. An asymptotically valid confidence interval at the coverage level of say 90% will then be constructed as

$$[\widehat{\alpha}_{\text{naive}} - \frac{1.65}{\sqrt{n}}\widehat{s}_{\text{naive}}, \widehat{\alpha}_{\text{naive}} + \frac{1.65}{\sqrt{n}}\widehat{s}_{\text{naive}}].$$

This 90% confidence interval, in turn, is connected with the notion of p-value associated with hypothesis testing. For example, if the null hypothesis is that average treatment effect is zero and the alternative is the average treatment effect is positive, then the observation of a p-value less than 0.05 corresponds to the observation that the above confidence interval has no negative number covered.

2.2 Non-stationary A/B tests

In general for A/B tests that are run over the horizon of several weeks, each sample collected from an A/B test has an associated timestamp. For example, this timestamp may represent the time that the customer arrives or the session starts. With the timestamp information, the data collected from running the experiment over $[0, T]$ are represented by

$$\{I_i, y_i, t_i\}_{i=1,2,\ldots,n} \tag{3}$$

where $t_i \in [0, T]$ represents the timestamp and n represents the realization of a random variable $N(T)$, denoting the total samples assigned to the experiment over $[0, T]$. Similar to the standard setting, I_i takes value 1/0 representing the label of of treatment/control and y_i represents the random outcome for the i-th sample. We introduce the following statistical model to describe the non-stationarity.

Non-stationarity in arrival rates. The time horizon is $[0, T]$. Agents (e.g., customers) sequentially arrive at an A/B test experiment according to a non-homogeneous Poisson process $N = (N(t) : t \in [0, T])$ with time-varying rate function $(\lambda(t) : t \in [0, T])$. For each $t \in (0, T]$, $N(t)$ represents the number of arrivals that occur in the time interval $[0, t]$, which is Poisson distributed with mean $\int_0^t \lambda(s)\, ds$. The time-varying rate function $(\lambda(t) : t \in [0, T])$ is written as

$$\lambda(t) = \bar{\lambda} f(t), \quad t \in [0, T],$$

where $\bar{\lambda} > 0$ is a scale parameter and $f(t)$ represents the normalized rate such that

$$\int_0^T f(t)\, dt = 1.$$

The expectation of the total number of arrivals over the entire time horizon $[0, T]$ is given by

$$\mathbb{E}N(T) = \int_0^T \lambda(t)\, dt = \bar{\lambda}.$$

Therefore the scale parameter $\bar{\lambda}$ also reflects the magnitude of the total number of agents that arrive at and participate in the A/B test within the time horizon of $[0, T]$.

Non-stationarity in treatment assignment probability. When an A/B test has one treatment and one control, an experiment design explicitly or implicitly specifies $(p(t) : t \in [0, T])$, where $p(t)$ denotes the probability that an agent who arrives at time t will be assigned to treatment, while $1 - p(t)$ denotes the probability that an agent who arrives at time t will be assigned to control. In some settings, $p(t)$ is simply set to be equal to $1/2$ over the horizon $[0, T]$. In other settings, $p(t)$ may be set as a step-wise increasing function of t, representing a ramping-up process. There are also settings where $p(t)$ are dynamically adjusted using different strategic plans. Conditional on the specified treatment assignment probabilities $(p(t) : t \in [0, T])$, the arrival flow that is assigned to the treatment follows a non-homogeneous Poisson process with arrival rate

$$\lambda(t)p(t), \quad t \in [0, T],$$

while the arrival flow that is assigned to the control follows a non-homogeneous Poisson process with arrival rate

$$\lambda(t)(1-p(t)), \quad t \in [0,T].$$

Non-stationarity in random outcomes. We build a model for the non-stationary random outcomes. Such model, for example, is able to capture the time-varying behavior of certain metrics related to an A/B test, as illustrated in Figure 1. Let

$$Y_1(t) = \mu_1(t) + \sigma_1(t) \cdot \epsilon_{1,t}$$

be the random outcome for a sample that arrives at time t and is assigned with treatment. Let

$$Y_0(t) = \mu_0(t) + \sigma_0(t) \cdot \epsilon_{0,t}$$

be the random outcome for a sample that arrives at time t and is assigned with control. Here $\mathbb{E}\epsilon_{l,t} = 0$ and $\mathbb{E}\epsilon_{l,t}^2 = 1$ for $l = 0, 1$ and $t \in [0,T]$. The estimation task is then for

$$\alpha = \mathbb{E}(Y_1(\tau) - Y_0(\tau)) = \int_0^T (\mu_1(t) - \mu_0(t)) f(t)\, dt.$$

where τ represents a random arrival time from the non-stationary arrival process $(N(t) : t \in [0,T])$.

Statistical inference for non-stationary A/B tests. We aim to address the following aspects on statistical inference for non-stationary A/B tests. First, suppose we have data $\{I_i, y_i, t_i\}_{i=1,2,\ldots,n}$ collected from an existing A/B test experiment over a time horizon $[0,T]$. We hope to develop an estimator for the average treatment effect that on one hand eliminates potential bias that is caused by ignoring or inadequately addressing non-stationary, and on the other hand makes the optimal use of data to improve test efficiency. We hope to prove central limit theorems. Second, in addition to the data analysis for A/B tests that have been conducted, suppose we are able to modify the experiment design phase for future A/B tests. We hope to develop a new experiment design where the randomization procedure makes the best use of non-stationarity to improve efficiency of A/B tests and prove associated central limit theorems to support valid statistical inference.

3 SAMPLE-BASED POST STRATIFICATION WITH CONTINUOUS COVARIATE

In this section, we consider scenarios where we have data collected from an A/B test experiment that has been run over a time horizon $[0,T]$. The data are represented as $\{I_i, y_i, t_i\}_{i=1,2,\ldots,n}$, defined in (3). We propose a sample-based post stratification with continuous covariate (short as SPS) estimator, develop statistical theory for the estimator, and prove that the SPS estimator achieves the optimal variance in the central limit theorem. The SPS estimator views time as a continuous covariate and performs post stratification with the number of strata asymptotically tending to infinity. Classical post stratification methods such as [1, 3, 15] only consider finite number of discrete strata and their statistical inference does not hold if a continuous strata were to be optimally utilized.

For the convenience of presentation, we presume in this section that the A/B test experiment in consideration randomly assigns each arriving customer to the treatment with probability p and to the control with probability $1-p$. The assignments across different customers are independent. The non-stationary arrival rate $\lambda(t)$ is assumed to be unknown in this section, which can accommodate settings where the A/B test is run over two weeks that involve unique events or market conditions that were never seen before. More general treatment assignment procedures, such as ramping-up by dynamically increase the treatment assignment probability, are considered in the next section (Section 4).

We start by three moderate assumptions.

Assumption 1. *The normalized rate $f(t)$ satisfies that $\Delta_0 < f(t) < \Delta_1$ for some $\Delta_1, \Delta_0 > 0$ and $t \in [0,T]$.*

Assumption 2. *The expected outcomes $\mu_1(t), \mu_0(t)$ as a function of time are Lipschitz continuous, and $\sigma_1(t), \sigma_0(t)$ are continuous. In addition, $\sigma_1(t), \sigma_0(t) > \Delta_2$ for some $\Delta_2 > 0$ and all $t \in [0,T]$.*

Assumption 3. $\lim_{M\to\infty} \sup_{t\in[0,T], l=0,1} \mathbb{E}[\epsilon_{l,t}^2 1_{\{|\epsilon_{l,t}|>M\}}] = 0.$

Assumption 1 assumes that the arrival rate is always positive and bounded from above over the time horizon $[0,T]$; this assumption is valid for many web-based applications. Assumption 2 assumes that the non-stationary changes over time are continuous, which can be relaxed into piecewise continuity for the theory to go through. Assumption 3 regularizes the tail behavior of the random noises, which is stated in a strong form for the ease of presentation. In fact, this assumption only needs to hold for all the random noises that are associated with each arrival epoch.

Given the observed data from an A/B test experiment, the SPS estimator $\widehat{\alpha}_{\text{sps}}$ is constructed as follows. Suppose we have data $\{I_i, y_i, t_i\}_{i=1,2,\ldots,n}$ with $n > 1$ (if $n = 1$, we can not estimate the treatment effect at all), and the total number of strata is $k(n) \in \mathbb{N}_+$. Define

$$T_j = \Big[\frac{j-1}{k(n)}T, \frac{j}{k(n)}T\Big)$$

for $1 \le j \le k(n) - 1$ and

$$T_{k(n)} = \Big[\frac{k(n)-1}{k(n)}T, T\Big].$$

Then define

$$S_{j,l} = \{i : 1 \le i \le n, t_i \in T_j, I_i = l\}$$

for $1 \le j \le k(n)$ and $l = 0, 1$. In addition, define $S_j = S_{j,1} \cup S_{j,0}$ for $1 \le j \le k(n)$. Denote the number of elements of any set A by $|A|$. Then the SPS estimator $\widehat{\alpha}_{\text{sps}}$ is given as

$$\widehat{\alpha}_{\text{sps}} = \sum_{j=1}^{k(n)} \frac{|S_j|}{n} \Big(\frac{\sum_{i\in S_{j,1}} y_i}{|S_{j,1}|} - \frac{\sum_{i\in S_{j,0}} y_i}{|S_{j,0}|} \Big).$$

Here if $S_{j,l} = 0$ for some j, l, we define the ratio $\frac{\sum_{i\in S_{j,1}} y_i}{|S_{j,1}|}$ to be 0, which means if there is no sample in a stratum, we can not estimate the treatment effect of this stratum. In fact, as long as $k(n) = O(n^\beta)$ with $\beta < 1$, it is evident to see

$$P(\pi_{j,1} = 0 \text{ for some } 1 \le j \le k(n)) \le k(n)\Big(1 - \frac{p\Delta_0}{k(n)}\Big)^n \to 0$$

as $n \to \infty$. Similar result holds for $\pi_{j,0}$'s. Since we are only interested in the asymptotic distribution of the estimator, we only consider the case $\pi_{j,l} > 0$ for all $1 \le j \le k(n)$ thereafter.

As we will later see, the choice of T_j can be more flexible. In fact, if we define $\big|[a,b]\big| = b - a$, then as long as $\sup_{1\le j\le k(n)} |T_j| \to 0$

as $n \to \infty$, all the results in this section still hold, but here for simplicity we assume all T_j have the same length.

Now the problems comes down to how to optimally choose $k(n)$ such that the estimator has the optimal asymptotic variance. Previously in classical stratification literature, people often set $k(n) = k$ to be a constant that does not depend on n. However, that may not be optimal. The following theorem gives an optimal result in terms of asymptotic variance.

Theorem 3.1. *Suppose Assumption 1, 2 and 3 are satisfied, $I_i (i \in \mathbb{N})$ are i.i.d. Bernoulli random variables with parameter p, our observations are $\{I_i, y_i, t_i\}_{i=1,2,\dots,n}$. If $\lim_{n\to\infty} k(n) = +\infty$ and $k(n) = o(n^{\frac{1}{2}})$, then*

$$\sqrt{\bar{\lambda}}(\widehat{\alpha}_{sps} - \alpha) \xrightarrow{d} \mathcal{N}(0, V_{sps}), \ \ as\ \bar{\lambda} \to +\infty.$$

Here

$$V_{sps} = \mathrm{Var}\big(\mu_1(\tau) - \mu_0(\tau)\big) + \mathbb{E}\Big[\frac{\sigma_1^2(\tau)}{p} + \frac{\sigma_0^2(\tau)}{1-p}\Big].$$

Proof. We first define some notations which will only be used in this proof. For $1 \le i \le n$, define $g(i) = j$ if and only if $i \in S_j$. Define $\pi_{j,l} = |S_{j,l}|$ for $1 \le j \le k(n)$ and $l = 0, 1$, $\pi_j = \pi_{j,1} + \pi_{j,0}$, $\widehat{p}(t) = p_j = \frac{\pi_{j,1}}{\pi_j}$ when $t \in T_j$. Finally, define $\hat{\mu}_l(t) = m_{j,l} = \mathbb{E}[\mu_l(t)|t \in T_j]$ when $t \in T_j$ for $l = 0, 1$.

Then

$$\begin{aligned}\widehat{\alpha}_{\mathrm{sps}} &= \sum_{j=1}^{k(n)} \frac{|S_j|}{n}\Big(\frac{\sum_{i\in S_{j,1}} y_i}{|S_{j,1}|} - \frac{\sum_{i\in S_{j,0}} y_i}{|S_{j,0}|}\Big)\\ &= \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{I_i Y_1(t_i)}{\widehat{p}(t_i)} - \frac{(1-I_i)Y_0(t_i)}{1-\widehat{p}(t_i)}\Big).\end{aligned}$$

Rearrange terms and we have:

$$\widehat{\alpha}_{\mathrm{ps}} - \alpha = A + B_1 - B_0 + C + D,$$

here

$$A = \frac{1}{n}\sum_{i=1}^{n}\big(\hat{\mu}_1(t_i) - \hat{\mu}_0(t_i) - \alpha\big),$$

$$B_1 = \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{I_i\hat{\mu}_1(t_i)}{p} - \hat{\mu}_1(t_i) + \frac{I_i Y_1(t_i)}{\widehat{p}(t_i)} - \frac{I_i Y_1(t_i)}{p}\Big),$$

$$B_0 = \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{(1-I_i)\hat{\mu}_0(t_i)}{1-p} - \hat{\mu}_0(t_i) + \frac{(1-I_i)Y_0(t_i)}{1-\widehat{p}(t_i)} - \frac{(1-I_i)Y_0(t_i)}{1-p}\Big),$$

$$C = \frac{1}{n}\sum_{i=1}^{n}\frac{I_i(\mu_1(t_i) - \hat{\mu}_1(t_i))}{p} - \frac{(1-I_i)(\mu_0(t_i) - \hat{\mu}_0(t_i))}{1-p},$$

and

$$D = \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{I_i\sigma_1(t_i)\epsilon_{1,t_i}}{p} - \frac{(1-I_i)\sigma_0(t_i)\epsilon_{0,t_i}}{1-p}\Big).$$

First let's consider $A + D$. Define $A_0 = \frac{1}{n}\sum_{i=1}^{n}\big(\mu_1(t_i) - \mu_0(t_i) - \alpha\big)$. Note that in this proof we don't need the order of t_i, that is, we don't require $t_1 \le t_2 \le \cdots \le t_n$, so we can regard $(t_1, \cdots, t_n)$ to be chosen uniformly from all the permutations of $(t_{(1)}, \cdots, t_{(n)})$, here $t_{(i)} (1 \le i \le n)$ are order statistics. Then we have:

$$\begin{aligned}&\sqrt{n}(A_0 + D)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big(\mu_1(t_i) - \mu_0(t_i) - \alpha + \frac{I_i\sigma_1(t_i)\epsilon_{1,t_i}}{p} - \frac{(1-I_i)\sigma_0(t_i)\epsilon_{0,t_i}}{1-p}\Big)\\ &\xrightarrow{d} \mathcal{N}\Big(0, \mathrm{Var}\big(\mu_1(\tau) - \mu_0(\tau)\big) + \mathbb{E}\Big[\frac{\sigma_1^2(\tau)}{p} + \frac{\sigma_0^2(\tau)}{1-p}\Big]\Big).\end{aligned}$$

Here, we use the fact that conditional on $N(T) = n$, $t_i (1 \le i \le n)$ are independent, so under Assumption 3 the Lindeberg-Feller central limit theorem (CLT) holds.

In addition, we have

$$\begin{aligned}&\sqrt{n}(A + D - A_0 - D) = \sqrt{n}(A - A_0)\\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\hat{\mu}_1(t_i) - \hat{\mu}_0(t_i) - \mu_1(t_i) + \mu_0(t_i)\big).\end{aligned}$$

Since μ_1, μ_0 are both Lipschitz continuous, we can take $L_1 > 0$ such that $|\mu_l(s_1) - \mu_l(s_0)| \le L_1|s_1 - s_0|$ for all $s_1, s_0 \in [0, T]$, then

$$|\hat{\mu}_l(t_i) - \mu_l(t_i)| \le \frac{L_1 T}{k(n)}, \ \text{for } l = 0, 1.$$

Thus,

$$\mathbb{E}\big(\sqrt{n}(A + D - A_0 - D)\big)^2 \le \frac{1}{n}\cdot n \cdot \frac{4L_1^2T^2}{k^2(n)} = o(1),$$

that is, $\sqrt{n}(A + D - A_0 - D) = o_p(1)$.

As for B_1, first from Chebyshev's inequality we can see

$$\begin{aligned}&P(\pi_j < n^{\frac{1}{2}} \text{ for some } 1 \le j \le n)\\ &\le \sum_{j=1}^{k(n)} P(n - \pi_j \ge n - n^{\frac{1}{2}})\\ &\le k(n)\frac{\frac{n\Delta_0}{k(n)}(1 - \frac{\Delta_1}{k(n)})}{(\frac{n\Delta_1}{k(n)} - n^{\frac{1}{2}})^2} \to 0 \text{ as } n \to \infty.\end{aligned}$$

Now if $\pi_j \ge n^{\frac{1}{2}}$, by Hoeffding's Inequality it's easy to see $P(p_j - p > \frac{p}{2}$ for some $1 \le j \le n) \to 0$, as $n \to \infty$. Combine above results together, we only need to consider the case that $p_j \le \frac{3p}{2}$ for all j. Rearrange terms and we have:

$$\begin{aligned}B_1 &= \frac{1}{n}\sum_{i=1}^{n}\Big(\frac{I_i\hat{\mu}_1(t_i)}{p} - \hat{\mu}_1(t_i) + \frac{I_i Y_1(t_i)}{\widehat{p}(t_i)} - \frac{I_i Y_1(t_i)}{p}\Big)\\ &= \sum_{j=1}^{k(n)}\frac{\pi_j}{n}\big(1 - \frac{p_j}{p}\big)\Big(\frac{\sum_{i\in S_{j,1}} Y_1(t_i)}{\pi_{j,1}} - m_{j,1}\Big).\end{aligned}$$

It's easy to see there exist constants $c_1, c_2 > 0$ which don't depend on n such that:

$$\mathbb{E}\Big[\Big(\frac{\sum_{i\in S_{j,1}} Y_1(t_i)}{\pi_{j,1}} - m_{j,1}\Big)^2|\pi_{j,1}\Big] \le \frac{c_1}{\pi_{j,1}},$$

$$\mathbb{E}\big[(p - p_j)^2|\pi_{j,1}, \pi_{j,0}\big] \le \frac{c_2}{\pi_j}.$$

Then by Cauchy Inequality we have:

$$\begin{aligned}\mathbb{E}|B_1| &= \mathbb{E}\big[\mathbb{E}[|B_1|\big|\pi_{j,l}(1 \le j \le k(n), l = 0, 1)]\big] \\ &\le \sum_{i=1}^{k(n)} \mathbb{E}\big[\frac{\pi_j}{np}\sqrt{\frac{c_1 c_2}{\pi_j \pi_{j,1}}}\big] \\ &\le \sqrt{\frac{3c_1 c_2}{2p}}\frac{k(n)}{n} = o(n^{\frac{1}{2}}).\end{aligned}$$

Thus, $\sqrt{n}B_1 = o_p(1)$. Similarly, $\sqrt{n}B_0 = o_p(1)$. Finally, for C, again by the Lipschitz continuity of μ_1 and μ_0 we have

$$\mathbb{E}(\sqrt{n}C)^2 \le 2\big(\frac{L_1^2 T^2}{pk^2(n)} + \frac{L_1^2 T^2}{(1-p)k^2(n)}\big) = o(1),$$

thus $\sqrt{n}C = o_p(1)$. Combine all above results together, we obtain:

$$\sqrt{n}(\widehat{\alpha}_{\text{sps}} - \alpha) \xrightarrow{d} \mathcal{N}(0, V_{\text{sps}}), \text{ as } \bar{\lambda} \to +\infty.$$

Here

$$V_{\text{sps}} = \text{Var}\big(\mu_1(\tau) - \mu_0(\tau)\big) + \mathbb{E}\big[\frac{\sigma_1^2(\tau)}{p} + \frac{\sigma_0^2(\tau)}{1-p}\big].$$

In addition, it's easy to see $\frac{n}{\bar{\lambda}} \xrightarrow{p.} 1$, then

$$\sqrt{\bar{\lambda}}(\widehat{\alpha}_{\text{sps}} - \alpha) \xrightarrow{d} \mathcal{N}(0, V_{\text{sps}}), \text{ as } \bar{\lambda} \to +\infty.$$

□

We note that in comparison, the naive standard estimator (1) has asymptotic variance

$$V_{\text{naive}} = \frac{\text{Var}(\mu_1(\tau))}{p} + \frac{\text{Var}(\mu_0(\tau))}{1-p} + \mathbb{E}\Big[\frac{\sigma_1^2(\tau)}{p} + \frac{\sigma_0^2(\tau)}{1-p}\Big],$$

which is larger than V_{sps} by Cauchy Inequality:

$$\big(p + (1-p)\big)\big(\frac{\text{Var}(\mu_1(\tau))}{p} + \frac{\text{Var}(\mu_0(\tau))}{1-p}\big) \ge \text{Var}(\mu_1(\tau) - \mu_0(\tau)).$$

In fact, we have the following proposition which shows that V_{sps} is asymptotically optimal.

Proposition 3.2. *Any semiparametric estimator $\widehat{\alpha}$ for α has following asymptotic variance lower bound:*

$$V \ge V_{optimal} = \text{Var}\big(\mu_1(\tau) - \mu_0(\tau)\big) + \mathbb{E}\Big[\frac{\sigma_1^2(\tau)}{p} + \frac{\sigma_0^2(\tau)}{1-p}\Big].$$

Proof. The proof can be adapted from the proofs for Theorem 1 and Theorem 2 of [5]. □

By Theorem 3.1, we can construct an asymptotically valid confidence interval for statistical inference. For $1 \le j \le k(n)$ and $l = 0, 1$, define

$$\hat{\mu}_{j,l} = \frac{\sum_{i \in S_{j,l}} y_i}{|S_{j,l}|}$$

and

$$\hat{\sigma}_{j,l}^2 = \frac{1}{|S_{j,l}| - 1}\sum_{i \in S_{j,l}}(y_i - \hat{\mu}_{j,l})^2,$$

then an estimator of V_{sps} can be given by:

$$\widehat{s}_{\text{sps}}^2 = \sum_{j=1}^{k(n)} \frac{|S_j|}{n}(\hat{\mu}_{j,1} - \hat{\mu}_{j,0} - \widehat{\alpha}_{\text{naive}})^2 + \sum_{j=1}^{k(n)} \frac{|S_j|}{n}\Big(\frac{\hat{\sigma}_{j,1}^2}{p} + \frac{\hat{\sigma}_{j,0}^2}{1-p}\Big).$$

It is easy to see $\widehat{s}_{\text{sps}}^2$ is a consistent estimator of V_{sps}, then an asymptotic confidence interval at the coverage level of say 90% can be constructed as:

$$[\widehat{\alpha}_{\text{sps}} - \frac{1.65}{\sqrt{n}}\widehat{s}_{\text{sps}}, \widehat{\alpha}_{\text{sps}} + \frac{1.65}{\sqrt{n}}\widehat{s}_{\text{sps}}].$$

Here asymptotic confidence interval means when $n \to \infty$, the coverage rate will converge to 90%.

3.1 Illustrative example on variance reduction

We use a simple example to illustrate the variance reduction performance of the proposed SPS estimator. That is, we compare the asymptotic variance V_{sps} for the SPS estimator and the asymptotic variance V_{naive} for the naive estimator using an illustrative example to show the magnitude of variance reduction. This simple example involves an A/B test that runs over $[0, T]$, following the same setting as in Section 2.2. To facilitate illustrate computation, we presume that the customer arrival process has a constant arrival rate over $[0, T]$. This example has non-stationary treatment/control random outcomes, whose mean and variance structure are given as

$$\mu_1(t) = a_1 + b_1 \cdot t$$
$$\mu_0(t) = a_0 + b_0 \cdot t$$

and

$$\sigma_1(t) = c_1 + d_1 \cdot t$$
$$\sigma_0(t) = c_0 + d_0 \cdot t.$$

We note that such linear structure in the time-dependence of the random outcome is a simple yet representative form of commonly arised non-stationarities, which can be naturally extended to the piecewise linear form. In fact, the non-stationarity pattern demonstrated in Figure 1 may be adequately approximated by piecewise linear patterns at reasonable resolution.

Suppose that $p = \frac{1}{2}$, then for naive estimator, we have

$$\begin{aligned}&V_{\text{naive}} \\ &= 2\text{Var}(\mu_1(\tau)) + 2\text{Var}(\mu_0(\tau)) + 2\mathbb{E}\big[\sigma_1^2(\tau) + \sigma_0^2(\tau)\big] \\ &= (b_1^2 + b_0^2)\frac{T^2}{6} + 2\big(c_1^2 + c_0^2 + (c_1 d_1 + c_0 d_0)T + (d_1^2 + d_0^2)\frac{T^2}{3}\big),\end{aligned}$$

and

$$\begin{aligned}V_{\text{sps}} &= \text{Var}\big(\mu_1(\tau) - \mu_0(\tau)\big) + 2\mathbb{E}\big[\sigma_1^2(\tau) + \sigma_0^2(\tau)\big] \\ &= (b_1 - b_0)^2\frac{T^2}{12} + 2\big(c_1^2 + c_0^2 + (c_1 d_1 + c_0 d_0)T + (d_1^2 + d_0^2)\frac{T^2}{3}\big).\end{aligned}$$

It is evident that $V_{\text{sps}} \le V_{\text{naive}}$. The exact magnitude of variance reduction depends on all the parameters, but in particular, when $\sigma_1(t), \sigma_0(t)$ are relatively small compared with $\mu_1(t), \mu_0(t)$, we have

$$\frac{V_{\text{sps}}}{V_{\text{naive}}} \approx \frac{(b_1 - b_0)^2 \frac{T^2}{12}}{(b_1^2 + b_0^2)\frac{T^2}{6}} = \frac{(b_1 - b_0)^2}{2(b_1^2 + b_0^2)}.$$

This ratio can be arbitrarily close to 0 when $b_1 - b_0$ is close to zero. The implication is that when the treatment and control has very close non-stationary pattern, adequately addressing the non-stationarity by viewing time as a continuous strata can bring significant variance reduction.

4 DE-BIASED POST STRATIFICATION WITH CONTINUOUS COVARIATE

In this section, we consider a more general setting (compared to Section 3) where the treatment assignment probability may have been time-varying in the experiment. That is, during the time horizon $[0, T]$ when the A/B test experiment has been run, the probability $p(t)$ of assigning an arriving customer to treatment may be non-stationary and have changed over time. Such non-stationarity can be caused by a commonly used ramping-up procedure where $p(t)$ starts off to be small and gradually increases. Such ramping-up procedure is usually implemented in the experiment when one hopes to reduce the downside risk that the treatment may be too bad. Such non-stationarity can also be caused by more sophisticated reinforcement learning procedures that are implemented to dynamically adjust the treatment assignment probability $p(t)$.

Such an A/B test generates experiment data $\{I_i, y_i, t_i\}_{i=1,2,\dots,n}$, as defined in (3). The goal is, similar to Section 3, using the data to estimate the average treatment effect

$$\alpha = \mathbb{E}(Y_1(\tau) - Y_0(\tau)) = \int_0^T (\mu_1(t) - \mu_0(t)) f(t)\, dt.$$

In presence of the non-stationarity in the treatment assignment probability $p(t)$, the naive estimator as defined in (1) is no longer guaranteed to be unbiased. In fact, the expectation of the naive estimator is given by

$$\mathbb{E}\widehat{\alpha}_{\text{naive}} = \int_0^T \Big(\frac{\mu_1(t)p(t)}{\bar{p}} - \frac{\mu_0(t)(1-p(t))}{1-\bar{p}}\Big) f(t) dt,$$

where $\bar{p} = \frac{1}{T}\int_0^T p(t)dt$. It is evident that

$$\mathbb{E}\widehat{\alpha}_{\text{naive}} \neq \alpha$$

in general, and the non-stationarity in $p(t)$, $\mu_1(t)$ and $\mu_0(t)$ combined may even make the bias $\mathbb{E}\widehat{\alpha}_{\text{naive}} - \alpha$ so large such that the sign of $\mathbb{E}\widehat{\alpha}_{\text{naive}}$ is different from the sign of α.

We next develop an estimator that we refer to as *de-biased post stratification with continuous covariate* (DPS) to address this bias issue. This estimator not only accommodates the non-stationary $p(t)$, but also accommodates another layer of non-stationarity and flexibility in the data generation process. That is, if we denote the arrival rate of customers assigned to this A/B test experiment as $\gamma(t)$, this $\gamma(t)$ is allowed by the DPS estimator to be different from the true underlying arrival rate $\lambda(t)$ to the platform. This flexibility can help the estimator cover broader classes of A/B tests. For example, it may be that an A/B test starts to use 5% of the entire traffic to the platform but after two days of experiment, the system decides to add additional 3% of traffic to the experiment. Such scenarios will make $\gamma(t)$, the arrival rate of customers assigned to the A/B test over the time horizon $[0, T]$ different from the true underlying arrival rate $\lambda(t)$ to the platform. In this situation, the estimator allows flexibility in $\gamma(t)$ and has knowledge about the underlying arrival rate $\lambda(t)$ and therefore $f(t)$ to the platform.

The DPS estimator $\widehat{\alpha}_{\text{dps}}$ is constructed as follows. We adopt the similar notation in Section 3, because the DPS estimator is closely connected to though different from the SPS estimator $\widehat{\alpha}_{\text{sps}}$. The DPS estimator writes as

$$\widehat{\alpha}_{\text{dps}} = \sum_{j=1}^{k(n)} q_j \Big(\frac{\sum_{i \in S_{j,1}} y_i}{|S_{j,1}|} - \frac{\sum_{i \in S_{j,0}} y_i}{|S_{j,0}|} \Big),$$

here

$$q_j = \int_{\frac{j-1}{k(n)}T}^{\frac{j}{k(n)}T} f(t) dt.$$

The key difference between this DPS estimator and SPS estimator is that, we replace the empirical probability of each stratum $\frac{|S_{j,1}|}{|S_j|}$ in the SPS estimator by the probability q_j that reflects the true underlying arrival rate to the platform.

To provide rigorous statistical inference theory for the DSP estimator, we need the following assumption:

Assumption 4. *There exist* $0 < M_1 < M_2$ *and* $\Delta_3 > 0$ *such that* $M_1 < \frac{\gamma(t)}{\bar{\gamma}} < M_2$ *and* $\Delta_3 < p(t) < 1 - \Delta_3$*, where* $\bar{\gamma} = \int_0^T \gamma(t)\, dt$.

This assumption is almost always met in practice, in the sense that the assumption requires that there is always a positive probability assigning a customer to treatment during the time horizon $[0, T]$ and there is always a positive portion of traffic assigned to the experiment during its running time $[0, T]$. We have the following result to first ensure that the DPS estimator is asymptotically unbiased.

Theorem 4.1. *Under Assumptions 1, 2 and 4, if* $\lim_{n\to\infty} k(n) = +\infty$ *and* $k(n) = o(n)$*, we have:*

$$\widehat{\alpha}_{dps} \xrightarrow{p.} \alpha, \text{ as } \bar{\gamma} \to \infty.$$

Proof. See Appendix. □

5 TIME-GROUPED RANDOMIZATION FOR EXPERIMENT DESIGN

In the previous two sections - Section 3 and Section 4, we have proposed estimators and proved statistical inference results to enhance post-experiment data analysis for non-stationary A/B tests. If in addition, one has the ability to modify the experiment design of the non-stationary A/B tests, we propose a time-grouped randomization approach that can be integrated with standard experiment design to improve efficiency.

We introduce the time-grouped randomization idea based on a simple A/B test experiment design setting where each arriving customer is randomly assigned to treatment with probability $1/2$ and to control with probability $1/2$. In this simple setting, if we denote I_i as the label of treatment/control for the i-th customer that arrives at the test experiment over $[0, T]$, then I_i is a Bernoulli random variable that takes value 1 (treatment) with probability $1/2$ and takes value 0 (control) with probability $1/2$. The sequence $I_1, I_2, \cdots$ are then i.i.d. Bernoulli random variables in the simple setting. Instead, the time-grouped randomization works as follows. First, set $J_1, J_2, \cdots$ to be a sequence of i.i.d Bernoulli random variables that take value 1 (treatment) with probability $1/2$ and takes value 0 (control) with probability $1/2$. Then we set

$$I_1 = J_1, I_2 = 1 - J_1, I_3 = J_2, I_4 = 1 - J_2, \cdots,$$

or summarized as $I_{2i-1} = J_i$ and $I_{2i} = 1 - J_i$ for $i = 1, 2, \cdots$. In this way, every two consecutive arrival customers are grouped into a pair. Within each pair, one customer is assigned to treatment chosen at random, and the other is assigned to control.

Compared to the standard i.i.d. Bernoulli randomization, such time-grouped randomization regularizes the treatment/control assignments to be less varying and more balanced within any given time window. On the other hand, the time-grouped randomization maintains the randomness at the customer level, in the sense that, a priori given any customer, the probability of that customer being assigned to treatment remains 1/2.

This time-grouped randomization can naturally be extended to more general treatment assignment probability p, as long as p is a rational number. In fact, many A/B tests run in practice has the range of p to be $\{0\%, 1\%, 2\%, \cdots, 99\%, 100\%\}$, at the resolution of 1%. For generality, we presume that p is always a rational number $\frac{u}{w}$, where $u, w \in \mathbb{N}_+$ and the greatest common divisor $(u, w) = 1$. The time-grouped randomization for the general choice of p is defined as follows: first, define $\mathcal{A}_i = \{x \in \mathbb{N} : (i-1)w + 1 \leq x < iw\}$ for $i \in \mathbb{N}_+$. For each $\mathcal{A}_i$, randomly pick up a subset $\mathcal{B}_i$ from all the u-element subsets of $\mathcal{A}_i$ with equal probability. Then, for the j-th arriving customer, we assign this customer to treatment if $j \in \mathcal{B}_i$ for some i, otherwise we assign the customer to control. This procedure is implemented for all customers that arrive at the A/B test experiment over $[0, T]$. In this section, we adopt the same data generation process as in Section 3.

Now, we are ready to construct the estimator $\widehat{\alpha}_{\text{tr}}$ based on data generated from the time-grouped randomization. Suppose we have n samples represented by

$$\{I_i, y_i, t_i\}_{i=1,2\cdots,n},$$

and in this section we assume $t_1 < t_2 < \cdots < t_n$. We first write $n = qw + r$ for some $q \in \mathbb{N}$ and $r \in \{0, 1, \cdots, w-1\}$. Define $\mathcal{B} = \cup_{i=1}^{q} \mathcal{B}_i$ and $\mathcal{A} = \cup_{i=1}^{q} \mathcal{A}_i$, then the estimator is constructed as follows:

$$\widehat{\alpha}_{\text{tr}} = \sum_{i \in \mathcal{B}} \frac{y_i}{|\mathcal{B}|} - \sum_{i \in \mathcal{A} \setminus \mathcal{B}} \frac{y_i}{|\mathcal{A} \setminus \mathcal{B}|}.$$

The estimator $\widehat{\alpha}_{\text{tr}}$ takes the same form as the naive estimator $\widehat{\alpha}_{\text{naive}}$. However, because the experiment design now includes the time-grouped randomization, the asymptotic variance of $\widehat{\alpha}_{\text{tr}}$ is smaller than and can sometimes be much smaller than the asymptotic variance V_{naive}. We formally prove a central limit theorem for the estimator $\widehat{\alpha}_{\text{tr}}$ under the time-grouped randomization experiment design.

Assumption 5. *$\mu_1(t), \mu_0(t), \sigma_1(t), \sigma_0(t)$ are all Lipschitz continuous functions. In addition, $\sigma_1(t), \sigma_0(t) > \Delta_2$ for some $\Delta_2 > 0$ and all $t \in [0, T]$.*

Assumption 6. *For all $t \in [0, T]$, $\epsilon_{1,t}$ are i.i.d. random variables, and $\epsilon_{0,t}$ are also i.i.d. random variables.*

Theorem 5.1. *Suppose Assumption 1, 5 and 6 are satisfied, under the time-grouped randomization design, we have*

$$\sqrt{\bar{\lambda}}(\widehat{\alpha}_{tr} - \alpha) \xrightarrow{d} \mathcal{N}(0, V_{tr}),$$

here

$$V_{tr} = \mathbb{E}\Big(\frac{\sigma_1^2(\tau)}{p} + \frac{\sigma_0^2(\tau)}{1-p}\Big) + \text{Var}(\mu_1(\tau) - \mu_0(\tau)).$$

Figure 2: The calibrated mean and standard deviation

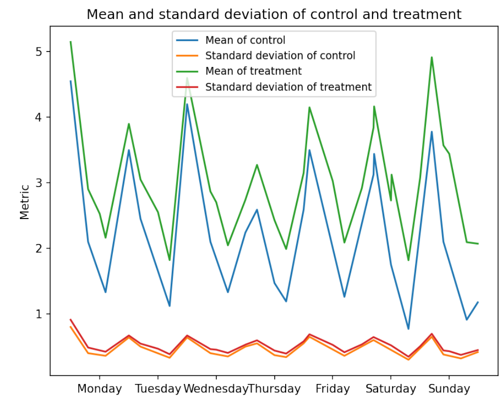
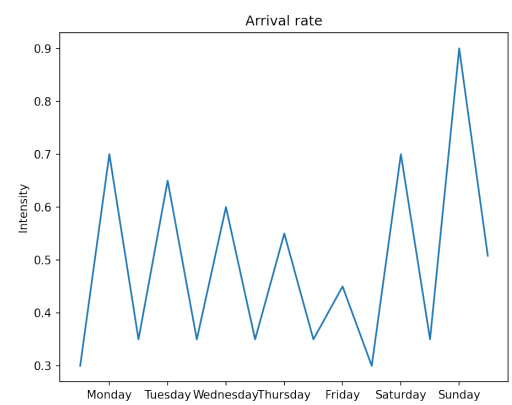


Figure 3: Arrival rate of the non-homogeneous Poisson process



Proof. See Appendix. □

6 NUMERICAL EXPERIMENTS

We validate the performances of the proposed SPS and DPS estimators and the time-grouped randomization on simulated data. The data generation process for the A/B test experiment is calibrated from real data from a leading e-commerce company. Figure 2 plots the non-stationary mean and variance of the random outcome model for treatment and control. Figure 3 plots the non-stationary customer arrival rate. The magnitudes in the two plots are deliberately scaled up to a constant multiplier compared to the real data, for data privacy reasons. For the numerical experiments, we validate the proposed estimators and randomization scheme for a variety of sample sizes through convergence rates and mean square errors, which is a reason that we calibrate a data generation process from the real data versus directly using the real data. Thus we are able to construct numerous parallel scenarios to validate

$\bar{\lambda}$	$\hat{\alpha}_{\text{naive}}$	$\hat{\alpha}_{\text{psf}}$	ratio	$\hat{\alpha}_{\text{sps}}$	ratio	$\hat{\alpha}_{\text{tr}}$	ratio
250	**1.18E-2**	1.18E-2	1.00	7.20E-3	0.61	**4.56E-3**	**0.39**
500	**6.26E-3**	6.25E-3	1.00	2.83E-3	0.45	**2.08E-3**	**0.33**
1000	**2.88E-3**	2.85E-3	0.99	1.28E-3	0.44	**1.04E-3**	**0.36**
2000	**1.51E-3**	1.50E-3	0.99	5.59E-4	0.37	**5.21E-4**	**0.35**
4000	**7.33E-4**	7.14E-4	0.97	2.74E-4	0.37	**2.64E-4**	**0.36**
8000	**3.84E-4**	3.79E-4	0.99	1.36E-4	0.35	**1.27E-4**	**0.33**
16000	**1.88E-4**	1.85E-4	0.98	6.56E-5	0.35	**6.44E-5**	**0.34**
32000	**9.01E-5**	8.86E-5	0.98	3.31E-5	0.37	**3.06E-5**	**0.34**
64000	**4.99E-5**	4.83E-5	0.97	1.66E-5	0.33	**1.60E-5**	**0.32**

Table 1: MSE of $\hat{\alpha}_{\text{naive}}$, $\hat{\alpha}_{\text{psf}}$, $\hat{\alpha}_{\text{sps}}$ and $\hat{\alpha}_{\text{tr}}$

$\bar{\lambda}$	Bias ($\hat{\alpha}_{\text{naive}}$)	MSE($\hat{\alpha}_{\text{naive}}$)	Bias ($\hat{\alpha}_{\text{dps}}$)	MSE ($\hat{\alpha}_{\text{dps}}$)	Var ($\hat{\alpha}_{\text{dps}}$)
250	8.36E-02	2.11E-02	6.02E-02	1.58E-02	1.21E-02
500	1.01E-01	1.90E-02	3.94E-02	5.66E-03	4.11E-03
1000	9.53E-02	1.36E-02	8.75E-03	2.59E-03	2.52E-03
2000	1.01E-01	1.22E0-2	3.27E-03	1.01E-03	9.98E-04
4000	1.02E-01	1.12E-02	1.48E-03	3.85E-04	3.83E-04
8000	1.01E-01	1.07E-02	2.33E-03	2.54E-04	2.49E-04
16000	1.02E-01	1.06E-02	1.03E-03	9.16E-05	9.06E-05
32000	9.97E-02	1.01E-02	2.17E-03	5.15E-05	5.15E-05
64000	9.89E-02	9.84E-03	4.76E-04	2.12E-05	2.09E-05

Table 2: MSE, bias and variance of $\hat{\alpha}_{\text{naive}}$ and $\hat{\alpha}_{\text{dps}}$

the convergence rate and mean square errors, while if we use only real data we would only be able to present a magnitude of variance reduction percentage without further validation on inference.

In our experiments, for the SPS estimator introduced in Section 3, we take $k(n) = 7[\frac{3n^{0.4}}{7}]$, so that $k(n) = o(n^{\frac{1}{2}})$, $\lim_{n\to\infty} k(n) = +\infty$ and $k(n)$ is a multiple of 7. We repeat each experiment $T = 100000$ times and calculate the empirical mean square error (MSE) for different values of aggregated arrival rate $\bar{\lambda}$ over [0,T] where T is fixed to be 7. We first compare $\widehat{\alpha}_{\text{naive}}$, $\widehat{\alpha}_{\text{psf}}$, $\widehat{\alpha}_{\text{sps}}$, $\widehat{\alpha}_{\text{tr}}$ under the situation that $p = \frac{1}{2}$. Here $\widehat{\alpha}_{\text{psf}}$ stands for the post stratification estimator with fixed strata number $k(n) = 7$, representing a standard practice that the day-of-week effect is considered. The results are illustrated by Table 1. Here, “ratio” in Table 1 means the ratio of the estimator’s MSE to the MSE of the naive estimator $\widehat{\alpha}_{\text{naive}}$.

From Table 1, we summarize that:

(1) Post-stratification with fixed strata number $k(n) = 7$ reduces only 1% to 3% of the MSE, which means if we only consider the day-of-week effect, the variance reduction can be on the weak side.
(2) Both $\hat{\alpha}_{\text{sps}}$ and $\hat{\alpha}_{\text{tr}}$ achieve a much better performance compared to the fixed strata estimator and the naive estimator, especially when $\bar{\lambda}$ is large. For $\bar{\lambda} \geq 2000$, $\hat{\alpha}_{\text{sps}}$ reduces more than 60% of the MSE. Similarly, $\hat{\alpha}_{\text{tr}}$ always reduces more than 60% of the MSE.

Though both $\hat{\alpha}_{\text{sps}}$ and $\hat{\alpha}_{\text{tr}}$ perform well in the experiments, we suggest to use $\hat{\alpha}_{\text{tr}}$ when it is possible to use time-grouped randomization, because its performance is more stable when the arrival rate is relatively small and its finite sample variance is smaller than $\hat{\alpha}_{\text{sps}}$. In practice, if the arrival rate is large, both estimators $\hat{\alpha}_{\text{sps}}$ and $\hat{\alpha}_{\text{tr}}$ work comparably well.

To illustrate the finite-sample performance for the debiased post-stratification (DPS) estimator $\widehat{\alpha}_{\text{dbs}}$ introduced in Section 4, we use the same data generation process for the non-stationary random outcomes, and we again take $k(n) = 7[\frac{3n^{0.4}}{7}]$. However, this time the treatment assignment probability $p(t)$ is changing over time from 10% t0 50% as follows $p(t) = 0.1 + \frac{2[2t]}{65}, t \in [0, 7)$. Note that in this case, the naive estimator and any post-stratification estimator with fixed number of strata may incur bias that can not be eliminated however large the sample size is.

As we can see in Table 2, in this case, $\hat{\alpha}_{\text{naive}}$ always incurs a bias about 0.1. This bias does not decrease when we simply increase $\bar{\lambda}$, i.e., increase the sample size, and the MSE of $\hat{\alpha}_{\text{naive}}$ are mainly contributed by the bias term. In contrast, the bias of $\hat{\alpha}_{\text{dps}}$ decreases effectively when $\bar{\lambda}$ increases. When $\bar{\lambda} \geq 1000$, the bias term can be neglected compared with variance.

7 ACKNOWLEDGMENTS

The authors are thankful to all four anonymous reviewers and one meta-reviewer for their valuable comments. The authors thank Junzi Zhang and Hongyu Zhu for helpful discussions.

REFERENCES

[1] Søren Asmussen and Peter W Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. Vol. 57. Springer Science & Business Media.
[2] Peter L Cohen and Colin B Fogarty. 2020. No-harm calibration for generalized oaxaca-blinder estimators. *arXiv preprint arXiv:2012.09246* (2020).
[3] Alex Deng, Ya Xu, Ron Kohavi, and Toby Walker. 2013. Improving the sensitivity of online controlled experiments by utilizing pre-experiment data. In *Proceedings of the sixth ACM international conference on Web search and data mining*. 123–132.
[4] Somit Gupta, Ronny Kohavi, Diane Tang, Ya Xu, Reid Andersen, Eytan Bakshy, Niall Cardin, Sumita Chandran, Nanyu Chen, Dominic Coey, et al. 2019. Top challenges from the first practical online controlled experiments summit. *ACM SIGKDD Explorations Newsletter* 21, 1 (2019), 20–35.
[5] Jinyong Hahn. 1998. On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica* (1998), 315–331.
[6] Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. 2021. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics* 49, 2 (2021), 1055–1080.
[7] Ramesh Johari, Pete Koomen, Leonid Pekelis, and David Walsh. 2017. Peeking at A/B tests: Why it matters, and what to do about it. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 1517–1525.
[8] Ramesh Johari, Pete Koomen, Leonid Pekelis, and David Walsh. 2021. Always valid inference: Continuous monitoring of A/B tests. *Operations Research* (2021).
[9] Ron Kohavi, Alex Deng, Brian Frasca, Toby Walker, Ya Xu, and Nils Pohlmann. 2013. Online controlled experiments at large scale. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1168–1176.
[10] Ron Kohavi and Roger Longbotham. 2017. Online Controlled Experiments and A/B Testing. *Encyclopedia of machine learning and data mining* 7, 8 (2017), 922–929.
[11] Ron Kohavi, Diane Tang, and Ya Xu. 2020. *Trustworthy online controlled experiments: A practical guide to A/B testing*. Cambridge University Press.
[12] Matt Taddy, Hedibert Freitas Lopes, and Matt Gardner. 2016. Scalable semiparametric inference for the means of heavy-tailed distributions. *arXiv preprint arXiv:1602.08066* (2016).
[13] Diane Tang, Ashish Agarwal, Deirdre O’Brien, and Mike Meyer. 2010. Overlapping experiment infrastructure: More, better, faster experimentation. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. 17–26.
[14] Edward Wu and Johann A Gagnon-Bartsch. 2018. The loop estimator: Adjusting for covariates in randomized experiments. *Evaluation review* 42, 4 (2018), 458–488.
[15] Huizhi Xie and Juliette Aurisset. 2016. Improving the sensitivity of online controlled experiments: Case studies at netflix. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 645–654.

A APPENDIX

A.1 Proof of Theorem 4.1

$$\begin{aligned}
&\widehat{\alpha}_{\rm ps}-\alpha\\
&=\sum_{j=1}^{k(n)} q_j\Big(\frac{\sum_{i\in S_{j,1}}\big(\mu_1(t_i)+\sigma_1(t_i)\epsilon_{1,t_i}\big)}{|S_{j,1}|}-\frac{\sum_{i\in S_{j,0}}\big(\mu_0(t_i)+\sigma_0(t_i)\epsilon_{0,t_i}\big)}{|S_{j,0}|}-\alpha\Big)\\
&=\sum_{j=1}^{k(n)} q_j\Big(\frac{\sum_{i\in S_{j,1}}\mu_1(t_i)}{|S_{j,1}|}-\frac{\sum_{i\in S_{j,0}}\mu_0(t_i)}{|S_{j,0}|}-\frac{\int_{\frac{j-1}{k(n)}T}^{\frac{j}{k(n)}T}(\mu_1(t)-\mu_0(t))f(t)dt}{q_j}\Big)\\
&\quad+\sum_{j=1}^{k(n)} q_j\Big(\frac{\sum_{i\in S_{j,1}}\sigma_1(t_i)\epsilon_{1,t_i}}{|S_{j,1}|}-\frac{\sum_{i\in S_{j,0}}\sigma_0(t_i)\epsilon_{0,t_i}}{|S_{j,0}|}\Big).
\end{aligned}$$

For the first summation, since μ_1,μ_0 are both Lipschitz continuous, there exists $L_1>0$ such that $|\mu_l(s_1)-\mu_0(s_0)|\le L_1|s_1-s_0|$ for all $s_1,s_0\in[0,1]$, thus for $l=0,1$,

$$\max_{t\in[\frac{j-1}{k(n)}T,\frac{j}{k(n)}T]}\Big|\frac{\sum_{i\in S_{j,1}}\mu_1(t_i)}{|S_{j,1}|}-\mu_l(t)\Big|\le\frac{L_1T}{k(n)}.$$

Plug in and we have:

$$\Big|\frac{\sum_{i\in S_{j,1}}\mu_1(t_i)}{|S_{j,1}|}-\frac{\sum_{i\in S_{j,0}}\mu_0(t_i)}{|S_{j,0}|}-\frac{\int_{\frac{j-1}{k(n)}T}^{\frac{j}{k(n)}T}(\mu_1(t)-\mu_0(t))f(t)dt}{q_j}\Big|\le\frac{2L_1T}{k(n)},$$

then

$$\begin{aligned}
&\Big|\sum_{j=1}^{k(n)}q_j\Big(\frac{\sum_{i\in S_{j,1}}\mu_1(t_i)}{|S_{j,1}|}-\frac{\sum_{i\in S_{j,0}}\mu_0(t_i)}{|S_{j,0}|}\\
&-\frac{\int_{\frac{j-1}{k(n)}T}^{\frac{j}{k(n)}T}(\mu_1(t)-\mu_0(t))f(t)dt}{q_j}\Big)\Big|\le\frac{2L_1T}{k(n)}=o(1).
\end{aligned}$$

For the second summation, since $\epsilon_{l,t}$ are all mean zero and independent, we have

$$\begin{aligned}
&\mathrm{Var}\sum_{j=1}^{k(n)}q_j\Big(\frac{\sum_{i\in S_{j,1}}\sigma_1(t_i)\epsilon_{1,t_i}}{|S_{j,1}|}-\frac{\sum_{i\in S_{j,0}}\sigma_0(t_i)\epsilon_{0,t_i}}{|S_{j,0}|}\Big)\\
&=\sum_{j=1}^{k(n)}q_j^2\mathbb{E}\Big(\frac{\sum_{i\in S_{j,1}}\sigma_1^2(t_i)\epsilon_{1,t_i}^2}{|S_{j,1}|^2}+\frac{\sum_{i\in S_{j,0}}\sigma_0^2(t_i)\epsilon_{0,t_i}^2}{|S_{j,0}|^2}\Big).
\end{aligned}$$

Since σ_1,σ_0 are both continuous functions on $[0,T]$, we have

$$\sigma_{\max}\triangleq\max_{t\in[0,T],l=0,1}|\sigma_l(t)|<+\infty.$$

In addition, note that $f(t)\le\Delta_1$, then $\sup_{1\le j\le k(n)}q_j\to0$ as $n\to0$, thus

$$\begin{aligned}
&\mathrm{Var}\sum_{j=1}^{k(n)}q_j\Big(\frac{\sum_{i\in S_{j,1}}\sigma_1(t_i)\epsilon_{1,t_i}}{|S_{j,1}|}-\frac{\sum_{i\in S_{j,0}}\sigma_0(t_i)\epsilon_{0,t_i}}{|S_{j,0}|}\Big)\\
&\le\sum_{j=1}^{k(n)}q_j^2\sigma_{\max}^2\Big(\frac{1}{|S_{j,1}|}+\frac{1}{|S_{j,0}|}\Big)\\
&\le\Big(\sup_{1\le j\le k(n)}q_j\Big)2\sigma_{\max}^2=o(1)
\end{aligned}$$

Here, since $k(n)=o(n)$, similar as we did before we know $p(|S_{j,l}|=0\text{ for some }j,l)\to0$ as $n\to\infty$, so we only consider the case $|S_{j,l}|\ge1$ for all $1\le j\le k(n)$ and $l=0,1$. Thus,

$$\sum_{j=1}^{k(n)}q_j\Big(\frac{\sum_{i\in S_{j,1}}\sigma_1(t_i)\epsilon_{1,t_i}}{|S_{j,1}|}-\frac{\sum_{i\in S_{j,0}}\sigma_0(t_i)\epsilon_{0,t_i}}{|S_{j,0}|}\Big)=o_p(1),$$

and $\widehat{\alpha}_{\rm ps}-\alpha=o_p(1)$. Finally, it's easy to see we can replace $n\to\infty$ by $\bar{\gamma}\to\infty$, which then gives the desired result.

A.2 Proof of Theorem 5.1

For any $x>0$, define $M=\lceil\frac{2}{x}\rceil$, and divide $[0,T]$ to M intervals: $[0,\frac{1}{M}],[\frac{1}{M},\frac{2}{M}],\cdots,[\frac{M-1}{M},1]$. If $\max_{1\le i\le n-1}|t_i-t_{i+1}|>x$, then there exists $j\in\{1,\cdots,M\}$ such that for any $1\le i\le n$, $t_i\notin[\frac{j-1}{M},\frac{j}{M}]$. Thus, $P(\max_{1\le i\le n-1}|t_i-t_{i+1}|>x)\le M(1-\frac{\Delta_0}{M})^n$. If $x=O(\frac{1}{n^{1-\varepsilon}})$ for some $\varepsilon>0$, we have $M=O(n^{1-\varepsilon})$, then $M(1-\frac{\Delta_0}{M})^n\to0$ as $n\to\infty$. Since we are only interested in the asymptotic distribution, from above results we only need to consider the case that $\max_{1\le i\le n-1}|t_i-t_{i+1}|\le\frac{1}{n^{1-\varepsilon}}$ for some $\varepsilon>0$ thereafter. That is, following discussions are all conditioned on $\max_{1\le i\le n-1}|t_i-t_{i+1}|\le\frac{1}{n^{1-\varepsilon}}$.

Now we can write $\widehat{\alpha}_{\rm tr}=A+B-C$, here we define:

$$A=\sum_{i\in\mathcal{B}}\frac{\mu_1(t_i)}{|\mathcal{B}|}-\sum_{i\in\mathcal{A}\setminus\mathcal{B}}\frac{\mu_0(t_i)}{|\mathcal{A}\setminus\mathcal{B}|}+\frac{1}{qw}\sum_{i=qw+1}^{n}(\mu_1(t_i)-\mu_0(t_i)),$$

$$B=\sum_{i\in\mathcal{B}}\frac{\sigma_1(t_i)\epsilon_{1,t_i}}{|\mathcal{B}|}-\sum_{i\in\mathcal{A}\setminus\mathcal{B}}\frac{\sigma_0(t_i)\epsilon_{0,t_i}}{|\mathcal{A}\setminus\mathcal{B}|}+\frac{1}{qw}\sum_{i=qw+1}^{n}\Big(\frac{\sigma_1(t_i)\epsilon_{1,t_i}}{p}-\frac{\sigma_0\epsilon_{0,t_i}}{1-p}\Big),$$

$$C=\frac{1}{qw}\sum_{i=qw+1}^{n}(\mu_1(t_i)-\mu_0(t_i))+\frac{1}{qw}\sum_{i=qw+1}^{n}\Big(\frac{\sigma_1(t_i)\epsilon_{1,t_i}}{p}-\frac{\sigma_0\epsilon_{0,t_i}}{1-p}\Big).$$

Since the summation in C contains at most w terms, then it's easy to see $\sqrt{n}C=o_p(1)$, which means C can be neglected in the asymptotic distribution. For A, we consider the summation of each group $\mathcal{A}_j$. For $j=1,\cdots,q$, we have

$$\begin{aligned}
&\frac{1}{q}\sum_{i\in\mathcal{B}_j}\frac{\mu_1(t_i)}{|\mathcal{B}_j|}-\frac{1}{q}\sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\mu_0(t_i)}{|\mathcal{A}_j\setminus\mathcal{B}_j|}-\frac{1}{qw}\sum_{i\in\mathcal{A}_j}(\mu_1(t_i)-\mu_0(t_i))\\
&=\frac{1}{q}\Big(\sum_{i\in\mathcal{B}_j}\Big(\frac{1}{u}-\frac{1}{w}\Big)\mu_1(t_i)-\sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{1}{w}\mu_1(t_i)\Big)\\
&-\frac{1}{q}\Big(\sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\Big(\frac{1}{u-w}-\frac{1}{w}\Big)\mu_0(t_i)-\sum_{i\in\mathcal{B}_j}\frac{1}{w}\mu_0(t_i)\Big).
\end{aligned}$$

Assumption 5 shows that $\mu_1(t),\mu_0(t)$ are both Lipschitz continuous, and in addition recall that we have $\max_{1\le i\le n-1}|t_i-t_{i+1}|\le\frac{1}{n^{1-\varepsilon}}$, thus there exists constant K_1 doesn't depend on q such that

$$\begin{aligned}
&\Big|\frac{1}{q}\sum_{i\in\mathcal{B}_j}\frac{\mu_1(t_i)}{|\mathcal{B}_j|}-\frac{1}{q}\sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\mu_0(t_i)}{|\mathcal{A}_j\setminus\mathcal{B}_j|}-\frac{1}{qw}\sum_{i\in\mathcal{A}_j}(\mu_1(t_i)-\mu_0(t_i))\Big|\\
&\le\frac{K_1}{qn^{1-\varepsilon}},
\end{aligned}$$

thus, as long as $\varepsilon < \frac{1}{2}$, we have

$$\sqrt{n}\Big|A - \frac{1}{qw}\sum_{i=1}^{n}(\mu_1(t_i) - \mu_0(t_i))\Big| \le \sum_{i=1}^{q}\frac{K_1}{qn^{\frac{1}{2}-\varepsilon}} = o(1),$$

which implies

$$A = \frac{1}{qw}\sum_{i=1}^{n}(\mu_1(t_i) - \mu_0(t_i)) + o_p(\frac{1}{\sqrt{n}}).$$

Now, as for B, for any $j \in \{1, \cdots, q\}$, from Assumption 5, σ_1, σ_0 are Lipschitz, combine this with $\max_{1\le i\le n-1}|t_i - t_{i+1}| \le \frac{1}{n^{1-\varepsilon}}$ we known there exists $K_2 > 0$ which doesn't depend on q such that

$$\sum_{i\in\mathcal{B}_j}\frac{\sigma_1(t_i)\epsilon_{1,t_i}}{|\mathcal{B}_j|} - \sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\sigma_0(t_i)\epsilon_{0,t_i}}{|\mathcal{A}_j\setminus\mathcal{B}_j|}$$
$$= \sum_{i\in\mathcal{B}_j}\frac{\sigma_1(t_{(j-1)w+1})\epsilon_{1,t_i}}{u} - \sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\sigma_0(t_{(j-1)w+1})\epsilon_{0,t_i}}{w-u} + o_p(\frac{1}{n^{1-\varepsilon}}).$$

For further calculation, we need following lemma:

Lemma 1. *For $q, w \in \mathbb{N}_+$ and $r \in \{0, 1\cdots, w-1\}$, define $\mathcal{S} = \{1, 2, \cdots, qw + r\}$. If we randomly choose a q-element subset $\mathcal{T}$ from all the q-element subsets of $\mathcal{S}$ with equal probability, and write the elements in $\mathcal{T}$ as $x_1 < x_2 < \cdots < x_q$, then*

$$\mathbb{E}\Big[\sum_{i=1}^{q}\big(x_i - (i-1)w - 1\big)^2\Big] = O(q^2), \text{ as } q \to \infty.$$

Proof of Lemma 1:

$$\mathbb{E}\sum_{i=1}^{q}\big(x_i - (i-1)w - 1\big)^2$$
$$=\mathbb{E}\Big[\sum_{i=1}^{q}x_i^2\Big] + \sum_{i=1}^{q}[(i-1)w+1]^2 - 2\sum_{i=1}^{q}\mathbb{E}x_i((i-1)w+1)$$
$$=\frac{2w^2q^3}{3} + O(q^2) - 2\sum_{i=1}^{q}\mathbb{E}x_i((i-1)w+1).$$

As for $\mathbb{E}x_i$, by direct calculation we have:

$$\mathbb{E}x_i = \frac{\sum_{k=i}^{qw+r-q+i}kC_{k-1}^{i-1}C_{qw+r-k}^{q-i}}{C_{qw+r}^{q}}$$
$$= \frac{i\sum_{k=i}^{qw+r-q+i}C_k^iC_{qw+r-k}^{q-i}}{C_{qw+r}^{q}} = \frac{iC_{qw+r+1}^{q+1}}{C_{qw+r}^{q}} = \frac{i(qw+r+1)}{q+1}.$$

Plug in and we immediately obtain

$$2\sum_{i=1}^{q}\mathbb{E}x_i((i-1)w+1) = \frac{2w^2q^3}{3} + O(q^2),$$

which finishes the proof.

Now, come back to the proof of Theorem 5.1. If we randomly pick a q-element subset $\mathcal{T} = \{x_1, \cdots, x_q\}$ of $\mathcal{S} = \{1, \cdots, qw + r\}$, then

$$B = \frac{1}{q}\sum_{j=1}^{q}\Bigg(\sigma_1(t_{x_j})\sum_{i\in\mathcal{B}_j}\frac{\epsilon_{1,t_i}}{u} - \sigma_0(t_{x_j})\sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\epsilon_{0,t_i}}{w-u}\Bigg) + D + o_p(\frac{1}{n^{1-\varepsilon}}).$$

Here,

$$D = \frac{1}{q}\sum_{j=1}^{q}\Bigg(\sum_{i\in\mathcal{B}_j}\frac{\big(\sigma_1(t_{(j-1)w+1}) - \sigma_1(t_{x_j})\big)\epsilon_{1,t_i}}{u} - \sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\big(\sigma_0(t_{(j-1)w+1}) - \sigma_0(t_{x_j})\big)\epsilon_{0,t_i}}{w-u}\Bigg).$$

Then again by the Lipschitz continuity of σ_1, σ_0, $\max_{1\le i\le n-1}|t_i - t_{i+1}| \le \frac{1}{n^{1-\varepsilon}}$ and the lemma, there exists K_3 doesn't depend on q such that

$$\mathbb{E}D^2 \le \frac{1}{q^2}\cdot K_3\frac{q^2}{n^{2-2\varepsilon}} = O(\frac{1}{n^{2-2\varepsilon}}),$$

then as long as we take $\varepsilon < \frac{1}{2}$, $\sqrt{qw}D = o_p(1)$, and

$$\sqrt{qw}B = \frac{\sqrt{qw}}{q}\sum_{j=1}^{q}\Bigg(\sum_{i\in\mathcal{B}_j}\frac{\sigma_1(t_{x_j})\epsilon_{1,t_i}}{u} - \sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\sigma_0(t_{x_j})\epsilon_{0,t_i}}{w-u}\Bigg) + o_p(1).$$

Combine above results together, we have

$$\sqrt{qw}(\widehat{\alpha}_{\mathrm{tr}} - \alpha) = \frac{1}{\sqrt{qw}}\sum_{i=1}^{n}(\mu_1(t_i) - \mu_0(t_i) - \alpha)$$
$$+ \frac{\sqrt{qw}}{q}\sum_{j=0}^{q-1}\Bigg(\sum_{i\in\mathcal{B}_j}\frac{\sigma_1(t_{x_{j+1}})\epsilon_{1,t_i}}{u} - \sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\sigma_0(t_{x_{j+1}})\epsilon_{0,t_i}}{w-u}\Bigg) + o_p(1)$$
$$= \frac{1}{\sqrt{qw}}\sum_{i=1, i\notin\mathcal{T}}^{n}(\mu_1(t_i) - \mu_0(t_i) - \alpha) + o_p(1) + \frac{\sqrt{qw}}{q}\sum_{j=1}^{q}$$
$$\Bigg((\mu_1(t_{x_j}) - \mu_0(t_{x_j}) - \alpha) + \sum_{i\in\mathcal{B}_j}\frac{\sigma_1(t_{x_j})\epsilon_{1,t_i}}{u} - \sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\sigma_0(t_{x_j})\epsilon_{0,t_i}}{w-u}\Bigg).$$

By the property of Poisson process, we know $t_i (1 \le i \le n)$ are i.i.d. random variables. In addition, since $\mathcal{T}$ is chosen randomly, $t_{x_1}, \cdots, t_{x_q}$ are independent to $t_i (i \notin \mathcal{T})$, then by the CLT for i.i.d. data, we have:

$$\frac{1}{\sqrt{qw}}\sum_{i=1, i\notin\mathcal{T}}^{n}(\mu_1(t_i) - \mu_0(t_i) - \alpha) \xrightarrow{d} \mathcal{N}\big(0, \frac{w-1}{w}\mathrm{Var}(\mu_1(\tau) - \mu_0(\tau))\big).$$

Also by CLT, we have:

$$\frac{\sqrt{qw}}{q}\sum_{j=1}^{q}\Bigg((\mu_1(t_{x_j}) - \mu_0(t_{x_j}) - \alpha) + \sigma_1(t_{x_j})\sum_{i\in\mathcal{B}_j}\frac{\epsilon_{1,t_i}}{u} - \sigma_0(t_{x_j})\times$$
$$\sum_{i\in\mathcal{A}_j\setminus\mathcal{B}_j}\frac{\epsilon_{0,t_i}}{w-u}\Bigg) \xrightarrow{d} \mathcal{N}\Bigg(0, \mathbb{E}\Big[\frac{\sigma_1^2(\tau)}{p} + \frac{\sigma_0^2(\tau)}{1-p}\Big] + \frac{\mathrm{Var}(\mu_1(\tau) - \mu_0(\tau))}{w}\Bigg).$$

Again by the independence of $t_i (i \in \mathcal{T})$ and $t_j (j \notin \mathcal{T})$, the two results above can be summed up and we obtain:

$$\sqrt{qw}(\widehat{\alpha}_{\mathrm{tr}} - \alpha) \xrightarrow{d} \mathcal{N}(0, V_{\mathrm{tr}}),$$

here

$$V_{\mathrm{tr}} = \mathbb{E}(\frac{\sigma_1^2(\tau)}{p} + \frac{\sigma_0^2(\tau)}{1-p}) + \mathrm{Var}(\mu_1(\tau) - \mu_0(\tau)).$$

Finally, it's easy to see $\frac{qw}{\lambda} \xrightarrow{p} 1$, then

$$\sqrt{\lambda}(\widehat{\alpha}_{\mathrm{tr}} - \alpha) \xrightarrow{d} \mathcal{N}(0, V_{\mathrm{tr}}).$$