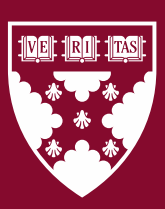
Working Paper 23-070

Design-Based Confidence Sequences: A General Approach to Risk Mitigation in Online Experimentation

Dae Woong Ham
Michael Lindon
Martin Tingley
Iavor Bojinov



Harvard
Business
School

Design-Based Confidence Sequences: A General Approach to Risk Mitigation in Online Experimentation

Dae Woong Ham
Harvard University

Michael Lindon
Netflix

Martin Tingley
Netflix

Iavor Bojinov
Harvard Business School

Working Paper 23-070

Funding for this research was provided in part by Harvard Business School.

Design-Based Confidence Sequences: A General Approach to Risk Mitigation in Online Experimentation

Dae Woong Ham[1], Michael Lindon[2], Martin Tingley[2], and Iavor Bojinov[3]

[1]*Department of Statistics, Harvard University*, [2]*Netflix, U.S.A*, [3]*Harvard School of Business*

May 25, 2023

Abstract

Randomized experiments have become the standard method for companies to evaluate the performance of new products or services. In addition to augmenting managers' decision-making, experimentation mitigates risk by limiting the proportion of customers exposed to innovation. Since many experiments are on customers arriving sequentially, a potential solution is to allow managers to "peek" at the results when new data becomes available and stop the test if the results are statistically significant. Unfortunately, peeking invalidates the statistical guarantees for standard statistical analysis and leads to uncontrolled type-1 error. Our paper provides valid design-based confidence sequences, sequences of confidence intervals with uniform type-1 error guarantees over time for various sequential experiments in an assumption-light manner. In particular, we focus on finite-sample estimands defined on the study participants as a direct measure of the incurred risks by companies. Our proposed confidence sequences are valid for a large class of experiments, including multi-arm bandits, time series, and panel experiments. We further provide a variance reduction technique incorporating modeling assumptions and covariates. Finally, we demonstrate the effectiveness of our proposed approach through a simulation study and three real-world applications from Netflix. Our results show that by using our confidence sequence, harmful experiments could be stopped after only observing a handful of units; for instance, an experiment that Netflix ran on its sign-up page on 30,000 potential customers would have been stopped by our method on the first day before 100 observations.

Keywords: Sequential Analysis, Anytime-valid inference, Asymptotic Confidence Sequence, A/B Test, Time Series Experiments, Panel Experiments, Switchback Experiments, Peeking, Multi-arm Bandits, Adaptive Testing

1 Introduction

Over the past decade, firms have widely adopted controlled experiments (often called A/B tests by practitioners) as the principled way of evaluating the performance of a new product or service (the treatment) relative to the current offering (the control). In a typical A/B test, the firm randomly allocates a proportion of its customers to experience the treatment ("B") while the rest continue to receive control ("A"). Then, after a prespecified amount of time, usually one to two weeks, the results are examined by a manager who uses them as input into the decision of widely deploying the treatment or rolling back to the control. Researchers have found encouraging evidence that companies adopting control experimentation and developing a culture of experimentation see a boost in product innovation, identifying and scaling promising ideas, and informing strategic decision-making (Camuffo et al., 2020; Cohen and Levinthal, 1994; Cohen, Nelson and Walsh, 2002; Koning, Hasan and Chatterji, 2022; Kohavi et al., 2013; Bakshy, Eckles and Bernstein, 2014; Li et al., 2015; Mao and Bojinov, 2021). For example, Koning, Hasan and Chatterji (2022) suggests that companies leveraging A/B testing could see an increase of 30-100% in revenue and visitation traffic after a year of initial adoption.

In addition to augmenting decision-making, experiments decrease the risk associated with innovation by limiting the initial proportion of treated customers to 1-5% of the total population(Kohavi, Tang and Xu, 2020; Bojinov and Gupta, 2022). This is especially relevant as many experiments are also not intended to be direct product improvements but have other justifications such as reducing costs or changing backend infrastructure. In these applications, a treatment effect is not expected, and experiments are effectively employed as quality control gates to mitigate risk. A common example are technology companies where product changes occur frequently through software rollouts (Lindon, Sanden and Shirikian, 2022). As large organizations often have millions of members, this small percentage still represents hundreds of thousands of users, providing sufficient power to detect even minor (negative) effects. However, if a treatment performs exceptionally poorly, prolonged exposure could lead to customer dissatisfaction and substantial revenue loss, a problem exacerbated by technology companies like Amazon, Netflix, and Microsoft conducting thousands of experiments on their customers each year (Thomke, 2020). From a manager's perspective, it's vital to stop such negative experiments before they cause too much harm to the people in the study (Bojinov and Gupta, 2022). For example, in a recent experiment at Netflix (discussed in detail in Section 8.1), the manager planned to terminate the experiment if the number of lost customers due to the new version exceeded a specific threshold, such as 10,000 out of 500,000 potential customers in the study. In order to minimize risk, it is necessary to stop harmful experiments *quickly*.

A potential solution involves allowing managers to "peek" at experimental results as new data becomes available and stopping the test if the results are significant. However, calculating a standard confidence interval each time new data is available would result in an uncontrolled type-1 error rate, as confidence intervals are only reliable when computed once at the end of the study (see Section 1.1 for more details). To ensure the resulting inference is valid—meaning it has an appropriate type-1 error rate—we need to use confidence sequences (a sequence of confidence intervals) that offer a uniform type-1 error guarantee for all time points. This type of inference is often called "anytime-valid" because it enables us to sequentially test the same hypothesis while maintaining type-1 error rate control (Johari, Pekelis and Walsh, 2015; Howard et al., 2020). Shifting from confidence intervals to confidence sequences allows managers to

"peek" at available data and stop the experiment if a statistically significant effect is observed at any time. Due to this increased flexibility, many companies, such as Adobe (Waudby-Smith et al., 2021), Microsoft (Waudby-Smith et al., 2022), and Netflix (Lindon and Malek, 2022; Lindon, Sanden and Shirikian, 2022), have recently started incorporating anytime-valid inference in their experimentation platforms. For example, in Section 8.1, we analyze an experiment conducted by Netflix for approximately 2 weeks on around 30,000 potential customers arriving to their sign-up page with a strong negative treatment effect, i.e., the experiment was harming potential new subscribers from subscribing. We find that using our proposed confidence sequence, Netflix could have stopped the experiment as early as the first day after observing only approximately 70 units, saving potentially hundreds and thousands of potential new subscribers.

A significant limitation of most existing research on this topic is the emphasis on estimands defined for a hypothetical (infinite) super population, such as the population average treatment effect, rather than the realized experimental sample, such as the sample average treatment effect (Waudby-Smith et al., 2021; Waudby-Smith and Ramdas, 2020b; Lindon and Malek, 2022; Waudby-Smith et al., 2022). This approach contrasts with the managerial risk mitigation perspective, which seeks to stop an experiment when the treatment's cost to customers in the study is deemed too high, as measured by, for instance, the sample average. Moreover, for the validity of super population estimands, customers need to be randomly sampled from the population—this assumption is often violated as heavier users are more likely to appear early in the study (Bojinov and Gupta, 2022). Therefore, sequential methods that aim to provide anytime-valid inference for a super population can lead to misleading results. Instead, we focus on developing anytime-valid inference procedures that focus on the participants in the study and estimands defined on the sample not the super population, removing the often violated assumption of random sampling from a super population and directly addressing the managerial perspective.

In particular, we provide an extension of the anytime-valid inference paradigm to the design-based framework. Design-based approaches conditions on the full set of (potential) outcomes and performs inference on quantities directly relevant to the experimental sample. This approach has a long history dating back to Fisher and Neyman but has seen a resurgence in popularity as it allows us to perform inference on the realized sample and handle complicated settings such as interference with minimal assumptions (Fisher, 1935; Neyman, Iwaszkiewicz and Kolodziejczyk, 1935; Bojinov and Shephard, 2019; Ding, Feller and Miratrix, 2016; Zhao and Ding, 2021; Harshaw, Sävje and Wang, 2022). Our design-based confidence sequence retains all the benefits of typical confidence sequences, allowing managers to terminate experiments at any time, for instance, upon detecting a strong harmful treatment effect, without specifying a fixed time horizon. Additionally, we estimate quantities directly related to the realized experimental sample, providing managers with a precise measure for risk mitigation.

Our work also extends existing research to address more complex yet standard experimental settings, such as multi-arm bandits, time series, and panel experiments, with minimal and practitioner-friendly assumptions. For example, multi-arm bandit (MAB) algorithms are a popular and well-established framework for sequential decision-making and adaptive A/B testing (Robbins, 1952; Bastani and Bayati, 2020; Rusmevichientong and Tsitsiklis, 2010) as they allow for rapid iteration due to their regret-minimizing properties that can adaptively adjust the probability a newly arrived customer is assigned to the treatment group. In Section 8.1, we apply our framework to analyze a MAB experiment

at Netflix where treatment assignment probabilities were updated based on the success of each treatment. Another important example we discuss is time series experiments, where the same units receive multiple treatments across time Bojinov and Shephard (2019); Xiong et al. (2019). This type of experimental design has seen rapid growth in adoption, especially by platform marketplace companies such as Doordash, Uber, and Lyft, which regularly use switchback experiments (a special case of time series experiments (Bojinov, Simchi-Levi and Zhao, 2020)). In Section 8.2, we apply our framework to analyze a panel experiment (the generalization of time series experiments to multiple units receiving different treatments over time (Bojinov, Rambachan and Shephard, 2021)) that Netflix ran on 2000 members for seven days to measure the effectiveness of sending push notification messages. In these more complex scenarios, it becomes harder to justify the stronger and untestable technical assumptions on the outcome distribution that are required by existing methods. Our methodology does not require such technical assumptions and can handle many complex adaptive designs, providing more flexibility for firms running these experiments.

Our primary contribution is developing valid design-based confidence sequences for the average treatment effect for n individuals (Imbens and Rubin, 2015), the mean reward difference in the multi-arm bandit setting (Sutton and Barto, 2018), the contemporaneous treatment effect for a single time series experiment with carryover effects (Bojinov and Shephard, 2019), and the average contemporaneous treatment effect for panel data settings for n individuals that are also observed across time t (Bojinov, Rambachan and Shephard, 2021). Our design-based approach allows us to naturally relax any independence or distributional assumptions and performs inference on the realized sample, producing interpretable and managerially relevant results. Moreover, our results are agnostic to the outcome distribution, working for continuous, count, and binary data. This is particularly important as most companies compute the results across thousands of business metrics that have very different distributions. Finally, our confidence sequences are exceptionally easy to compute as they only rely on means and variance, which can be efficiently computed for large streaming and batched data.

1.1 Confidence intervals and sequences

We begin by formally defining confidence intervals. We say I_t is a valid confidence interval with type-1 error α for the true treatment effect μ_t if

$$\Pr(\mu_t \in I_t) \geq 1 - \alpha \tag{1}$$

holds. Equation (1) shows that I_t covers the truth with at least $1 - \alpha$ probability, where α denotes the probability of failing to contain the truth, i.e., type-1 error. For instance setting $\alpha = 0.1$ allows practitioners to interpret I_t as a random interval containing the truth parameter μ_t 90% of the time. It is well known that peeking at the data and taking any action will invalidate the type-1 error guarantee in Equation (1) (Johari et al., 2017). In other words, if we construct a valid confidence interval I_t for each time t, and then verify if all the intervals contain the truth, we would no longer be able to assert that our error rate is only α.

To illustrate this issue, we show in Figure 1 the inflated type-1 error when performing a t-test every time a new observation is observed, where each observation is independently drawn from a $N(0, 1)$. For example, when there are 500 observations, the type-1 error is uncontrolled at 70% when $\alpha = 0.10$ as shown by the blue line. The issue is further

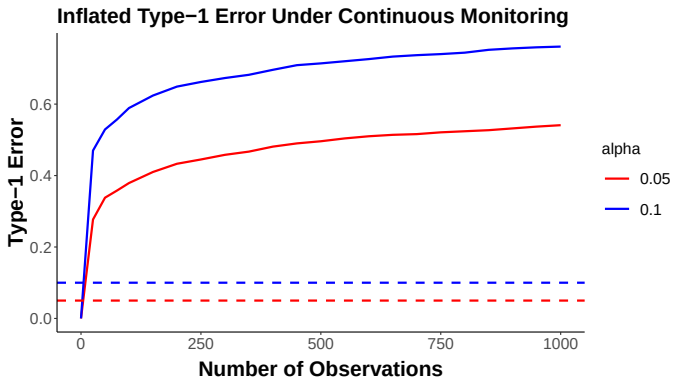


Figure 1: The figure plots the type-1 error when an analyst performs an $\alpha = 0.05$ and $\alpha = 0.1$ level t-test every time a new observation arrives (x-axis), when the data is independently and identically drawn from a $N(0, 1)$ distribution. The solid red and blue lines show the inflated type-1 error rates as the number of observations grows for $\alpha = 0.05$ and $\alpha = 0.10$, respectively. The corresponding horizontal dotted lines show the expected type-1 error at 0.05 and 0.10, respectively.

exacerbated if the data is dynamically and adaptively collected, i.e., if the next observation is sampled as a function of the previous observations.

To overcome this, companies and researchers have started using methods to compute confidence sequences to allow analysts to continuously monitor the experiment (Lindon and Malek, 2022; Waudby-Smith and Ramdas, 2020a; Darling and Robbins, 1967; Wang and Ramdas, 2022; Howard et al., 2020). Confidence sequences are sequences of confidence sets with time-uniform coverage guarantees, that is, the probability that *all* confidence sets cover the estimand is at least $1 - \alpha$. Formally, a sequence of intervals $\{V_t\}_{t=1}^{\infty}$ is a $1 - \alpha$ confidence sequence for a parameter μ_t, e.g., time-varying ATE, if

$$\Pr(\forall t, \mu_t \in V_t) \geq 1 - \alpha \iff \Pr(\exists t, \mu_t \notin V_t) \leq \alpha. \tag{2}$$

It follows that $\Pr(\mu_T \notin V_T) \leq \alpha$ for arbitrary data-dependent stopping rule T. In words, Equation (2) allows the analyst to continuously monitor ongoing experiments through V_t and also immediately stop the experiment whenever the manager desires (for example detecting a significant negative treatment effect for the Netflix sign-up page experiment) while controlling the error below the nominal threshold, as depicted by the dotted blue and red lines in Figure 1, uniformly at *all times*.

Additionally, as companies rapidly scale the number of experiments that are performed, there is an increasing demand to automate the experiment lifecycle and have experiments managed algorithmically. Anytime-valid inference allows companies to automate the stopping time of experiments by defining algorithmic stopping rules in terms of the confidence sequence, such as when the confidence sequence excludes zero, or when the confidence sequence is sufficiently small around zero indicating that any difference, if it exists, is not practically meaningful. As we show in our empirical application, this type of large-scale automation has the potential to significantly de-risk innovation.

1.2 Outline

In Section 2, we introduce the design-based framework and the relevant causal estimands and estimators for the most classical and simple setting, where the goal is to infer the average treatment effect from n individuals. In Section 3, we first give a brief overview of how confidence sequences are constructed via martingales. Then, we propose an exact confidence sequence that is valid for general settings but depends on potentially unknown parameters of the data-generating process. Section 4 introduces an improved design-based asymptotic confidence sequence. In Section 4.2, we introduce the improved asymptotic confidence sequences in the simplest case for the average treatment effect of individuals with independent data. In Section 4.3, we extend the results to allow fully dynamic and adaptive treatment assignments, allowing sequential testing in the multi-arm bandit literature. Section 5.1- 5.2 generalizes the results to a time series setting with potential carryover effects and the panel data setting. We also provide illustrative examples of using our design-based confidence sequence for each of the aforementioned setting. Section 6 proposes an important variance reduction technique by incorporating prior information or covariates. Section 7 shows simulations that demonstrate the favorable properties of our main results and contrasts confidence sequences with confidence intervals to guide practitioners. In Section 8, we present three real world applications of our methodology to experiments that Netflix. Finally, Section 9 presents our concluding remarks.

We purposefully present results in the simplest settings first before generalizing the results to more complicated settings to emphasize how our results cover a wide variety of problems. We also believe that separating each cases and providing an example brings clarity both notationally and conceptually for what the goal and causal estimand is for each setting.

2 Causal estimands, estimators and inference

Suppose N independent units arrive sequentially; for each unit, we randomly assign them to a treatment W_i and subsequently observe a corresponding outcome Y_i so that we observe data $\{W_i, Y_i\}_{i=1}^N$. In our Netflix application, the N units represent the different customers arriving on the company's sign-up page, upon arrival they are assigned to either a new version $W_i = 1$ or the current offering $W_i = 0$; we then observe if they signed up $Y_i = 1$ or did not $Y_i = 0$. Throughout, we focus on the binary treatment assignments for simplicity, but our results generalize to treatments with multiple levels.

Under the potential outcomes formulation of causal inference, each user has a pair of potential outcomes $\{Y_i(1), Y_i(0)\}$, corresponding to what would happen if the user is assigned to the treatment $Y_i(1)$ or the control $Y_i(0)$; see (Neyman, Iwaszkiewicz and Kolodziejczyk, 1935; Rubin, 1974; Robins, 1986). Note that we are explicitly assuming that there is no interference between experimental units (i.e., assumes one unit's treatment assignment does not impact another unit's outcome (Cox, 1958)). Often this assumption is combined with the requirement that there are no alternative versions of the treatment and is called the stable treatment value assumption (Rubin, 1986). Of course, there are many practical examples where the no-interference assumption is violated; however, we leave the full exploration of such settings to future work. Later, we relax the no-interference assumption in Section 5. We can connect the potential

outcomes to the observed outcome by noting that

$$Y_i = W_i Y_i(1) + (1 - W_i) Y_i(0), \tag{3}$$

under the assumption of fully compliance that is trivially satisfied in our experimental setup.

Our paper crucially treats the potential outcomes as fixed and leverages the randomization due to the experiment for inference. Therefore, Equation (3) shows that Y_i is only random through W_i because $\{Y_i(1), Y_i(0)\}$ are fixed constants. This is typically referred to as a design-based approach and can equivalently be derived by conditioning on the full set of potential outcomes (Abadie et al., 2020; Rubin, 1974; Basse and Airoldi, 2017). To make this formal, we denote $\mathcal{F}_n$ as the sigma-algebra that contains all pairs of N potential outcomes $\{Y_i(1), Y_i(0)\}_{i=1}^N$ and all observed independent data up to the n^{th} user $\{W_i, Y_i\}_{i=1}^n$ conditional on the N pairs of potential outcomes. A big advantage of the design-based approach is that it avoids making any assumptions about the distribution of the potential outcomes and solely leverages the randomness of the treatment W for inference.

Our goal is to estimate the finite-sample average treatment effect τ_N for the N individuals in our obtained experimental sample.

Definition 2.1 (Finite-Sample Average Treatment Effect)**.**

$$\tau_N := \frac{1}{N} \sum_{i=1}^{N} Y_i(1) - Y_i(0),$$

where N denotes our finite-sample population of interest. In a sequential experimental setting, N changes as more data arrives (further formalized in Theorem 3.2). Definition 2.1 represents an average treatment directly related to the N experimental sample as opposed to the more commonly proposed estimand related to an infinite general population defined as

$$\tau_{\text{sp}} := E_{\text{sp}}[Y_i(1) - Y_i(0)],$$

where the expectation is taken over the distribution of the potential outcomes assuming the potential outcomes are identical and independently distributed for all units. Comparing the two estimands, we see that τ_N is directly related to the effect of the treatment on the N units enrolled in the experiment while τ_{sp} has no relation (having no N in the definition). Therefore, unlike the super population estimand τ_{sp}, our finite-sample estimand τ_N is directly aligned with the managerial risk mitigation perspective. For example, if $\tau_N = -0.1$ in the sign-up page experiment for $N = 100,000$ customers in our experiment, the manager can directly conclude that running this experiment led to a potential loss of $10,000$ sign-ups. Additionally, our estimand does not require us to assume that the potential outcomes are identically and independently distributed, an assumption that is often violated in more complex settings.

Although the individual treatment effect, $Y_i(1) - Y_i(0)$, is never jointly observed, we use the inverse propensity score estimator (Horvitz and Thompson, 1952; Imbens and Rubin, 2015) $\hat{\tau}_i$ that is unbiased for for the individual treatment effect, where

$$\hat{\tau}_i := \frac{\mathbb{1}\{W_i = 1\} Y_i}{p_i(1)} - \frac{\mathbb{1}\{W_i = 0\} Y_i}{p_i(0)} \tag{4}$$

and $p_i(w) := \Pr(W_i = w)$. To ensure that $\hat{\tau}_i$ is well defined, we make the following positivity assumption.

Assumption 1 (Probabilistic Treatment Assignment)**.** For every $i = 1, 2, \ldots, \infty$ and $w \in \{0, 1\}$,

$$0 < p_i(w) < 1,$$

where $p_\infty(w)$ is defined as the limiting propensity score.

In words, Assumption 1 states that the propensity scores for every individual is bounded away from zero and one. For example, this states that each customers arriving to the sign-up page has some positive probability to see both the new or current offering. Lastly, classical results show that the average of $\hat{\tau}_i$ is an unbiased estimator of $Y_i(1) - Y_i(0)$, providing also a simple estimate of the (upper bound) of the variance. The variance estimate is an upper bound because the actual variance term contains the product $Y_i(1)Y_i(0)$, which are never jointly observed. Thus, we use the following inequality $a^2 + b^2 \geq 2ab$ for any $a, b \in \mathbb{R}$ to obtain the following upper bound.

Lemma 2.1 (Mean and Variance of Inverse Propensity Score Estimator)**.** Under Assumption 1,

$$\frac{1}{N}\sum_{i=1}^{N} E(\hat{\tau}_i \mid \mathcal{F}_{i-1}) = \tau_N.$$

Furthermore, the upper bound of the variance has the following unbiased estimator

$$\mathrm{Var}(\hat{\tau}_i \mid \mathcal{F}_{i-1}) \leq E(\hat{\sigma}_i^2 \mid \mathcal{F}_{i-1}), \quad \hat{\sigma}_i^2 := \frac{\mathbb{1}\{W_i = 1\}Y_i^2}{p_i(1)^2} + \frac{\mathbb{1}\{W_i = 0\}Y_i^2}{p_i(0)^2}, \tag{5}$$

where all expectations throughout the rest of this paper is taken with respect to the randomness of the treatment assignment conditioning on the potential outcomes.

The proof is provided in Appendix A. Before stating our first result, we also assume the potential outcomes are bounded.

Assumption 2 (Bounded Potential Outcomes)**.**

$$|Y_i(w)| \leq M$$

for all i and $w \in \{0, 1\}$, where $M \in \mathbb{R}$.

The value of M can be chosen to be extreme, ensuring that this assumption holds. This strategy is commonly employed to satisfy the regularity conditions necessary for design-based inference (Bojinov and Shephard, 2019; Lei and Ding, 2020). In the finite-population framework, it is likely implausible for any units to have infinite potential outcome. Furthermore, Assumption 2 is about the *realized* potential outcome. For example, if N user's potential outcomes were generated from a $N(0, 1)$ distribution, an unbounded distribution, each of the N user's realized potential outcome is still bounded. More concretely, this assumption states that the N Netflix customers all have finite responses on metrics Netflix cares about, e.g., daily streaming hours, successful sign-ups, etc. Therefore, we view Assumption 2 as a mild regularity condition.

3 Exact design-based confidence sequences

3.1 Confidence sequence construction through martingales

We begin by introducing how confidence sequences are constructed through martingales. Equation (2) requires that the confidence sequence V_t contains the target parameter uniformly, allowing managers to perform hypothesis tests at *every* time. On one hand, this can be viewed as a multiple-testing problem, where we control for type-1 error while performing multiple hypothesis tests. Although any naive approach such as Bonferroni correction will technically lead to valid type-1 error guarantees, such an approach leads to uncontrolled type-2 error since $\alpha \to 0$ as time grows. Thus, the state-of-the-art approaches have used martingale constructions that automatically give a time-uniform guarantee through Ville's maximal inequality (Ville, 1939).

Lemma 3.1 (Ville's Maximal Inequality)**.** Let M_n be a non-negative supermartingale with respect to a filtration $\mathcal{F}$ with initial value $M_0 = 1$. Then,

$$\Pr\left(\exists n \in \mathbb{N}_0 \,:\, M_n \geq \frac{1}{\alpha}\right) \leq \alpha.$$

Contrasting Lemma 3.1 with the desired gaurantee in Equation (2), we see that Ville's Maximal Inequality directly gives the desired *uniform* type-1 error guarantee. Thus, most constructions of confidence sequences rely on constructing a non-negative supermartingale and applying Lemma 3.1 to achieve the time-uniform guarantee (Lindon and Malek, 2022; Howard et al., 2020; Waudby-Smith and Ramdas, 2020a).

3.2 Design-based exact confidence sequence

We now apply the empirical Bernstein inequalities to derive a design-based confidence sequence for τ_N (Maurer and Pontil, 2009).

Theorem 3.2 (Design-based Non-asymptotic Confidence Sequence for the ATE)**.** Suppose independent $\{W_i, Y_i\}_{i=1}^N$ are observed for any arbitrary data dependent stopping time N,[1] where Assumptions 1- 2 are satisfied. Let $m := M/p_{min}$, where $p_{min} = \min_{i,w} p_i(w)$, i.e., m is the most extreme value our estimate $\hat{\tau}_i$ can take for any i. Denote $S_n := \sum_{i=1}^n \hat{\sigma}_i^2$. Then,

$$\frac{1}{n}\sum_{i=1}^n \hat{\tau}_i \pm \left[\frac{m(m+1)}{n}\log\left(\frac{2}{\alpha}\right) + \frac{S_n}{n}\left(\frac{m+1}{m}\log\left(1+\frac{1}{m}\right) - \frac{1}{m}\right)\right]$$

forms a valid $(1-\alpha)$ confidence sequence for the average treatment effect τ_n defined in Definition 2.1.

The proof is provided in Appendix B. The key part of the proof is that we show

$$\exp\left[\frac{\sum_{i=1}^n(\hat{\tau}_i - \tau_n)}{m(m+1)} + \frac{S_n}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right]$$

[1] With a slight abuse of notation we also use N (the population size) as a stopping time because once the experiment is terminated at "time" N, the confidence sequence infers the average treatment effect of the N individuals in the available data, making it the "population". Furthermore, because N is a data-dependent stopping time, we require that it is formally a well-defined stopping time (a measurable function dependent on the current and previous data).

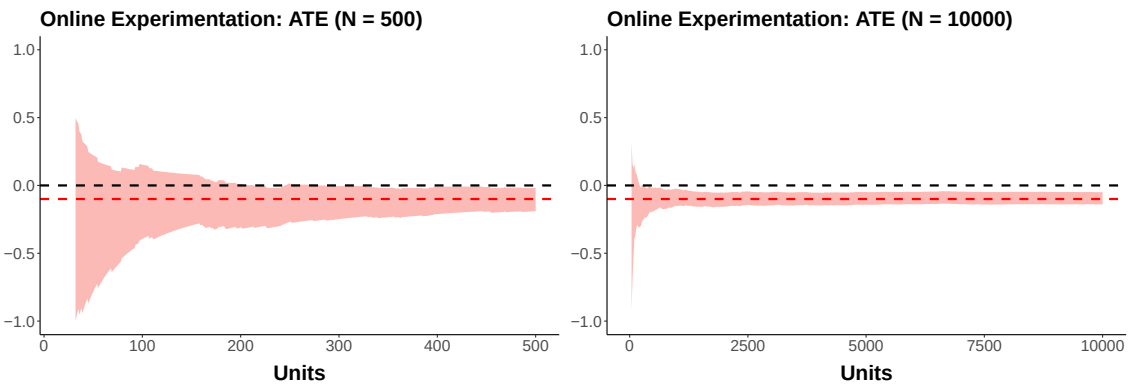


Figure 2: Online Experimentation (Example 1). The red contours show the lower and upper confidence sequence as units enter in the experiment at $\alpha = 0.05$ using Theorem 3.2. The left and right panel show the confidence sequences for $N = 500, 5000$ units, respectively. The horizontal red dotted line represent the true treatment effect of $\tau_n = 0.1$ and the black horizontal dotted line represents the zero (null) line.

forms a supermartingale by applying Fan's inequality in (Fan, Grama and Liu, 2015). The rest follows from algebraically manipulating the supermartingale after applying Lemma 3.1.

There are three practical limitations to Theorem 3.2. First, the confidence sequence scales with M, which is dependent on both the observed and missing potential outcomes. Although Assumption 2 can be seen as a mild regularity condition as finite-sample units likely do not have infinite potential outcomes, analysts often do not know this extreme value, except in special cases such as binary outcomes. This issue is further exacerbated if there exists one extreme potential outcome.

Second, in the spirit of a sequential test, the analyst may desire to change $p_i(w)$ as more units enter. However, Theorem 3.2 requires p_{min} to be determined before the start of the experiment, thus restricting the treatment assignment probabilities to a pre-specified range. Although this is common practice for many experiments, the confidence width further scales inversely with p_{min}.

Third, the confidence width is of order approximately S_n/n, which does not shrink asymptotically ($n \to \infty$) to zero unless S_n grows sub-linearly ($\hat{\sigma}_i$ vanishes to zero) or there are stronger assumptions on the potential outcomes. This issue was recognized in Howard et al. (2020) who propose a solution under additional conditions. Nevertheless, we present the exact confidence sequence under general conditions because it is still useful for special cases such as binary outcomes (see Example 1) and emits a closed-form design-based confidence sequence that is easy to implement. We show in Appendix C how to leverage a mixture distribution with the truncated gamma distribution to build another confidence sequence with order roughly $O(\sqrt{S_n \log S_n}/n)$, which does not have a closed-form and requires root-solving algorithms. We illustrate the confidence sequence in practice in Example 1 below.

Example 1 (Online Experimentation)**.** Suppose individuals visit a website sequentially over time. When the page loads, each person is randomly assigned with probability $1/2$ to the original version "A" (control) or the new version "B" (treatment). Suppose we are interested in a binary outcome that tracks if the user clicks on the sign-up page. Then,

in the notation from Theorem 3.2, $M = 1$ and $m = 1/0.5 = 2$. Finally, suppose the ground truth is that each user has a 0.15 and 0.05 chance of clicking on the sign-up button for versions "A" and "B," respectively; mathematically, $\Pr(Y_i(0) = 1) = 0.15$ and $\Pr(Y_i(1)) = 0.05$ for all i. The average treatment effect is roughly -0.1. The left and right panels of Figure 2 show the confidence sequence from applying Theorem 3.2 on one simulated experimental data in the aforementioned setting for $N = 500$ and $N = 10,000$ units, respectively. The left panel shows that the analyst would likely stop at the 210^{th} individual because the confidence sequence shows a statistically significant negative treatment effect (the first time the red contours do not overlap the black dotted line), mitigating further harm. However, the right panel shows that the confidence width does not shrink despite a large $N = 10,000$, illustrating the aforementioned limitations.

4 Design-based asymptotic confidence sequences

We now improve the design-based *non-asymptotic* confidence sequence introduced in Section 3.2 by relaxing assumptions and obtaining an *asymptotically* valid confidence sequence. Asymptotic confidence sequences were first introduced by (Waudby-Smith et al., 2021), and we extend their work to the design-based framework.

4.1 Asymptotic confidence sequence

Informally, asymptotic confidence sequences are valid confidence sequences after a "sufficiently large" time. Although this may worry practitioners since confidence sequences, by definition, should have valid coverage at all times (including early times), we show through simulations in Section 7 that the empirical coverage remains robust in practice because the confidence sequence width for early times is wide in part due to our upper bound variance estimator.

Definition 4.1 (Asymptotic Confidence Sequences)**.** We say that $(\hat{\mu}_i \pm V_i)$ is a two-sided $(1-\alpha)$ asymptotic confidence sequence for a target parameter μ_i if there exists a non-asymptotic confidence sequence $(\hat{\mu}_i \pm \tilde{V}_i)$ for μ_i such that

$$\frac{\tilde{V}_t}{V_t} \xrightarrow{a.s.} 1. \tag{6}$$

Furthermore, we say V_t has an approximation rate R if $\tilde{V}_t - V_t = O_{a.s.}(R)$, where R can be interpreted as how fast the approximation error $\tilde{V}_t - V_t$ is decreasing.

To readers familiar with asymptotic confidence *intervals*, the above definition may be puzzling. To give some intuition, the idea of "couplings" has been used in the literature of strong approximations and invariance principals (Einmahl, 1987; Komlos, Major and Tusnády, 1976) to formally define asymptotic confidence *intervals*. This literature defines a $(1-\alpha)$ asymptotic confidence interval if there exists a *non-asymptotic* (unknown) confidence interval centered at the same statistic such that the difference between this non-asymptotic and asymptotic confidence interval is negligible asymptotically. Equation (6) captures the same notion except we replace all convergence statement with almost-sure convergence to satisfy the time *uniform* guarantee required of confidence sequences.[2] Throughout the rest of the paper (including the appendix), we omit subscript "a.s." from $o_{a.s.}(.)$ and $O_{a.s.}(.)$ to simplify notation.

[2]This is further discussed and proven in Appendix C.4 of (Waudby-Smith et al., 2021)

4.2 Design-based asymptotic confidence sequence for the ATE

We now state one of the main result of our paper, namely that we can construct asymptotically valid confidence sequences for τ_N using $\hat{\tau}_i$ and our estimated variance $\hat{\sigma}_i^2$. Before stating the theorem, we require one additional assumption so that our variance is well-behaved and does not vanish in the limit to allow for asymptotic approximations.

Assumption 3 (None Vanishing Variance)**.** In Appendix A, we show that $\text{Var}(\hat{\tau}_i) \leq \sigma_i^2$, where

$$\sigma_i^2 = \frac{Y_i(1)^2}{p_i(1)} + \frac{Y_i(0)^2}{p_i(0)}. \tag{7}$$

We say Assumption 3 holds if

$$\frac{1}{\sum_{i=1}^n \sigma_i^2} = o(1) \iff \tilde{S}_n := \sum_{i=1}^n \sigma_i^2 \xrightarrow{n\to\infty} \infty \text{ almost surely.}$$

Assumption 3 holds if $\frac{1}{n}\tilde{S}_n \xrightarrow{a.s.} \sigma_*^2$ or if $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_N^2$. Informally, Assumption 3 is satisfied as long as the potential outcomes do not vanish to zero as time grows. Conversely, one way Assumption 3 does not hold is if the response of *all* the users arriving over time suddenly vanish to zero after a certain time. We believe this is unlikely in practice.

Theorem 4.1 (Design-based Asymptotic Confidence Sequence for the ATE)**.** Assume the same setting as that in Theorem 3.2, where additionally Assumption 3 is satisfied. Then,

$$\frac{1}{n}\sum_{i=1}^n \hat{\tau}_i \pm \frac{1}{n}\sqrt{\frac{S_n\eta^2+1}{\eta^2}\log\left(\frac{S_n\eta^2+1}{\alpha^2}\right)}$$

forms a valid $(1-\alpha)$ asymptotic confidence sequence for the average treatment effect τ_n with approximation rate $o\left(\sqrt{\tilde{S}_n \log \tilde{S}_n}/n\right)$ for any pre-specified constant $\eta > 0$.

The proof is in Appendix D and follows from using strong martingale difference sequence approximation, constructing a martingale from the limiting Gaussian distribution, and plugging in a consistent variance estimator. The key difference between this result and Theorem 3.2 is that the confidence width scales approximately (ignoring log terms) as $O(\sqrt{S_n}/n)$ as opposed to $O(S_n/n)$. In contrast to the non-asymptotic confidence sequence presented in Section 3, the confidence width has no expressions related to p_{min} or M, thus making it a fully general result. Ignoring constant terms, Theorem 4.1 also shows that the width scales similar to a confidence interval, which usually scales as $O(\sqrt{S_n}/n)$. However, the confidence sequence has an extra log penalty term to achieve the necessary time-uniform guarantees (see Section 7.2 for simulations comparing confidence intervals and sequences).

Lastly, η is a tuning parameter typically chosen by the analyst to minimize the confidence width at a certain fixed time. Given $\tilde{S}_n$ and n, the confidence width can be minimized for a choice of η, and we give a closed form expression for the minimum in Appendix E. Although the choice of η can impact the confidence sequence especially at the beginning when sample size is small, the choice of η becomes less significant as more data arrives. Consequently, for all examples and simulations we choose η so that the width is minimized at time 10 or, equivalently, the 10^{th} individual since the

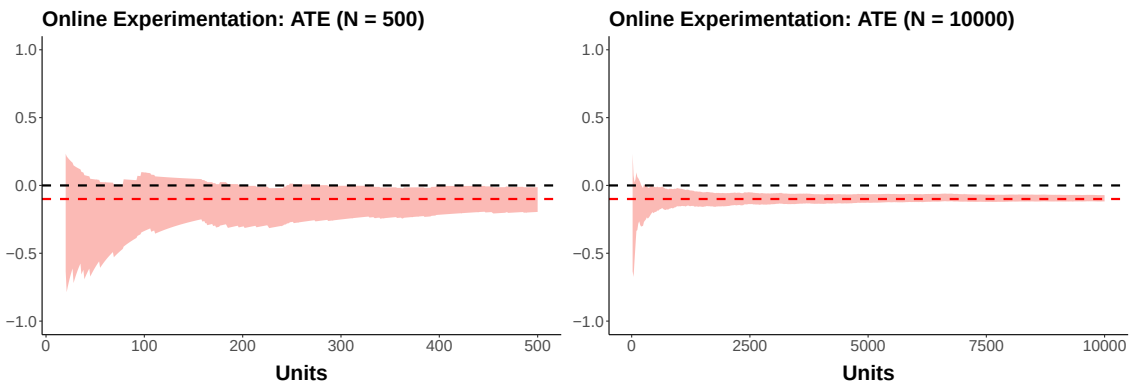


Figure 3: Online Experimentation Continued (Example 1). The setting is identical to that of Figure 2 except the red contours show the lower and upper confidence sequence as units enter in the experiment at $\alpha = 0.05$ using the design-based asymptotic confidence sequences in Theorem 4.1.

confidence sequence width is the largest at early times. For practitioners who do not want to calculate η, we also recommend using the same η we do throughout the paper, i.e, $\eta \approx 0.77$. We now return to Example 1 to illustrate Theorem 4.1.

Example 2 (Online Experimentation Continued)**.** Suppose the same setting as that presented in Example 1. Figure 3 shows the confidence sequence from applying Theorem 4.1 on the same simulated experimental data in the aforementioned setting for $N = 500$ and $N = 10,000$ units. The left panel shows the analyst would likely terminate the experiment at approximately the 200^{th} unit. Although the left panel of Figure 3 is not substantially different than the left panel of Figure 2 (largely because m is not extreme due to the binary outcomes), the right panel shows that the confidence sequence width shrinks to zero unlike the confidence sequence in the right panel of Figure 2.

4.3 Dynamic updating and bandit setting

We now extend the previous confidence sequence when the treatment assignments are adaptive based on previous data. This now allows us to tackle the common bandit settings, where an agent is typically tasked to maximize reward and pull the next arm as a function of all the previous treatment assignments (Slivkins, 2019; Sutton and Barto, 2018).

To formalize this, we define the *adaptive* probability assignments the following way with the same positivity assumption.

Assumption 4 (Adaptive Probabilistic Treatment Assignment)**.** For every $i \geq 1$ and $w \in \{0, 1\}$,

$$0 < p_{i|i-1}(w) := \Pr(W_i = w \mid \mathcal{F}_{i-1}) < 1.$$

We also require the limiting adaptive probability $p_\infty(w)$ to be bounded away from zero or one. This is often a "stricter" assumption for multi-arm bandits as many adaptive algorithms such as Thompson Sampling (Thompson, 1933; Russo, 2020) make the adaptive probabilities converge to one or zero. To solve this concern, we recommend practitioners to

use a mixture of a uniform explore policy and an adaptive exploit policy, e.g., assigning an epsilon mixture weight to the uniform explore policy ensures the propensity scores are bounded.

We next redefine $\hat{\tau}_i, \hat{\sigma}_i$ to incorporate this new adaptive probability with the following notation.

$$\hat{\tau}_{i|i-1} := \frac{\mathbb{1}\{W_i = 1\}Y_i}{p_{i|i-1}(1)} - \frac{\mathbb{1}\{W_i = 0\}Y_i}{p_{i|i-1}(0)}, \quad \hat{\sigma}^2_{i|i-1} := \frac{\mathbb{1}\{W_i = 1\}Y_i^2}{p_{i|i-1}(1)^2} + \frac{\mathbb{1}\{W_i = 0\}Y_i^2}{p_{i|i-1}(0)^2}.$$

Similar to Assumption 3, we assume the variance based on the new adaptive probabilities do not vanish.

Assumption 5 (None Vanishing Variance in Dynamic Settings)**.** Let $\text{Var}(\hat{\tau}_{i|i-1} \mid \mathcal{F}_{i-1}) \leq \sigma^2_{i|i-1}$, where $\sigma^2_{i|i-1}$ is identical to Equation (7) except $p_i(w)$ is replaced with the adaptive probability assignments. Then we assume that

$$\frac{1}{\sum_{i=1}^n \sigma^2_{i|i-1}} = o(1) \iff \tilde{S}_{n|n-1} := \sum_{i=1}^n \sigma^2_{i|i-1} \xrightarrow{n\to\infty} \infty \text{ almost surely.}$$

With these new definitions and assumptions, we directly extend Theorem 4.1 for the bandit setting with the following corollary that is proved in Appendix D.

Corollary 4.1 (Design-based Asymptotic Confidence Sequences for Bandit Settings)**.** Suppose (non-independent) $\{W_i, Y_i\}_{i=1}^N$ are observed for arbitrary data dependent stopping time N, where Assumptions 2, 4, and 5 are satisfied. Let $S_{n|n-1} := \sum_{i=1}^n \hat{\sigma}^2_{i|i-1}$. Then,

$$\frac{1}{n}\sum_{i=1}^n \hat{\tau}_{i|i-1} \pm \frac{1}{n}\sqrt{\frac{S_{n|n-1}\eta^2 + 1}{\eta^2} \log\left(\frac{S_{n|n-1}\eta^2 + 1}{\alpha^2}\right)}$$

forms a valid $(1-\alpha)$ asymptotic confidence sequence for the average treatment effect τ_n with approximation rate $o\left(\sqrt{\tilde{S}_{n|n-1}\log \tilde{S}_{n|n-1}}/n\right)$ for any pre-specified constant $\eta > 0$.

The confidence width in Corollary 4.1 has all the same benefits as those enjoyed by the confidence width in Theorem 4.1. For example, it does not depend on any hyperparameters of the data (M, p_{min}) and the width also asymptotically shrinks to zero.

Example 3 (Two Arm Bandit)**.** Suppose for simplicity the two-arm bandit problem, where an agent pulls either arm A (control) or arm B (treatment). Suppose the rewards under arms A and B have distributions $N(1,1)$ and $N(2,1)$, respectively. Consider the following adaptive probabilities for the n^{th} individual

$$p_{n|n-1}(1) = \frac{\bar{Y}_{1,n-1}}{\bar{Y}_{1,n-1} + \bar{Y}_{0,n-1}}, \quad n > 10$$

and $p_{n|n-1}(0) = 1 - p_{n|n-1}(1)$, where $\bar{Y}_{1,n-1}$ is the sample mean of arm B using the realized samples for $i = 1, 2, \ldots, n-1$ and $\bar{Y}_{0,n-1}$ is defined similarly. In other words, the agent upweights the arm that produce a higher mean reward based on the sample means. Lastly, the agent flips a fair coin for the first ten time periods (exploration period). Figure 4 shows that even under adaptively sampled data, our confidence sequence tightens to the desired truth. For this case, the agent would likely terminate the experiment at approximately $n = 150$ because the reward from arm B is statistically significantly higher than the reward from arm A.

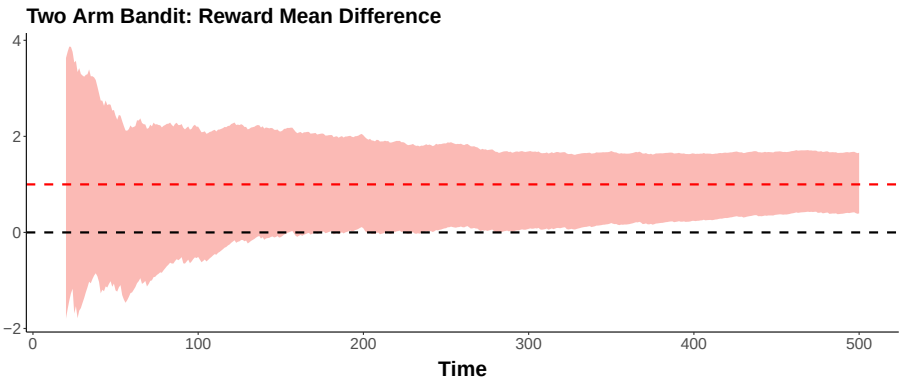


Figure 4: Two Arm Bandit (Example 3). The red contours show the lower and upper confidence sequence as the agent pulls each arm adaptively at $\alpha = 0.05$ using Corollary 4.1. The horizontal red dotted line represent the true mean difference of the rewards and the black horizontal dotted line represents the zero (null) line.

5 Extensions to time series experiments and panel data

We now extend our confidence sequences to time series experiments where a single unit, as opposed to different individuals, receives multiple treatments over time with potential carryover effects (that is, past treatments impact current outcomes). We then further generalize our results to the panel data experiments, where we have time series experiments for all N units. The panel data setting is especially common for many organizations where they randomly assign N customers to treatment and control and observe them over time.

5.1 Time series experiments with carryover effects

We begin by changing notation by using subscript time t instead of unit i, e.g., Y_t, so readers can conceptually understand that we are now in a time series setting. In this setting, there are often strong carryover effects, i.e., the potential outcome is not only a function of its current treatment assignment but of the whole treatment assignment path $Y_t(w_1, w_2, \dots, w_t)$. For example, organizations such as Uber and Lyft use switchback experiments to employ a different algorithm (W_t) at each time t and the potential outcome may depend on the previous treatment paths (Bojinov, Simchi-Levi and Zhao, 2020; Farronato, MacCormack and Mehta, 2018; Aronow et al., 2020; Toulis and Kao, 2013). Consequently, we weaken the no-interference assumption for potential outcomes and build confidence sequences for causal effects that account for carryover effects.

For any random variable O_t, we first denote $O_{1:t} = (O_1, O_2, \dots, O_t)$, thus $W_{1:t}$ denotes the vector of treatment paths up to time t and $w_{1:t}$ is a realization of the random variable $W_{1:t}$. We allow our potential outcomes to be dependent on all the treatment assignment path by writing $Y_t(w_{1:t})$, where we assume that the potential outcome at time t does not depend on future treatment assignments. Our observed outcome is $Y_t = Y_t(w^{obs}_{1:t})$, where $w^{obs}_{1:t}$ is our observed treatment assignment path. We then denote the entire collection of potential outcomes up to time t as

$$Y_{1:t}(\bullet) = \{Y_1(\bullet), Y_2(\bullet), \dots, Y_t(\bullet)\},$$

where $Y_t(\bullet) = \{Y_t(w_{1:t}) : w_{1:t} \in \{0,1\}^t\}$ denotes the entire possible collection of potential outcome at time t. As similarly done before, the design-based approach conditions on $Y_t(\bullet)$ for all t. Unfortunately, the number of potential outcomes grow exponentially with t, thus we focus on the contemporaneous causal effect that is a function of our observed treatment path.

Definition 5.1 (Contemporaneous Causal Effect)**.**

$$\tau_t(w^{obs}_{1:(t-1)}) := Y_t(w^{obs}_{1:(t-1)}, 1) - Y_t(w^{obs}_{1:(t-1)}, 0),$$

Definition 5.1 captures the contemporaneous treatment effect (CTE) had the unit received treatment at time t conditioning on our past treatment path. We specifically define the causal estimand as a function of our observed treatment path to show that our causal estimand changes as a function of our treatment path. Defining the treatment effect in this way is similar to focusing on the average effect on the treated, which is a widely accepted causal estimand (Imbens and Rubin, 2015) (see Section 3.3 of (Bojinov and Shephard, 2019) for further discussions of Definition 5.1). We have that $\hat{\tau}_{t|t-1}, \hat{\sigma}^2_{t|t-1}$ are still (conditionally) unbiased estimators for $\tau_t(w^{obs}_{1:(t-1)})$ and the (upper bound) variance of $\hat{\tau}_{t|t-1}$, respectively.

Lemma 5.1 (Mean and variance)**.** Under Assumption 1, we have that $E(\hat{\tau}_t \mid \mathcal{F}_{t-1}) = \tau_t(w^{obs}_{1:(t-1)})$ and that $\text{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1}) \leq E(\hat{\sigma}_{t|t-1} \mid \mathcal{F}_{t-1})$, where with a slight abuse of notation $\mathcal{F}_{t-1}$ is the sigma algebra containing all possible potential outcomes $Y_{1:T}(\bullet)$ and observed data $\{W_j, Y_j\}_{j=1}^{t-1}$ up to time $t-1$.

The proof is provided in Appendix A. Like Assumption 2, we similarly assume these new potential outcomes (with potential carryover effects) are bounded.

Assumption 6 (Bounded Potential Outcomes Under Carryover Effects)**.**

$$|Y_t(w_{1:t})| \leq M$$

for all t and any $w_{1:t} \in \{0,1\}^t$, where $M \in \mathbb{R}$.

Theorem 5.2 (Design-based Asymptotic Confidence Sequences for the Contemporaneous Treatment Effect with Carryover Effects)**.** Suppose $\{W_t, Y_t\}_{t=1}^T$ are observed for arbitrary data dependent stopping time T, where Assumptions 4- 6[3] are satisfied. Denote $\bar{\tau}_t(w^{obs}_{1:(t-1)}) := \frac{1}{t}\sum_{j=1}^t \tau_j(w^{obs}_{1:(j-1)})$ as the running mean of the contemporaneous treatment effect. Then,

$$\frac{1}{t}\sum_{j=1}^{t} \hat{\tau}_{j|j-1} \pm \frac{1}{t}\sqrt{\frac{S_{t|t-1}\eta^2 + 1}{\eta^2} \log\left(\frac{S_{t|t-1}\eta^2 + 1}{\alpha^2}\right)}$$

forms a valid $(1-\alpha)$ asymptotic confidence sequence for $\bar{\tau}_t(w^{obs}_{1:(t-1)})$ with approximation rate $o\left(\sqrt{\tilde{S}_{t|t-1} \log \tilde{S}_{t|t-1}}/t\right)$ for any pre-specified constant $\eta > 0$, where $S_{t|t-1}$ is defined in Corollary 4.1.

[3]In Assumptions 4- 6 all statements are respect to filtration defined in Lemma 5.1 and subscript i is replaced with t.

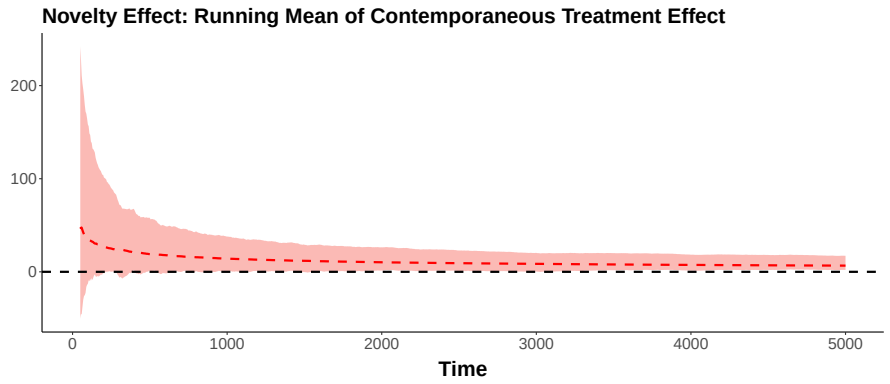


Figure 5: Novelty Effect of Message Alerting Treatment (Example 4). The red contours show the lower and upper confidence sequence for the time varying novelty treatment effect of message alerts at $\alpha = 0.05$ using Theorem 5.2. The red dotted line represents the true time varying treatment effect that diminishes to zero and the black horizontal dotted line represents the zero (null) line.

The confidence sequence in Theorem 5.2 is identical to that in Corollary 4.1 except the confidence sequence covers the running mean of the contemporaneous treatment effect as opposed to the average treatment effect and the notation is defined with respect to t as opposed to i to conceptually illustrate the different settings. The proof is in Appendix D, where this proof proves all asymptotic design-based confidence sequences provided in this paper. For example, Theorem 4.1 can be recovered by assuming that all treatment assignments are independent and there are no carryover effects. Corollary 4.1 can also be recovered by assuming away carryover effects. We choose to present it this way so that readers can understand the wide variety of problems our confidence sequence can tackle. Lastly, although all examples thus far have considered stationary treatment effects, i.e., treatment effects that do not vary by time, our confidence sequences can cover time varying causal effects. We demonstrate this in the following example.

Example 4 (Novelty Effect of Message Alerting Treatment)**.** Suppose at each time t the treatment is to alert the one user in our study with a message to engage with a product. Suppose the user's engagement increases substantially in the beginning based on the new treatment but the treatment effect diminishes over time (to zero) because the user has grown used to the treatment (often known as novelty effect). Furthermore, we assume that the message alerting treatment is only effective if the user did not receive a message alert at the previous time. Formally, we have that

$$Y_t(w_{t-1} = 0, w_t = 1) = Y_t(w_{t-1} = 0, w_t = 0) + 500/\sqrt{t}$$

and $Y_t(w_{t-1} = 0, w_t = 0) = Y_t(w_{t-1} = 1, w_t = 0) = Y_t(w_{t-1} = 1, w_t = 1) \sim N(25, 10^2)$. We exaggerate the initial treatment effect to start at 500, which decreases with order $1/\sqrt{t}$, to clearly show the time varying effect. Figure 5 shows that the confidence sequence uniformly covers the true time varying running mean of the contemporaneous treatment effect (in red) at all times. Although the analyst would reject the null effect at approximately $t = 200$, the analyst would also see a diminishing time varying treatment effect had the analyst suspected a novelty effect and continued running the experiment.

5.2 Panel data setting

5.2.1 Setting and notation

The above theorems and results are in the context of time series experiment. In practice, many organizations run multiple time series experiments for all n fixed units, where we observe multiple responses $Y_{i,t}$ for each unit $i = 1, 2, \ldots, n$ (Bojinov, Rambachan and Shephard, 2021). In this setting, we also observe the treatment assignment $W_{i,t}$ for each unit i across time t, where as before $w_{i,t}$ is a realization of this random variable and $w_{i,1:t}$ denotes the entire treatment path for unit i until time t. Although our above setting allows for any general carryover effects, in this setting we assume that each unit's potential outcome is only a function of its own treatment assignment path. More formally,

Assumption 7 (Independence of Potential Outcome Across Units)**.** $Y_{i,t}(w_{-i,1:t}, w_{i,1:t}) = Y_{i,t}(w'_{-i,1:t}, w_{i,1:t})$ for all $w_{-i,1:t}, w'_{-i,1:t}$, $i = 1, 2, \ldots, n$, and $t = 1, 2, \ldots, T$, where $w_{-i,1:t}$ denotes the treatment assignment paths for all units up to time t except unit i.

Concretely, Assumption 7 states that one Netflix member's treatment does not impact another, likely disconnected, Netflix member's outcome. This would be violated, for example, if units in the experiment communicate about their Netflix experience to each other. This assumption allows us to denote each unit's potential outcome as $Y_{i,t}(w_{i,1:t})$, where we again assume the potential outcome is also not a function of the future treatment assignments. As before, we denote the entire collection of potential outcome for all units at time t as

$$Y_{n,t}(\bullet) = \{Y_{1,t}(w_{1,1:t}), Y_{2,t}(w_{2,1:t}), \ldots, Y_{n,t}(w_{n,1:t}) : w_{i,1:t} \in \{0,1\}^t\}$$

and consequently define the entire collection of potential outcome for all units up to time t as $Y_{n,1:t}(\bullet) = \{Y_{n,1}(\bullet), Y_{n,2}(\bullet), \ldots, Y_{n,t}(\bullet)\}$.

The new filtration $\mathcal{F}_{n,t}$ denotes the sigma algebra containing the information set of all treatment assignment $W_{i,1:t}$ and observed outcome $Y_{i,t}$ up to time t for all units i. Furthermore, the filtration always contains the entire potential outcome set $Y_{n,1:T}(\bullet)$. Similar to Assumptions 4 and 6, we assume all the potential outcomes are bounded and the treatment assignments are bounded away from zero or one, which we now state under one assumption.

Assumption 8 (Bounded Potential Outcomes and Treatment Assignment in Panel Data Setting)**.**

$$|Y_{i,t}(w_{i,1:t})| \leq M, \quad 0 < p_{i,t|t-1}(w) := \Pr(W_{i,t} = w \mid \mathcal{F}_{n,t-1}) < 1,$$

for $M \in \mathbb{R}$, all $i = 1, 2, \ldots, n$, $t = 1, 2, \ldots, T$, and $w_{i,1:t} \in \{0,1\}^t$.

Assumption 8 allows the experimenter to adapt the treatment assignments for unit i not only as a function of the previous unit i's treatment assignment and response but also all the other units' treatment assignment and response. Lastly, since we have multiple users n, our causal estimand changes to the contemporaneous treatment effect averaged over n units.

$$\tau_{n,t}(w^{obs}_{1:n,1:(t-1)}) := \frac{1}{n}\sum_{i=1}^{n} Y_{i,t}(w^{obs}_{i,1:(t-1)}, 1) - Y_{i,t}(w^{obs}_{i,1:(t-1)}, 0), \tag{8}$$

where $w^{obs}_{1:n,1:(t-1)}$ denotes the entire observed treatment path for all units $i = 1, 2, \ldots, n$ up to time $t-1$. As a reminder, all causal estimands in Equation (8) and Definition 5.1 also capture quantities directly related to the experimental sample, allowing managers to perform efficient risk mitigation.

5.2.2 Aggregation

One naive, but powerful, approach to sequentially test for $\tau_{n,t}(w^{obs}_{1:n,1:(t-1)})$ is to simply "stack" or aggregate the data and pretend there is one single time series of nT observations.

To illustrate this, consider our full panel data matrix

$$\underbrace{\left.\begin{bmatrix} Y_{1,1} & Y_{1,2} & \dots & Y_{1,T} \\ Y_{2,1} & Y_{2,2} & \dots & Y_{2,T} \\ \vdots & \vdots & \vdots & \vdots \\ Y_{n,1} & Y_{n,2} & \dots & Y_{n,T} \end{bmatrix}\right\}\text{Units}}_{\text{Time}}$$

where the rows denote time horizon and the columns denote the n units. At each time t, we observe the t^{th} column of the above matrix. The naive aggregation method concatenates all columns from the above matrix into one row and pretends the data comes from one time series

$$(\underbrace{Y_{1,1}, \dots, Y_{n,1}}_{t=1}, \underbrace{Y_{1,2}, \dots, Y_{n,2}}_{t=2}, \dots, \underbrace{Y_{1,T}, \dots, Y_{n,T}}_{t=T}).$$

With this aggregation, we observe $\tilde{t} = 1, 2, \dots, nT$ time points and apply Theorem 5.2 on the above aggregated time series. Because at each time t we simultaneously observe n observations, the width of the confidence sequence further decreases with n. Since many organizations typically have large number of experimental units, the confidence sequence width can be small even at very early times t and the asymptotics more credible. To formalize this, we further denote the aggregated causal estimates and variance estimates as

$$\hat{\tau}_{\bullet,t|t-1} = \frac{1}{n}\sum_{i=1}^{n} \hat{\tau}_{i,t|t-1}, \qquad \hat{\sigma}^2_{\bullet,t|t-1} = \sum_{i=1}^{n} \hat{\sigma}^2_{i,t|t-1},$$

respectively, where

$$\hat{\tau}_{i,t|t-1} := \frac{\mathbb{1}\{W_{i,t}=1\}Y_{i,t}}{p_{i,t|t-1}(1)} - \frac{\mathbb{1}\{W_{i,t}=0\}Y_{i,t}}{p_{i,t|t-1}(0)}, \quad \hat{\sigma}^2_{i,t|t-1} := \frac{\mathbb{1}\{W_{i,t}=1\}Y^2_{i,t}}{p_{i,t|t-1}(1)^2} + \frac{\mathbb{1}\{W_{i,t}=0\}Y^2_{i,t}}{p_{i,t|t-1}(0)^2}$$

are the corresponding individual level counterparts of $\hat{\tau}_{t|t-1}$ and $\hat{\sigma}_{t|t-1}$, respectively. Finally, we also have the corresponding assumption for Assumption 5.

Assumption 9 (None Vanishing Variance for Panel Data Setting)**.** Let $\text{Var}(\hat{\tau}_{i,t|t-1} \mid \mathcal{F}_{n,t-1}) \le \sigma^2_{i,t|t-1}$, where $\sigma^2_{i,t|t-1}$ is equivalent to $\sigma^2_{t|t-1}$ (defined in Assumption 5) except it is defined for individual level potential outcomes and adaptive probability assignments. Then we assume that

$$\frac{1}{\sum_{j=1}^{t}\sigma^2_{\bullet,j|j-1}} = o(1) \iff \tilde{S}_{\bullet,t|t-1} := \sum_{j=1}^{t}\sigma^2_{\bullet,j|j-1} \xrightarrow{t\to\infty} \infty \text{ almost surely,}$$

where $\sigma^2_{\bullet,t|t-1} = \sum_{i=1}^{n}\sigma^2_{i,t|t-1}$.

Theorem 5.3 (Design-Based Confidence Sequence for the Average Contemporaneous Treatment Effect for Panel Data)**.** Suppose $\{W_{i,t}, Y_{i,t}\}_{i,t}$ are observed for all units $i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$ for any arbitrary data dependent stopping time T, where Assumptions 7-9 are satisfied. Denote $\bar{\tau}_{n,t}(w^{obs}_{1:n,1:(t-1)}) := \frac{1}{t}\sum_{j=1}^{t} \tau_{n,j}(w^{obs}_{1:n,1:(j-1)})$ as the running mean for the average contemporaneous treatment effect defined in Equation (8) and $S_{\bullet,t|t-1} := \sum_{j=1}^{t} \hat{\sigma}^2_{\bullet,j|j-1}$. Then,

$$\frac{1}{t}\sum_{j=1}^{t} \hat{\tau}_{\bullet,j|j-1} \pm \frac{1}{tn}\sqrt{\frac{S_{\bullet,t|t-1}\eta^2 + 1}{\eta^2} \log\left(\frac{S_{\bullet,t|t-1}\eta^2 + 1}{\alpha^2}\right)}$$

forms a valid $(1 - \alpha)$ asymptotic confidence sequence for $\bar{\tau}_{n,t}(w^{obs}_{1:n,1:(t-1)})$ with approximation rate $o\left(\sqrt{\tilde{S}_{\bullet,t|t-1} \log \tilde{S}_{\bullet,t|t-1}}/tn\right)$ and for any pre-specified constant $\eta > 0$.

The proof is omitted because under the stated assumptions it is identical to that of Theorem 5.2 replacing time with $\tilde{t} = 1, 2, \ldots, nT$ and stacking the panel data into one time series. Theorem 5.3 shows that our confidence sequence width further decrease with $1/n$.

One advantage of the aggregation approach is that it can also account for units that enter the experiment at different times. Although for clarity we present the results when there are exactly the same n units at each time t, our method also accommodates staggered adoption when the total number of units n are different at each time t, i.e., units can both be entering and leaving at any time point t. In such a case the causal estimand will also consequently change according to the different units that enter and exit. We present an example using Theorem 5.3 in Example 5.

6 Variance reduction technique via proxy outcomes

In the previous sections, our confidence sequences were constructed with estimators that only leverage the observed treatment and outcomes. In practice, there may be available covariates X or prior information that an analyst can leverage to further reduce the confidence sequence width through the use of proxy outcomes.

Note that, although we present our results in the time series and panel experiment setting, they immediately apply to the independent units setting studied in Section 4.

6.1 Proxy outcomes in single time series experiment

We first illustrate how to incorporate covariates in the general case for the single time series experiment (Section 5.1). We first denote $\hat{f}_t(X_{1:T}, Y_{1:(t-1)}, W_{1:(t-1)})$ to be the prediction for Y_t using any available covariate information $X_{1:T} = (X_1, X_2, \ldots, X_T)$ (each X_t may be multi-dimensional) and our previous data (W, Y) until time $t - 1$. For brevity, we define $\hat{Y}_{t|t-1} := \hat{f}_t(X_{1:T}, Y_{1:(t-1)}, W_{1:(t-1)})$. Since our prediction is based only on a function of our filtration $\mathcal{F}_{X,t-1}$, defined by enriching $\mathcal{F}_{t-1}$ to include the information set for all covariate $X_{1:T}$, (Bojinov and Shephard, 2019) defines $\hat{Y}_{t|t-1}$ as a "time series proxy outcome".

The prediction $\hat{Y}_{t|t-1}$ and the filtration are functions of *all* covariate information $X_{1:T}$, including covariate information not necessarily available at time $t < T$. We present it this way to clarify two points. First, in scenarios such

as Example 1, where the experiment involves N users, the covariate information for all the users may be known beforehand. For instance, in member-based experiments, the analyst may have access to the covariates of the existing members before the beginning of the experiment. In such cases, conditioning the covariates for all members at any time t accurately reflects the aforementioned scenario. Second, there may be cases where the covariates change over time or the experiment's units are not initially known, such as in new member experiments. In such situations, the analyst may still want to use X_t to create the prediction $\hat{Y}_t$ for Y_t To allow this, we allow both the prediction and the filtration to condition on all X_t, including the future, similar to how we always condition on all the potential outcomes. In practice, however, the analyst will typically only leverage the available covariates, $X_{1:t}$, for predicting Y_t at time t.

Using this proxy outcome, our causal estimand in Definition 5.1 can be rewritten as,

$$\tau_t(w^{obs}_{1:(t-1)}) = \{Y_t(w^{obs}_{1:(t-1)}, 1) - \hat{Y}_{t|t-1}\} - \{Y_t(w^{obs}_{1:(t-1)}, 0) - \hat{Y}_{t|t-1}\}.$$

Using the above formulation, the corresponding estimator using the proxy outcome is

$$\hat{\tau}^X_{t|t-1} := \frac{\mathbb{1}\{W_t = 1\}\{Y_t - \hat{Y}_{t|t-1}\}}{p_{t|t-1}(1)} - \frac{\mathbb{1}\{W_t = 0\}\{Y_t - \hat{Y}_{t|t-1}\}}{p_{t|t-1}(0)}, \tag{9}$$

with a similar upper-bound estimate of the variance of

$$\hat{\gamma}^2_{t|t-1} := \frac{\mathbb{1}\{W_t = 1\}\{Y_t - \hat{Y}_{t|t-1}\}^2}{p_{t|t-1}(1)^2} + \frac{\mathbb{1}\{W_t = 0\}\{Y_t - \hat{Y}_{t|t-1}\}^2}{p_{t|t-1}(0)^2}, \tag{10}$$

respectively. One can directly see that $\hat{\tau}^X_{t|t-1}$ is again unbiased for $\tau_t(w^{obs}_{1:(t-1)})$ conditional on $\mathcal{F}_{X,t-1}$, where conditioning on all the $X_{1:T}$ does not harm our inference because $\hat{Y}_{t|t-1}$ still remains a constant conditional on the filtration (hence the predictions are constructed only using past data without using current W_t). This allows the analyst to formally use the proxy outcome $\hat{Y}_{t|t-1}$ to incorporate any machine learning algorithm or prior knowledge to reduce the variance. This reduction is proportional to how small $\{Y_t - \hat{Y}_{t|t-1}\}^2$ is, i.e., how well the analyst can use the prior data to predict the next response.

Furthermore, we can also allow treatment assignment probabilities $p_{t|t-1}(w)$, defined in Assumption 4, to depend on the covariates since we always condition on $\mathcal{F}_{X,t-1}$. Lastly, we also require that the new variances with the proxy outcome do not disappear and that the proxy outcomes do not output infinity so that our new outcome $Y_t - \hat{Y}_{t|t-1}$ remains bounded.

Assumption 10 (None Vanishing Variance with Proxy Outcomes)**.** $\text{Var}(\hat{\tau}^X_{t|t-1} \mid \mathcal{F}_{X,t-1}) \leq \gamma^2_{t|t-1}$, where $\gamma^2_{t|t-1}$ is provided in Appendix A Equation (12) and similar to the expression in Equation (7) except we replace each potential outcome in the numerator with $Y_t(w^{obs}_{1:(t-1)}, \bullet) - \hat{Y}_{t|t-1}$. Then we assume that

$$\frac{1}{\sum_{j=1}^t \gamma^2_{j|j-1}} = o(1) \iff \tilde{S}^X_{t|t-1} := \sum_{j=1}^t \gamma^2_{j|j-1} \xrightarrow{t\to\infty} \infty \text{ almost surely.}$$

Further, we have that the predictions do not return infinity, i.e., $|\hat{Y}_{t|t-1}| \leq M'$ for all t and $M' \in \mathbb{R}$.

Theorem 6.1 (Design-based Asymptotic Confidence Sequence Using Proxy Outcomes)**.** Suppose $\{W_t, Y_t, X_t\}_{t=1}^T$ are observed for arbitrary data dependent stopping time T, where Assumptions 4[4], 6, and 10 are satisfied. Let $S^X_{t|t-1} := \sum_{j=1}^t \hat{\gamma}^2_{j|j-1}$. Then,

$$\frac{1}{t}\sum_{j=1}^{t} \hat{\tau}^X_{j|j-1} \pm \frac{1}{t}\sqrt{\frac{S^X_{t|t-1}\eta^2+1}{\eta^2}\log\left(\frac{S^X_{t|t-1}\eta^2+1}{\alpha^2}\right)}$$

forms a valid $(1-\alpha)$ asymptotic confidence sequence for the running mean of the contemporaneous treatment effect $\bar{\tau}_t(w^{obs}_{1:(t-1)})$ with approximation rate $o\left(\sqrt{\tilde{S}^X_{t|t-1}\log \tilde{S}^X_{t|t-1}}/t\right)$ for any pre-specified constant $\eta > 0$.

The proof is omitted because the setting as well as the assumptions are identical to that of Theorem 5.2 except we replace $S_{t|t-1}$ with $S^X_{t|t-1}$ and introducing $\hat{Y}^X_{t|t-1}$ is equivalent to changing the (fixed) potential outcome $Y_t(w^{obs}_{1:(t-1)}, \bullet)$ to a new constant $Y_t(w^{obs}_{1:(t-1)}, \bullet) - \hat{Y}^X_{t|t-1}$. Since the entire proof always conditions on the filtration, the proof remains identical. However, this difference allows the confidence sequence in Theorem 6.1 to incorporate covariates and other prior information to potentially reduce the confidence sequence width. Furthermore, our covariates X_t can contain both pre-treatment covariates that do not evolve over time, e.g., user sex and race, browser and device type, etc., and time varying covariates that can evolve over time (even as a function of previous treatment assignments making it a post-treatment confounder). Additionally, even if there are no available covariates, one could still likely reduce variance by using the sample mean of $Y_{1:(t-1)}$ for $\hat{Y}_{t|t-1}$, thus making Theorem 6.1 a useful practical extension for many cases.

Although we present Theorem 6.1 in the general time series setting with carryover effects, the same variance reduction technique also extends to non-time series setting in Theorem 4.1 and Corollary 4.1. In the aforementioned setting, Theorem 6.1 would be identical except all expressions with subscript t are replaced with n, allowing the proxy outcome $\hat{f}_{n|n-1}$ to instead predict the next user's response as a function of all the previous data and available covariate information.

6.2 Proxy outcomes in panel data setting

The variance reduction technique via proxy outcomes is most applicable in the panel data setting because we can leverage common information shared across n users (as opposed to only one user) to make predictions for the next time point. Although generalizing the results in Section 6.1 to the panel data setting is straightforward, we formalize this for completeness.

Starting in a similar fashion, we can rewrite the average contemporaneous treatment effect as

$$\tau_{n,t}(w^{obs}_{1:n,1:(t-1)}) = \left\{\frac{1}{n}\sum_{i=1}^{n}(Y_{i,t}(w^{obs}_{i,1:(t-1)}, 1) - \hat{Y}_{i,t|t-1})\right\} - \left\{\frac{1}{n}\sum_{i=1}^{n}(Y_{i,t}(w^{obs}_{i,1:(t-1)}, 0) - \hat{Y}_{i,t|t-1})\right\},$$

where $\hat{Y}_{i,t|t-1}$ is the prediction for the i^{th} individual's outcome at time t as a function of $\{W_{i,1:(t-1)}, X_{i,1:T}, Y_{i,1:(t-1)}\}_{i=1}^n$ and $X_{i,t}$ is the (multivariate) covariate value(s) for individual i at time t. We denote $\mathcal{F}_{X,n,t-1}$ as the sigma algebra containing $\{W_{i,1:(t-1)}, X_{i,1:T}, Y_{i,1:(t-1)}\}_{i=1}^n$ and all potential outcomes for all n units up to time T.

[4]In Assumption 4, we allow our adaptive probability treatment assignments to adapt to the covariate values X by replacing $\mathcal{F}_{t-1}$ with $\mathcal{F}_{X,t-1}$ for this theorem.

For the panel data setting, the corresponding estimators using the proxy outcome are

$$\hat{\tau}^X_{\bullet,t|t-1} = \frac{1}{n}\sum_{i=1}^{n} \hat{\tau}^X_{i,t|t-1}, \qquad \hat{\gamma}^2_{\bullet,t|t-1} = \sum_{i=1}^{n} \hat{\gamma}^2_{i,t|t-1},$$

respectively, where

$$\hat{\tau}^X_{i,t|t-1} := \frac{\mathbb{1}\{W_{i,t}=1\}\{Y_{i,t}-\hat{Y}_{i,t|t-1}\}}{p_{i,t|t-1}(1)} - \frac{\mathbb{1}\{W_{i,t}=0\}\{Y_{i,t}-\hat{Y}_{i,t|t-1}\}}{p_{i,t|t-1}(0)}$$
$$\hat{\gamma}^2_{i,t|t-1} := \frac{\mathbb{1}\{W_{i,t}=1\}\{Y_{i,t}-\hat{Y}_{i,t|t-1}\}^2}{p_{i,t|t-1}(1)^2} + \frac{\mathbb{1}\{W_{i,t}=0\}\{Y_{i,t}-\hat{Y}_{i,t|t-1}\}^2}{p_{i,t|t-1}(0)^2}.$$

Similar to Assumption 10, we have the following assumption.

Assumption 11 (None Vanishing Variance with Proxy Outcome Variances for Panel Data)**.** Let $\mathrm{Var}(\hat{\tau}^X_{i,t|t-1} \mid \mathcal{F}_{X,n,t-1}) \leq \gamma^2_{i,t|t-1}$, where $\gamma^2_{i,t|t-1}$ is equivalent to $\gamma^2_{t|t-1}$ (defined in Assumption 10) except for individual level potential outcomes and adaptive probability assignments. Then we assume that

$$\frac{1}{\sum_{j=1}^{t}\gamma^2_{\bullet,j|j-1}} = o(1) \iff \tilde{S}^X_{\bullet,t|t-1} := \sum_{j=1}^{t}\gamma^2_{\bullet,j|j-1} \xrightarrow{t\to\infty} \infty \text{ almost surely,}$$

where $\gamma^2_{\bullet,t|t-1} = \sum_{i=1}^{n}\gamma^2_{i,t|t-1}$. Further, we have that the predictions do not return infinity, i.e., $|\hat{Y}_{i,t|t-1}| \leq M'$ for all i, t and $M' \in \mathbb{R}$.

Theorem 6.2 (Design-based Asymptotic Confidence Sequence Using Proxy Outcomes for Panel Data)**.** Suppose $\{W_{i,t}, Y_{i,t}, X_{i,t}\}_{i,t}$ are observed for arbitrary data dependent stopping time T and $i = 1, 2, \ldots, n$ users, where Assumptions 7, 8[5], and 11 are satisfied. Let $S^X_{\bullet,t|t-1} := \sum_{j=1}^{t}\hat{\gamma}^2_{\bullet,j|j-1}$. Then,

$$\frac{1}{t}\sum_{j=1}^{t}\hat{\tau}^X_{\bullet,j|j-1} \pm \frac{1}{tn}\sqrt{\frac{S^X_{\bullet,t|t-1}\eta^2+1}{\eta^2}\log\left(\frac{S^X_{\bullet,t|t-1}\eta^2+1}{\alpha^2}\right)}$$

forms a valid $(1-\alpha)$ asymptotic confidence sequence for the running mean of the average contemporaneous treatment effect $\bar{\tau}_{n,t}(w^{obs}_{1:n,1:(t-1)})$ with approximation rate $o\left(\sqrt{\tilde{S}^X_{\bullet,t|t-1}\log\tilde{S}^X_{\bullet,t|t-1}}/t\right)$ for any pre-specified constant $\eta > 0$.

Comparing Theorem 6.2 and Theorem 5.3, we see that we can again get a reduction in variance depending on how small we can make $\{Y_{i,t} - \hat{Y}_{i,t|t-1}\}^2$, i.e., how well we can predict the next response for each individuals at every time using past data. We end this section with an example demonstrating Theorem 5.3 and the consequent reduction in variance we can get using Theorem 6.2.

Example 5 (Unit Varying Treatment Effects in Linear Models)**.** Suppose that every individual has an engagement score that is a function of the previous day's engagement and covariate X. Further suppose that each user has a user-specific

[5]In Assumption 8, we further allow our adaptive probability treatment assignments to adapt to covariate values X by replacing $\mathcal{F}_{n,t-1}$ with $\mathcal{F}_{X,n,t-1}$ for this theorem.

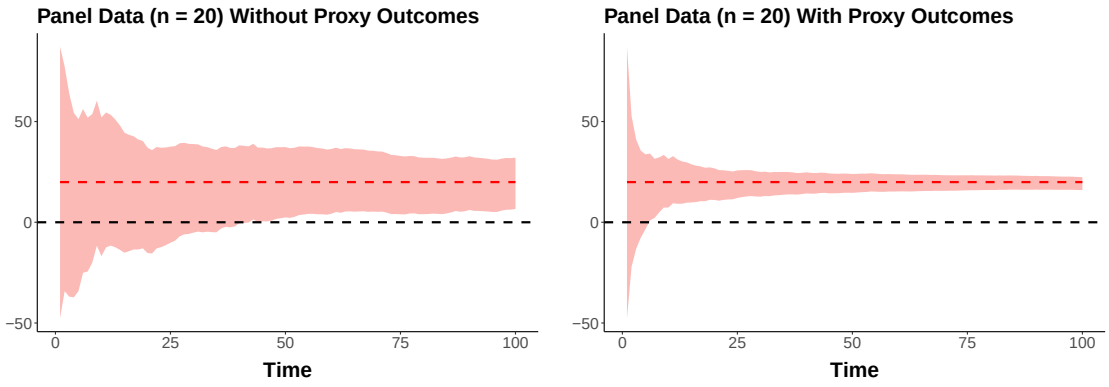


Figure 6: Unit Varying Treatment Effects in Linear Models (Example 5). The red contours show the lower and upper confidence sequence for average treatment effect for the 20 individuals in our sample at $\alpha = 0.05$ using Theorem 5.3 (left panel) and Theorem 6.2 (right panel). The parameters for this example are $n = 20, T = 100, \beta = 1, \rho = 0.5, \mu = 20$. The red dotted line represents the true average treatment effect that diminishes to zero and the black horizontal dotted line represents the zero (null) line.

treatment effect. More formally,

$$
\begin{aligned}
&Y_{i,t}(0) = \rho Y_{i,t-1}(0) + \beta X_i + \epsilon_{i,t}, \quad |\rho| \leq 1, \quad \epsilon_{i,t} \overset{iid}{\sim} N(0, 10^2) \\
&Y_{i,t}(1) = Y_{i,t}(0) + \mu_i, \quad \mu_i \sim N(\mu, 10^2) \\
&Y_{i,0} \overset{iid}{\sim} \beta X_i + \epsilon_{i,0}, \quad X_i \overset{iid}{\sim} N(25, 5^2), \quad W_{i,t} \overset{iid}{\sim} \text{Bern}(0.5) \text{ for all } i, t
\end{aligned}
$$

This scenario reflects a practical example where each unit may react differently to the treatment, but overall the treatment effect changes the engagement level by μ on average. For simplicity, we use a time-invariant covariate X_i that also has a stationary relationship β across time. Lastly, all examples thus far have assumed the potential outcomes come from some independent distribution. Because the design-based approach conditions on the potential outcome, we can allow for any arbitrary dependence. In this case, we demonstrate it for an AR(1) process, where each user's current engagement is only dependent on previous engagement. We set $n = 20, T = 100, \beta = 1, \rho = 0.5, \mu = 20$ for this example.

For the left panel of Figure 6, we build the confidence sequence using Theorem 5.3 by pretending we do not observe X_i. For the right panel of Figure 6, we build the confidence sequence using Theorem 6.2, where $\hat{Y}^X_{i,t|t-1}$ is the predicted response from an ordinary least square regression of all responses Y on all covariates X available at time t. Consequently, as t grows, $\hat{\beta}$ becomes more accurate, allowing a reduction in variance proportional to $(\beta X_i)^2$. Figure 6 shows that using proxy outcomes substantially reduces the confidence sequence width. For example, the right panel would reject the null average treatment effect by $t = 6$ while the left panel would reject it by $t = 43$, approximately a seven times reduction.

7 Simulation study

7.1 Demonstrating theoretical properties

Although all Examples 1- 5 are technically simulations, they only demonstrate the properties for one confidence sequence, thus it is still unclear whether our proposed confidence sequences have the time-uniform type-1 error guarantee. In this section, we build upon the linear model in Example 5 with three goals. First, we show the empirical *uniform* type-1 error guarantees obtained from our design-based asymptotic confidence sequences for all times (even early times). Second, we show the expected reduction in time to detect a statistically significant effect when incorporating proxy outcomes. Third, we further show the expected reduction in time to detect an effect when incorporating proxy outcomes from a misspecified prediction model.

To achieve this, we first replicate the simulation in Example 5 for 5000 empirically computed confidence sequences and record the proportion of times it contained the true treatment effect for all times t and the average time it took to reject the point null of zero treatment effect, which we refer to as "average stopping time." For the second scenario, we adjust the linear model with the following nonlinear model

$$Y_{i,t}(0) = \rho Y_{i,t-1}(0) + |x \sin(x)| + \epsilon_{i,t} \tag{11}$$

while using the same OLS prediction so that the prediction model is misspecified.

Lastly, we also create a third scenario where we only have $n = 1$ unit, representing a single time series experiment. The asymptotics are more credible in the panel data setting because even at $t = 1$, there is effectively n time steps already. Consequently, we should expect the panel data setting to have strong type-1 error guarantees even at early times. For this reason, we show the results when $n = 1$ under the same linear model in Example 5, where we fix $X = 25, \mu = 20$ (we drop the subscript because there is only one unit) and the remaining scenario is identical. Although we could also use proxy outcomes for a single time series by again using the OLS estimator after $t \geq 3$ (so that the OLS estimator exists), we choose to not use the proxy outcome for brevity. We also only report the type-1 error and omit the average stopping time to facilitate comparison because the single time series with $n = 1$ will have a substantially larger stopping time compared to the other panel data settings with $n = 20$.

Table 1 show the results of the simulation. As expected, scenarios 1-2 show strong type-1 error control even at early times. The over conservative coverage is due to our variance estimate being an upper bound estimate. Further, we see approximately a seven times reduction in the average stopping time using proxy outcomes under scenario 1. Although the difference is not as substantial when the prediction model is misspecified for a highly non-linear outcome, we still roughly see a 15% decrease in the average stopping time. Lastly, the third scenario shows that even when $n = 1$, the time-uniform coverage guarantee holds for all time. Although the theory guarantees time-uniform coverage after a sufficiently large t, our simulations suggest the coverage is strong even at early times likely because our estimated variance is conservative and the confidence width is large at early times. Nevertheless, for a single time series, we recommend practitioners to start "peeking" after some initial t, e.g., $t \geq 10$, to allow the asymptotics to take effect.

Different Scenarios	Average Empirical Results (5000 MC Repetitions)	
	Type-1 Error	Stopping Time
Scenario 1: Without proxy outcome (linear model: $n = 20$)	0.002	36
Scenario 1: With well specified proxy Outcome (linear model: $n = 20$)	0.002	5.5
Scenario 2: Without proxy outcome (non-linear model: $n = 20$)	0.001	34
Scenario 2: With misspecified proxy outcome (non-linear model: $n = 20$)	0.001	29
Scenario 3: Single time series (linear model: $n = 1$)	0.010	NA

Table 1: The first two rows represent simulations under the same setting as that in Example 5. The third and fourth rows are also under the same setting except we change the potential outcome model to Equation (11). The last row shows the empirical type-1 error control when $n = 1$ under scenario one. The second column represents the empirical proportion of times each respective confidence sequence covers the true treatment effect for all times t when $\alpha = 0.05$. The third column represents the average time it took to reject the point null of zero treatment effect. The proportion and average are taken over 5000 Monte-Carlo simulated confidence sequences for each scenario.

7.2 Comparison to confidence intervals and hybrid approaches

We next aim to quantify the tradeoffs practitioners face when deciding whether to perform anytime-valid inference (confidence sequence), fixed-time standard inference (confidence interval), or a mixture of both. To facilitate this comparison, we additionally comparable confidence intervals and a naive interim hybrid approach under a similar simulation setting as above.

To construct valid confidence intervals, we use the design-based asymptotically valid confidence interval proposed in (Bojinov and Shephard, 2019). We choose this comparison because it is centered around the same inverse propensity score estimator and uses the same design-based variance estimator in Equation (5), allowing a fair comparison. Alternatively, practitioners may opt for a hybrid approach that is the middle ground of confidence intervals and confidence sequences such as group sequential test that conducts hypothesis testing for only a finite number of times (Gordon Lan and Demets, 1983; Wassmer and Brannath, 2016). This strategy uses multiple-testing correction procedures to achieve the uniform type-1 error control in Equation (2) for finite number of times. As suggested by (Gordon Lan and Demets, 1983), we use a Bonferroni-corrected confidence interval from (Bojinov and Shephard, 2019) and construct K confidence intervals at uniform times between $t = 1, \dots, T$. For example, if $T = 100$ and the hybrid approach tests at $K = 5$ uniformly spaced points, then we construct a Bonferroni-corrected confidence interval at $t = 20, 40, 60, 80, 100$.

For our simulation, we further simplify the data generating process in Example 5. First we set $\rho = \beta = 0$ to get rid of any structural dependency and covariates. Next, we set the treatment mean $\mu = 10$, reducing the signal by half and set $T = 100, n = 5$ for the length of the experiment and number of units, respectively. We report both the type-1 error and the average stopping time as similarly done in Table 1. However, to facilitate a fair comparison of "stopping time" with confidence intervals and the hybrid interim approach, we redefine stopping time as the fastest time to detect a positively significant effect assuming the experiment is always terminated at $T = 100$ (even if a positive effect is

	Average Empirical Results (5000 MC Repetitions)			
Method	Type-1 Error	Stopping Time	Width	Power
Confidence Interval	97%	100	4.91	95%
Hybrid Interim Approach	99%	33	6.45	94%
Confidence Sequence	98%	31	9.90	90%

Table 2: Simulations comparing confidence sequence and intervals under the setting in Example 5 with $\rho = \beta = 0, \mu = 10, T = 100, n = 5$. The first row shows the performance of asymptotically valid design-based confidence intervals proposed in (Bojinov and Shephard, 2019). The second row performs a hybrid interim approach by constructing the aforementioned valid confidence interval at five uniform finite times $t = 20, 40, 60, 80, 100$ and performing a test at each time. The last row constructs the confidence sequence using Theorem 6.1 that is valid for all times. The second and third column are defined similarly as the simulation in Table 1. The last two columns denote the average width at $t = 100$ and the statistical power of each method, respectively.

not detected). Since confidence intervals are tested only once at $T = 100$, the stopping time for confidence intervals is trivially 100. The stopping time for confidence sequences is the average times it stopped early before $t < 100$ and averaging all the times it stopped at $t = 100$ because it was unable to find a statistically significant effect.

We additionally report both the average width and statistical power. The average width is defined as the expected width of the confidence interval, sequence, or hybrid confidence interval at the final time $t = T$. The statistical power is the probability of rejecting the null treatment effect by time $t = T$. For example, if the statistical power of the confidence sequence is 90%, then by $t = T$ the confidence sequence on average rejects the null treatment effect with a 90% chance before the end of the experiment.

Table 2 shows that the confidence interval, as expected, has a higher statistical power than the confidence sequence. However, the confidence sequence, on average, can terminate much earlier by $t = 31$ rather than $t = 100$, reducing the total expected experimentation time ($T = 100$) by roughly 70%. While confidence sequences have a wider interval than confidence intervals by definition, we quantify this tradeoff by demonstrating that the confidence sequence width is penalized by approximately a factor of two compared to the width of a confidence interval. Finally, the hybrid interim approach ($K = 5$) performs worse than both the confidence interval in terms of power and width, but it may have the potential to stop earlier than confidence intervals while still having competitive stopping times with confidence sequences. Although this approach provides a suitable middle ground between the two extremes, the advantages of potentially stopping early is sensitive to the pre-specified finite-times. Furthermore, the analyst can not continue to run the experiment even if the analyst desires to after the final pre-determined time, posing similar issues to confidence intervals.

8 Application to Netflix experiments

8.1 Sign-up page experiments

A crucial component of any subscription-based business is designing a simple and efficient sign-up page that attracts new members. Consequently, Netflix runs thousands of experiments to adapt and improve the sign-up page while attempting to mitigate the experimental risk of losing potential members. In this section, we analyze two such experiments.

The first experiment tests whether providing users with verification and assistance when entering credit card information can help more prospective members complete the sign-up process. Netflix has observed that users often benefit from assistance when completing their subscriptions after incorrectly inputting credit card information— a well-known phenomenon studied in UX research (Holst, 2017). To improve the process, Netflix designed a new interface that automatically detects the brand of the card (e.g., Mastercard or Visa) after the first few numbers and automatically formats the card numbers (e.g., XXXX-XXXX-XXXX-XXXX) corresponding to the brand. To test the effectiveness of the new design, Netflix ran an experiment in a single country, comparing the new design against the standard offering without such assistance; we refer to this experiment/treatment as the "Credit Card Assistance" experiment/treatment.

Over approximately[6] two weeks, $N \approx 30,000$ visitors to the Netflix sign up page were randomly assigned (with equal probability) to either the treatment (the new sign-up page) or the control (the original sign-up page). The primary outcome was whether or not the individual successfully joined Netflix by providing a working payment method (binary outcome).

The second experiment tests whether clearly indicating that Netflix accepts prepaid gift cards drove additional sign-ups. Currently, Netflix accepts prepaid gift cards, but this is not indicated in the payment selection. Managers at Netflix believed that indicating this information clearly would help more prospective members join Netflix; we refer to this experiment as the "Prepaid Card" experiment.

The Prepaid Card experiment was conducted for approximately 2 months in Italy with a sample size of $N \approx 125,000$ total visitors to the Netflix sign up page that were either shown the default experience (control) or a different version where Netflix clearly indicates that they accept Postepay cards, a popular Italian prepaid card (treatment). The company used the standard Thompson adaptive sampling scheme as they believed that rapidly moving towards the best treatment would help maximize subscriptions (Thompson, 1933). Lastly, although there were two treatment versions, we focused on one for simplicity.

We re-analyze the two sign-up page experiments in Figure 7 using Theorem 4.1 for the first non-adaptive experiment (left panel) and Corollary 4.1 for the second adaptive experiment (right panel). The left panel plots the confidence sequence using only the first $n = 1000$ customers out of a total sample size of $N \approx 30,000$. We see that the confidence sequence detects a statistically significant negative treatment effect before the 70^{th} unit and a treatment effect at least as harmful as 5% by approximately the 80^{th} unit. This result is consistent with the finding at Netflix, where this experiment was known to have a negative treatment effect due to technical issues related to implementing the treatment for

[6]For the purposes of this paper, we omit the exact details of the data, such as precise sample sizes and test durations.

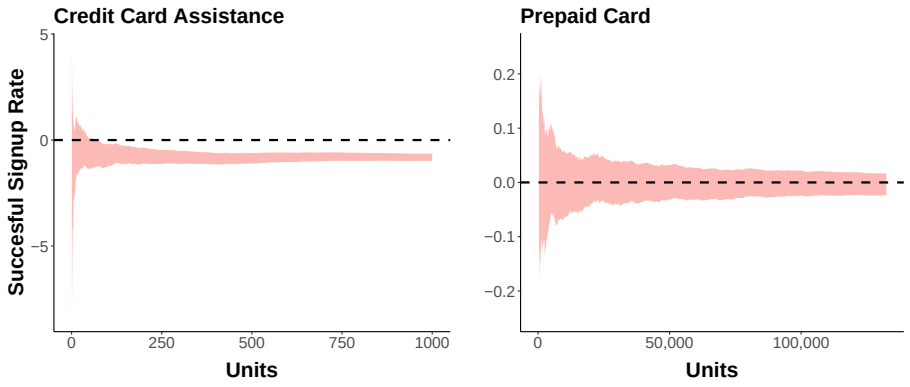


Figure 7: Sign-up page experiments. The left panel shows the confidence sequence using Theorem 4.1 for the Credit Card Assistance experiment. We plot the confidence sequence for the first $n \approx 1000$ units out of a total sample of $N \approx 30,000$. The right panel shows the confidence sequence using Corollary 4.1 to analyze the adaptive Prepaid Card experiment. The black horizontal dotted line represents the zero (null) line with $\alpha = 0.05$.

the Firefox browser. However, this experiment ran for approximately two weeks with this harmful treatment on over $N = 30,000$ customers with a final estimated treatment effect of approximately -20%[7]. The left panel of Figure 7 shows this could have been avoided had the business manager used confidence sequences and terminated the harmful experiment as early as the first day before incurring further costs.

The right panel shows the confidence sequence when analyzing the Prepaid Card experiment. We find that the confidence sequence covers zero for all units, showing that indicating that pre-paid cards are accepted does not lead to higher conversion rates with a point estimate of -0.0033. This result is also consistent with other findings, where Netflix ran another non-adaptive experiment for the same setting and found no significant treatment effect using a t-test with over $100,000$ samples. Suppose the analyst wanted to terminate the experiment as soon as there is evidence the treatment effect is less than ϵ. In that case, Figure 7 shows the analyst would terminate the experiment approximately as early as the $60,000^{\text{th}}$ unit for $\epsilon = 0.05$, i.e., the first unit where the analyst can confidently conclude the treatment effect is less than 5%. For $\epsilon = 0.02$, the analyst would approximately terminate at the $100,000^{\text{th}}$ unit.

8.2 Messaging treatment

We now reanalyze an experiment that tested whether or not sending a push messaging notification (for example, the notification could remind subscribers that a specific show has been released) increases engagement among Netflix members. To do this, we leverage Theorem 5.3 to build a confidence sequence due to the special panel data structure.

Netflix ran the experiment on $n \approx 2000$ members that have some probability of seeing a message or not (our binary treatment) on each day t. The probabilities were computed from a machine learning model (details omitted) that predicts whether each member would benefit from a message. Finally, the experiment was run for approximately one week, where it is known that sending messages this way has a strong positive impact on engagement. The primary

[7]The point estimate is an estimate for the average treatment effect τ_N, i.e., $\sum_{i=1}^{N} \tau_i / N \approx -0.20$

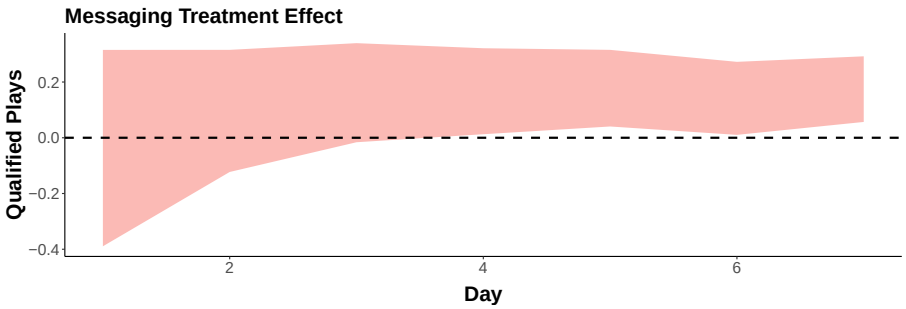


Figure 8: Messaging treatment application. The red contours show the lower and upper confidence sequence as approximately 2000 units continue to see messages or not throughout approximately one week at $\alpha = 0.05$ using Theorem 5.3. The black horizontal dotted line represents the zero (null) line. The confidence sequence covers zero at all times.

response of interest is "Qualified Play", an internal metrics that records if a user has played a video for a meaningful amount of time on that day. Figure 8 shows that even with only $n \approx 2000$ users, we see a positive treatment effect as early as the fourth day with a final estimated treatment effect of 0.17. This allows the manager to terminate the experiment earlier in this more complex panel data setting.

9 Concluding remarks

Our work bridges the sequential testing literature with the design-based literature to perform continuous monitoring for the average treatment effect, contemporaneous treatment effect relevant also to the bandit settings, and the average contemporaneous treatment effect in panel data settings with time-uniform guarantees. We summarize our contribution and a few key remarks in Table 3. Our confidence sequences formally allow managers to "peek" at any time and stop the experiment in a data-dependent way, e.g., as soon as detecting a statistically significant harmful effect. Additionally, the design-based approach allows managers to perform risk mitigation by quantifying an estimand directly relevant to the harm incurred on the obtained experimental sample in an assumption-light approach.

Acknowledgement

We thank Ian Waudby-Smith and participants of the 2022 Conference on Digital Experimentation for advice and feedback.

Setting of Confidence Sequence (CS)	**Remarks**
Exact design-based CS - ATE	Theorem 3.2
	Dependence on M; CS does not shrink to zero asymptotically
Asymptotic design-based CS - ATE	Theorem 4.1
	No dependence on M; CS shrinks to zero asymptotically
Asymptotic design-based CS - bandit settings	Corollary 4.1
	Adaptive treatment assignment
Asymptotic design-based CS - CTE	Theorem 5.2
	Single time series with carryover effects
Asymptotic design-based CS - panel setting	Theorem 5.3
	Panel data setting with n units observed across T time periods
Introduction of proxy outcomes	Theorem 6.1- 6.2
	Incorporates covariates or any modeling assumption to reduce CS width

Table 3: Summary of our contribution. The first column describes the confidence sequence of interest. The second column describes where to find the respective confidence sequence followed by key generalizations and remarks about the confidence sequence.

References

Abadie, Alberto, Susan Athey, Guido W Imbens and Jeffrey M Wooldridge. 2020. "Sampling-Based versus Design-Based Uncertainty in Regression Analysis." Econometrica 88:265–296.

Aronow, Peter M., Dean Eckles, Cyrus Samii and Stephanie Zonszein. 2020. "Spillover Effects in Experimental Data." ArXiv abs/2001.05444.

Bakshy, Eytan, Dean Eckles and Michael Bernstein. 2014. "Designing and Deploying Online Field Experiments." WWW 2014 - Proceedings of the 23rd International Conference on World Wide Web.

Basse, Guillaume and Edoardo Airoldi. 2017. "Limitations of Design-based Causal Inference and A/B Testing under Arbitrary and Network Interference." Sociological Methodology 48.

Bastani, Hamsa and Mohsen Bayati. 2020. "Online Decision Making with High-Dimensional Covariates." Operations Research 68:276–294.
URL: *https://doi.org/10.1287/opre.2019.1902*

Bojinov, Iavor, Ashesh Rambachan and Neil Shephard. 2021. "Panel experiments and dynamic causal effects: A finite population perspective." Quantitative Economics 12:1171–1196.

Bojinov, Iavor, David Simchi-Levi and Jinglong Zhao. 2020. "Design and Analysis of Switchback Experiments."

Management Science 0:null.
URL: *https://doi.org/10.1287/mnsc.2022.4583*

Bojinov, Iavor and Neil Shephard. 2019. "Time Series Experiments and Causal Estimands: Exact Randomization Tests and Trading." Journal of the American Statistical Association 114:1665–1682.
URL: *https://doi.org/10.1080/01621459.2018.1527225*

Bojinov, Iavor and Somit Gupta. 2022. "Online Experimentation: Benefits, Operational and Methodological Challenges, and Scaling Guide." Harvard Data Science Review 4. https://hdsr.mitpress.mit.edu/pub/aj31wj81.

Camuffo, Arnaldo, Alessandro Cordova, Alfonso Gambardella and Chiara Spina. 2020. "A Scientific Approach to Entrepreneurial Decision Making: Evidence from a Randomized Control Trial." Management Science 66:564–586.
URL: *https://doi.org/10.1287/mnsc.2018.3249*

Chow, Y. S. 1971. "On the L_p-Convergence." The Annals of Mathematical Statistics 42:393 – 394.
URL: *https://doi.org/10.1214/aoms/1177693530*

Cohen, Wesley M. and Daniel A. Levinthal. 1994. "Fortune Favors the Prepared Firm." Management Science 40:227–251.
URL: *https://doi.org/10.1287/mnsc.40.2.227*

Cohen, Wesley M., Richard R. Nelson and John P. Walsh. 2002. "Links and Impacts: The Influence of Public Research on Industrial R&D." Management Science 48:1–23.
URL: *https://doi.org/10.1287/mnsc.48.1.1.14273*

Cox, D. R. 1958. Planning of Experiments. New York: Wiley.

Darling, D. A. and Herbert Robbins. 1967. "CONFIDENCE SEQUENCES FOR MEAN, VARIANCE, AND MEDIAN." Proceedings of the National Academy of Sciences 58:66–68.
URL: *https://www.pnas.org/doi/abs/10.1073/pnas.58.1.66*

Ding, Peng, Avi Feller and Luke Miratrix. 2016. "Randomization inference for treatment effect variation." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 78:655–671.
URL: *https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/rssb.12124*

Einmahl, Uwe. 1987. "Strong Invariance Principles for Partial Sums of Independent Random Vectors." The Annals of Probability 15:1419 – 1440.
URL: *https://doi.org/10.1214/aop/1176991985*

Fan, Xiequan, Ion Grama and Quansheng Liu. 2015. "Exponential inequalities for martingales with applications." Electronic Journal of Probability 20:1 – 22.
URL: *https://doi.org/10.1214/EJP.v20-3496*

Farronato, Chiara, Alan MacCormack and Sarah Mehta. 2018. "Innovation at Uber: The Launch of Express POOL." Harvard Business School Case) 82.

Fisher, Ronald A. 1935. The Design of Experiments. London: Oliver and Boyd Chapter II.

Gordon Lan, K. K. and David L. Demets. 1983. "Discrete sequential boundaries for clinical trials." Biometrika 70:659–663.
URL: *https://doi.org/10.1093/biomet/70.3.659*

Harshaw, Christopher, Fredrik Sävje and Yitan Wang. 2022. "A Design-Based Riesz Representation Framework for Randomized Experiments.".

Holst, Christian. 2017. "The Credit Card Number Field Must Allow and Auto-Format Spaces (80% Don't)." Baymard Institute.

Horvitz, Daniel G. and D. J. Thompson. 1952. "A Generalization of Sampling Without Replacement from a Finite Universe." Journal of the American Statistical Association 47:663–685.

Howard, Steven R., Aaditya Ramdas, Jon D. McAuliffe and Jasjeet S. Sekhon. 2020. "Time-uniform, nonparametric, nonasymptotic confidence sequences." The Annals of Statistics.

Imbens, Guido W. and Donald B. Rubin. 2015. Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge University Press.

Johari, Ramesh, Leonid Pekelis and David Walsh. 2015. "Always Valid Inference: Bringing Sequential Analysis to A/B Testing.".

Johari, Ramesh, 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. KDD '17 New York, NY, USA: Association for Computing Machinery p. 1517–1525.
URL: *https://doi.org/10.1145/3097983.3097992*

Kohavi, Ron, 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. KDD '13 New York, NY, USA: Association for Computing Machinery p. 1168–1176.
URL: *https://doi.org/10.1145/2487575.2488217*

Kohavi, Ron, Diane Tang and Ya Xu. 2020. Trustworthy online controlled experiments: A practical guide to a/b testing. Cambridge University Press.

Komlos, Janos, Peter Major and G. Tusnády. 1976. "An approximation of partial sums of independent RV's, and the sample DF. II." Probability Theory and Related Fields 34:33–58.

Koning, Rembrand, Sharique Hasan and Aaron Chatterji. 2022. "Experimentation and Start-up Performance: Evidence from A/B Testing." Management Science 68:6434–6453.
URL: *https://doi.org/10.1287/mnsc.2021.4209*

Lei, Lihua and Peng Ding. 2020. "Regression adjustment in completely randomized experiments with a diverging number of covariates." Biometrika 108:815–828.
URL: *https://doi.org/10.1093/biomet/asaa103*

Li, Jimmy Q., Paat Rusmevichientong, Duncan Simester, John N. Tsitsiklis and Spyros I. Zoumpoulis. 2015. "The Value of Field Experiments." Management Science 61:1722–1740.
URL: *https://doi.org/10.1287/mnsc.2014.2066*

Lindon, Michael and Alan Malek. 2022. Anytime-Valid Inference For Multinomial Count Data. In Advances in Neural Information Processing Systems, ed. Alice H. Oh, Alekh Agarwal, Danielle Belgrave and Kyunghyun Cho.
URL: *https://openreview.net/forum?id=a4zg0jiuVi*

Lindon, Michael, Chris Sanden and Vaché Shirikian. 2022. Rapid Regression Detection in Software Deployments through Sequential Testing. KDD '22 New York, NY, USA: Association for Computing Machinery p. 3336–3346.
URL: *https://doi.org/10.1145/3534678.3539099*

Mao, Jialiang and Iavor Bojinov. 2021. "Quantifying the Value of Iterative Experimentation." arXiv preprint arXiv:2111.02334.

Maurer, Andreas and Massimiliano Pontil. 2009. Empirical Bernstein Bounds and Sample-Variance Penalization. In Annual Conference Computational Learning Theory.

Neyman, J., K. Iwaszkiewicz and St. Kolodziejczyk. 1935. "Statistical Problems in Agricultural Experimentation (with discussion)." Supplement of Journal of the Royal Statistical Society 2:107–180.

Ramdas, Aaditya, Johannes Ruf, Martin Larsson and Wouter M. Koolen. 2020. "Admissible anytime-valid sequential inference must rely on nonnegative martingales." arXiv: Statistics Theory.

Robbins, Herbert. 1952. "Some aspects of the sequential design of experiments." Bulletin of the American Mathematical Society 58:527 – 535.
URL: *https://doi.org/*

Robbins, Herbert. 1970. "Statistical Methods Related to the Law of the Iterated Logarithm." The Annals of Mathematical Statistics 41:1397 – 1409.
URL: *https://doi.org/10.1214/aoms/1177696786*

Robins, James. 1986. "A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect." Mathematical Modelling 7:1393–1512.
URL: *https://www.sciencedirect.com/science/article/pii/0270025586900886*

Rubin, Donald B. 1974. "Estimating causal effects of treatments in randomized and nonrandomized studies." Journal of Educational Psychology 66:688–701.

Rubin, Donald B. 1986. "Comment." Journal of the American Statistical Association 81:961–962.
URL: *https://doi.org/10.1080/01621459.1986.10478355*

Rusmevichientong, Paat and John N. Tsitsiklis. 2010. "Linearly Parameterized Bandits." Mathematics of Operations Research 35:395–411.
URL: *https://doi.org/10.1287/moor.1100.0446*

Russo, Daniel. 2020. "Simple Bayesian Algorithms for Best-Arm Identification." Operations Research 68:1625–1647.
URL: *https://doi.org/10.1287/opre.2019.1911*

Slivkins, Aleksandrs. 2019. "Introduction to Multi-Armed Bandits." Foundations and Trends® in Machine Learning 12:1–286.
URL: *http://dx.doi.org/10.1561/2200000068*

Strassen, Volker. 1967. Almost sure behavior of sums of independent random variables and martingales.

Sutton, Richard and Andrew Barto. 2018. Reinforcement learning: an introduction. Adaptive Computation and Machine Learning. MIT Press.

Thomke, Stefan H. 2020. Experimentation works: The surprising power of business experiments. Harvard Business Press.

Thompson, William R. 1933. "ON THE LIKELIHOOD THAT ONE UNKNOWN PROBABILITY EXCEEDS ANOTHER IN VIEW OF THE EVIDENCE OF TWO SAMPLES." Biometrika 25:285–294.
URL: *https://doi.org/10.1093/biomet/25.3-4.285*

Toulis, Panos and Edward Kao. 2013. Estimation of Causal Peer Influence Effects. In Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28. ICML'13 JMLR.org p. III–1489–III–1497.

Ville, Jean-Luc. 1939. Étude critique de la notion de collectif.

Wang, Hongjian and Aaditya Ramdas. 2022. Catoni-style confidence sequences for heavy-tailed mean estimation.

Wassmer, Gernot and Werner Brannath. 2016. Group Sequential and Confirmatory Adaptive Designs in Clinical Trials.

Waudby-Smith, Ian and Aaditya Ramdas. 2020a. Confidence sequences for sampling without replacement. In Advances in Neural Information Processing Systems, ed. H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan and H. Lin. Vol. 33 Curran Associates, Inc. pp. 20204–20214.

Waudby-Smith, Ian and Aaditya Ramdas. 2020b. "Estimating means of bounded random variables by betting." arXiv e-prints.

Waudby-Smith, Ian, David Arbour, Ritwik Sinha, Edward H. Kennedy and Aaditya Ramdas. 2021. "Time-uniform central limit theory with applications to anytime-valid causal inference.".

Waudby-Smith, Ian, Lili Wu, Aaditya Ramdas, Nikos Karampatziakis and Paul Mineiro. 2022. "Anytime-valid off-policy inference for contextual bandits.".
URL: *https://arxiv.org/abs/2210.10768*

Xiong, Ruoxuan, Susan Athey, Mohsen Bayati and Guido Imbens. 2019. "Optimal Experimental Design for Staggered Rollouts." SSRN Electronic Journal.

Zhao, Anqi and Peng Ding. 2021. "Reconciling design-based and model-based causal inferences for split-plot experiments.".

A Proof of Lemma 5.1

We prove it under the most general setting in Section 5.1, where we have adaptive probability treatment assignment $p_{t|t-1}(w)$ and carryover effects in the potential outcome. This proof is nearly identical to that in Appendix A of (Bojinov and Shephard, 2019) but we provide it under our setting for completeness.

$$E(\hat{\tau}_t \mid \mathcal{F}_{t-1}) = \tau_t(w^{obs}_{1:(t-1)})$$

Proof.

$$\begin{aligned}
E(\hat{\tau}_t \mid \mathcal{F}_{t-1}) &= E\left(\frac{\mathbb{1}\{W_t = 1\}Y_t}{p_{t|t-1}(1)} - \frac{\mathbb{1}\{W_t = 0\}Y_t}{p_{t|t-1}(0)} \mid \mathcal{F}_{t-1}\right) \\
&= \frac{p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 1)}{p_{t|t-1}(1)} - \frac{p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 0)}{p_{t|t-1}(0)} \\
&= \tau_t(w^{obs}_{1:(t-1)})
\end{aligned}$$

□

Next, we calculate the closed form expression of $\text{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1})$.

$$\begin{aligned}
\text{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1}) &= \text{Var}\left(\frac{\mathbb{1}\{W_t = 1\}Y_t}{p_{t|t-1}(1)} - \frac{\mathbb{1}\{W_t = 0\}Y_t}{p_{t|t-1}(0)} \mid \mathcal{F}_{t-1}\right) \\
&= \frac{p_{t|t-1}(1)p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)^2} + \frac{p_{t|t-1}(0)p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)^2} \\
&+ 2Y_t(w^{obs}_{1:(t-1)}, 1)Y_t(w^{obs}_{1:(t-1)}, 0) \\
&= \frac{p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)} + \frac{p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)} + 2Y_t(w^{obs}_{1:(t-1)}, 1)Y_t(w^{obs}_{1:(t-1)}, 0) \\
&= \frac{\left(p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1) + p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)\right)^2}{p_{t|t-1}(1)p_{t|t-1}(0)},
\end{aligned}$$

where the third line follows because $\text{Cov}(\mathbb{1}\{W_t = 1\}, \mathbb{1}\{W_t = 0\} \mid \mathcal{F}_{t-1}) = -p_{t|t-1}(1)p_{t|t-1}(0)$. Now we show that

$$\text{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1}) \leq \sigma_{t|t-1} := \frac{Y_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)} + \frac{Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)},$$

which would complete the proof because it is straight forward to show that $E(\hat{\sigma}_{t|t-1} \mid \mathcal{F}_{t-1}) = \sigma_{t|t-1}$.

Proof.

$$\begin{aligned}
\text{Var}(\hat{\tau}_{t|t-1} \mid \mathcal{F}_{t-1}) &= \frac{p_{t|t-1}(0)^2Y_t(w^{obs}_{1:(t-1)}, 1)^2 + p_{t|t-1}(1)^2Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(1)p_{t|t-1}(0)} \\
&+ \frac{2p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1)p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)}{p_{t|t-1}(1)p_{t|t-1}(0)} \\
&\leq \frac{Y_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)} + \frac{Y_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)},
\end{aligned}$$

where the last line follows because $(a-b)^2 \geq 0 \rightarrow a^2 + b^2 \geq 2ab$ and we let $a = p_{t|t-1}(0)Y_t(w^{obs}_{1:(t-1)}, 1)$ and $b = p_{t|t-1}(1)Y_t(w^{obs}_{1:(t-1)}, 0)$. □

Lastly, $\gamma^2_{t|t-1}$ defined in Assumption 10 can be obtained by replacing $Y_t(w^{obs}_{1:(t-1)}, \bullet)$ with a new "residualized" potential outcome $\tilde{Y}_t(w^{obs}_{1:(t-1)}, \bullet) := Y_t(w^{obs}_{1:(t-1)}, \bullet) - \hat{Y}_{t|t-1}$. Since all the proof is conditioned on the filtration, the above is equivalent to introducing a constant and therefore we have that

$$\gamma^2_{t|t-1} = \frac{\tilde{Y}_t(w^{obs}_{1:(t-1)}, 1)^2}{p_{t|t-1}(1)} + \frac{\tilde{Y}_t(w^{obs}_{1:(t-1)}, 0)^2}{p_{t|t-1}(0)} \tag{12}$$

B Proof of Theorem 3.2

We build off the proof of Theorem 4 in (Howard et al., 2020). As hinted in Section 3.2, we will first show that

$$\exp\left[\frac{\sum_{i=1}^{n}(\hat{\tau}_i - \tau_n)}{m(m+1)} + \frac{S_n}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right]$$

is a non-negative supermartingale with respect to the filtration $\mathcal{F}_{N,n-1}$. (Fan, Grama and Liu, 2015) show that

$$\exp\left(\lambda\kappa + \kappa^2(\lambda + \log(1-\lambda))\right) \leq 1 + \lambda\kappa$$

for $\kappa \geq -1$ and $\lambda \in [0, 1)$ from Fan's Inequality. We let

$$\kappa = \frac{\hat{\tau}_n}{m},$$

where $\kappa \geq -1$ since $|\hat{\tau}_i| \leq m$ for every i by Assumption 2. Therefore, we have

$$\begin{aligned}
\exp\left(\lambda\frac{\hat{\tau}_n}{m} + \frac{\hat{\tau}_n^2}{m^2}(\lambda + \log(1-\lambda))\right) &\leq 1 + \lambda\frac{\hat{\tau}_n}{m} \\
E\left[\exp\left(\frac{\lambda\hat{\tau}_n}{m} + \frac{\hat{\sigma}_n^2}{m^2}(\lambda + \log(1-\lambda))\right) \mid \mathcal{F}_{N,n-1}\right] &\leq 1 + \lambda\frac{\tau_n}{m} \\
E\left[\exp\left(\frac{\lambda(\hat{\tau}_n - \tau_n)}{m} + \frac{\hat{\sigma}_n^2}{m^2}(\lambda + \log(1-\lambda))\right) \mid \mathcal{F}_{N,n-1}\right] &\leq \exp(-\frac{\lambda\tau_n}{m})[1 + \lambda\frac{\tau_n}{m}] \\
E\left[\exp\left(\frac{\lambda(\hat{\tau}_n - \tau_n)}{m} + \frac{\hat{\sigma}_n^2}{m^2}(\lambda + \log(1-\lambda))\right) \mid \mathcal{F}_{N,n-1}\right] &\leq 1
\end{aligned}$$

where the second line follows because $\hat{\tau}_n^2 = \hat{\sigma}_n^2$ and Lemma 2.1 and the last line follows because $1 - x \leq \exp(-x)$. We plug $\lambda = 1/(m+1)$ and because the above is a non-negative quantity this directly implies that

$$\exp\left[\frac{\sum_{i=1}^{n}(\hat{\tau}_i - \tau_n)}{m(m+1)} + \frac{S_n}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right]$$

is indeed a non-negative super martingale with respect to $\mathcal{F}_{N,n-1}$ as desired with initial value less than one. We apply Lemma 3.1 and have that

$$\Pr\left(\exists n : \exp\left[\frac{\sum_{i=1}^{n}(\hat{\tau}_i - \tau_n)}{m(m+1)} + \frac{S_n}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right] \geq \frac{1}{\tilde{\alpha}}\right) \leq \tilde{\alpha}$$

$$\Pr\left(\exists n : \left[\frac{\sum_{i=1}^{n}(\hat{\tau}_i - \tau_n)}{m(m+1)} + \frac{S_n}{m^2}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right] \geq \log\left(\frac{1}{\tilde{\alpha}}\right)\right) \leq \tilde{\alpha}$$

$$\Pr\left(\exists n : \sum_{i=1}^{n}(\hat{\tau}_i - \tau_n) \geq m(m+1)\log\left(\frac{1}{\tilde{\alpha}}\right) - \frac{(m+1)S_n}{m}\left(\log\left(\frac{m}{m+1}\right) + \frac{1}{m+1}\right)\right) \leq \tilde{\alpha}$$

$$\Pr\left(\exists n : \sum_{i=1}^{n}(\hat{\tau}_i - \tau_n) \geq \left[m(m+1)\log\left(\frac{1}{\tilde{\alpha}}\right) + S_n\left(\frac{m+1}{m}\log\left(1+\frac{1}{m}\right) - \frac{1}{m}\right)\right]\right) \leq \tilde{\alpha}$$

Consequently, we have that

$$\Pr\left(\exists n : \frac{1}{n}\sum_{i=1}^{n}(\hat{\tau}_i - \tau_n) \geq \left[\frac{m(m+1)}{n}\log\left(\frac{1}{\tilde{\alpha}}\right) + \frac{S_n}{n}\left(\frac{m+1}{m}\log\left(1+\frac{1}{m}\right) - \frac{1}{m}\right)\right]\right) \leq \tilde{\alpha}$$

This gives the one-sided confidence sequence and we can do the same trick and build the same statement instead for $\kappa = -\hat{\tau}_n/m$. Then we get

$$\Pr\left(\exists n : -\frac{1}{n}\sum_{i=1}^{n}(\hat{\tau}_i - \tau_n) \geq \left[\frac{m(m+1)}{n}\log\left(\frac{1}{\tilde{\alpha}}\right) + \frac{S_n}{n}\left(\frac{m+1}{m}\log\left(1+\frac{1}{m}\right) - \frac{1}{m}\right)\right]\right) \leq \tilde{\alpha}$$

Taking $\alpha = \tilde{\alpha}/2$ and applying the union bound completes the proof.

C Theoretical extension of Theorem 3.2

In this section, we correct the order of the confidence sequence presented in Theorem 3.2 for theoretical completeness. We leverage the results presented in (Waudby-Smith et al., 2022; Howard et al., 2020) by applying a mixture martingale over a truncated gamma distribution.

The above proof shows that

$$M_n := \exp\left(\lambda A_n + B_n(\lambda + \log(1-\lambda))\right) \tag{13}$$

is a super-martingale with initial value 1, where

$$A_n := \frac{\sum_{i=1}^{n}(\hat{\tau}_i - \tau_n)}{m}, \quad B_n := \frac{S_n}{m^2}.$$

For any distribution F on $(0,1)$, we have by Fubini's theorem that

$$\tilde{M}_n := \int_{\lambda\in(0,1)} M_n dF(\lambda)$$

is again another super-martingale with initial value 1. Following the proof of Theorem 2 in (Waudby-Smith et al., 2022), we choose the truncated gamma distribution given by

$$f(\lambda) = \frac{\rho^\rho e^{-\rho(1-\lambda)}(1-\lambda)^{\rho-1}}{\Gamma(\rho) - \Gamma(\rho,\rho)}$$

for any $\rho \geq 0$. Therefore, we have that

$$\begin{aligned}
\tilde{M}_n &= \int_0^1 \exp\left\{\lambda A_n + B_n(\lambda + \log(1-\lambda)\right\} f(\lambda) d\lambda \\
&= \int_0^1 \exp\left\{\lambda A_n + B_n(\lambda + \log(1-\lambda)\right\} \frac{\rho^\rho e^{-\rho(1-\lambda)}(1-\lambda)^{\rho-1}}{\Gamma(\rho) - \Gamma(\rho,\rho)} d\lambda \\
&= \frac{\rho^\rho e^{-\rho}}{\Gamma(\rho) - \Gamma(\rho,\rho)} \int_0^1 \exp\{\lambda\left(A_n + B_n + \rho\right)\}(1-\lambda)^{B_n+\rho-1} d\lambda \\
&= \left(\frac{\rho^\rho e^{-\rho}}{\Gamma(\rho) - \Gamma(\rho,\rho)}\right)\left(\frac{1}{B_n+\rho}\right) {}_1F_1(1, B_n+\rho+1, A_n+B_n+\rho),
\end{aligned}$$

where the last line follows from the definition of the Kummer's confluent hypergeometric function.

Therefore, we have by Lemma 3.1 that

$$\Pr\left(\exists n : \left(\frac{\rho^\rho e^{-\rho}}{\Gamma(\rho) - \Gamma(\rho,\rho)}\right)\left(\frac{1}{B_n+\rho}\right) {}_1F_1(1, B_n+\rho+1, A_n+B_n+\rho) \geq \frac{1}{\tilde{\alpha}}\right) \leq \tilde{\alpha}$$

Consequently, a one-sided lower confidence sequence, i.e., a confidence sequence that is bounded from below but covers up to infinity, can be obtained by a root-finding algorithm to find all

$$\{\tau_n : V_n(\tau_n) \geq \frac{1}{\tilde{\alpha}}\},$$

where

$$V_n(\tau_n) := \left(\frac{\rho^\rho e^{-\rho}}{\Gamma(\rho) - \Gamma(\rho,\rho)}\right)\left(\frac{1}{B_n+\rho}\right) {}_1F_1(1, B_n+\rho+1, A_n+B_n+\rho).$$

An upper confidence sequence can be obtained in a similar way. Furthermore, this confidence sequence does not solve the issue where it requires the analyst to know M and p_{min} before the experiment. Furthermore, this confidence sequence does not have a closed-form expression, thus it requires a root-solving algorithm to build the confidence sequence. However, (Waudby-Smith et al., 2022) show that this provably has an asymptotic rate of $O(\sqrt{B_n \log(B_n)}/n)$, which does solve the issue related to the order of the confidence sequence width.

D Proof of Theorem 5.2

The proof proceeds in three steps. We note that this proof leverages the proof of Theorem 2.3 in (Waudby-Smith et al., 2021) but extended to our setting.

Step 1: Building martingale using Gaussian distribution Recently, (Ramdas et al., 2020) shows that all sequential tests must have an explicit or implicit construction of a non-negative martingale. Although one of the major advantages of an asymptotic confidence sequences is that it avoids explicitly constructing a martingale, the proof still relies on constructing a martingale with the asymptotic Gaussian distribution. Consequently, the first step of the proof builds a martingale from a sequence of iid standard Gaussian random variables.

Let $(Z_t)_{t=1}^\infty$ be a sequence of iid standard Gaussian random variable. We note that

$$M_t(\lambda) := \exp\left(\sum_{j=1}^t (\lambda \sigma_{j|j-1} Z_j - \lambda^2 \sigma_{j|j-1}^2/2)\right)$$

is a non-negative martingale starting at one for any $\lambda \in \mathbb{R}$ with respect to the canonical filtration (Robbins, 1970). For algebraic simplicity, we also define $L_t := \sum_{j=1}^{t} \sigma_{j|j-1} Z_j$ and $\bar{\sigma}_t^2 = \frac{1}{t} \sum_{j=1}^{t} \sigma_{j|j-1}^2$. Moreover, for any probability distribution $F(\lambda)$ on $\mathbb{R}$, we also have the mixture,

$$\int_{\lambda \in \mathbb{R}} M_t(\lambda) dF(\lambda)$$

is again a non-negative martingale with initial value one (Robbins, 1970). In particular, we consider the probability distribution function $f(\lambda; 0, \eta^2)$ for the Gaussian distribution with mean zero and variance η^2 as the mixing distribution. The resulting martingale is

$$\begin{aligned}
M_t :=& \int_{\lambda \in \mathbb{R}} M_t(\lambda) f(\lambda; 0, \eta^2) d\lambda \\
=& \frac{1}{\sqrt{2\pi\eta^2}} \int_{\lambda} \exp\left(\lambda L_t - \frac{t\lambda^2 \bar{\sigma}_t^2}{2} \right) \exp\left(\frac{-\lambda^2}{2\eta^2} \right) d\lambda \\
=& \frac{1}{\sqrt{2\pi\eta^2}} \int_{\lambda} \exp\left(\lambda L_t - \frac{\lambda^2 (1 + t\eta^2 \bar{\sigma}_t^2)}{2\eta^2} \right) d\lambda \\
=& \frac{1}{\sqrt{2\pi\eta^2}} \int_{\lambda} \exp\left(\frac{-\lambda^2 (1 + t\eta^2 \bar{\sigma}_t^2) + 2\lambda\eta^2 L_t}{2\eta^2} \right) d\lambda \\
=& \frac{1}{\sqrt{2\pi\eta^2}} \int_{\lambda} \exp\left(\frac{-a(\lambda^2 + \frac{b}{a} 2\lambda}{2\eta^2} \right) d\lambda,
\end{aligned}$$

where $a = t\eta^2 \bar{\sigma}_t^2 + 1$ and $b = \eta^2 L_t$. Completing the square, we have that the integrand is:

$$\exp\left(\frac{-a(\lambda^2 + \frac{b}{a} 2\lambda}{2\eta^2} \right) = \exp\left(\frac{-(\lambda - b/a)^2}{2\eta^2/a} \right) \exp\left(\frac{b^2}{2a\eta^2} \right).$$

Putting the expression back into M_t we have that,

$$\begin{aligned}
M_t &= \frac{1}{\sqrt{2\pi\eta^2/a}} \int_{\lambda} \exp\left(\frac{-(\lambda - b/a)^2}{2\eta^2/a} \right) d\lambda \frac{\exp\left(\frac{b^2}{2a\eta^2} \right)}{\sqrt{a}} \\
&= \exp\left(\frac{\eta^2 (\sum_{j=1}^{t} \sigma_{j|j-1} Z_j)^2}{2(t\bar{\sigma}_t^2 \eta^2 + 1)} \right) (t\bar{\sigma}_t^2 \eta^2 + 1)^{-1/2},
\end{aligned}$$

where the last line follows because the first part of the first line is one and we plug back in the definition of a and b.

Since M_t is a non-negative martingale with initial value one we can use Lemma 3.1 to claim that

$$\begin{aligned}
&\Pr(\forall t \geq 1, M_t < 1/\alpha) \geq 1 - \alpha \\
&= \Pr\left(\forall t \geq 1, \left| \frac{1}{t} \sum_{j=1}^{t} \sigma_{j|j-1} Z_j \right| < \sqrt{\frac{2(t\bar{\sigma}_t^2 \eta^2 + 1)}{t^2 \eta^2} \log\left(\frac{\sqrt{t\bar{\sigma}_t^2 \eta^2 + 1}}{\alpha} \right)} \right) \geq 1 - \alpha,
\end{aligned} \tag{14}$$

where the last line follows from taking the logarithm and simple algebraic manipulation.

Step 2: Strong Approximation via Martingale Sequence Differences We first define the estimation error as

$$u_t = \hat{\tau}_{t|t-1} - \tau_t(w^{obs}_{1:(t-1)}).$$

By Lemma 5.1, $\{u_t\}$ is a martingale difference sequence with respect to $\mathcal{F}_{t-1}$. Similar to the proof of Step 2 of (Waudby-Smith et al., 2021), we also use the strong approximation theorem presented in (Strassen, 1967). In particular, we require Equation (159) in Theorem 4.4 of Strassen's paper (further details in Lemma A.3 of (Waudby-Smith et al., 2021)) for our strong approximation theorem. However, our proof is different than that in Step 2 of (Waudby-Smith et al., 2021) for the following reason.

The original Theorem 4.4 in (Strassen, 1967) is stated for martingales difference sequence of the form $E(D_n \mid \sigma(D_1, \dots, D_{n-1})) = 0$, where $D_i := (W_i, Y_i)$ is the data at time i. Although our martingale is of the form $E(f(D_n) \mid \sigma(D_1, \dots, D_{n-1})) = 0$, where $f(.)$ is the function that maps the data to $\hat{\tau}_{t|t-1} - \tau_t(w^{obs}_{1:(t-1)})$. More formally, to use the strong approximation theorem in (Strassen, 1967), we replace the beginning conditions of Theorem 4.4 in the following way.

"Let $X_1, X_2, \dots$ be random variables such that $0 \leq E(f(X_n)^2 \mid X_1, \dots, X_{n-1}) \leq C$ is bounded by some constant C (this directly holds under Assumption 6) and $E(f(X_n) \mid X_1, \dots, X_{n-1}) = 0$, a.s. for all n. Put $S_n = \sum_{i \leq n} f(X_i)$ and $V_n = \sum_{i \leq n} E(f(X_i)^2 \mid X_1, \dots, X_{i-1})$, where, in order to avoid trivial complications, we assume $V_1 = E(f(X_1)^2) > 0$."

The remaining conditions are identical and we omit the uniform integrability condition in Equation (138) of (Strassen, 1967) since it holds trivially under our bounded potential outcome for Assumption 6. The proof leading to Equation (159) remains identical and valid except replacing X_n with $f(X_n)$ in the appropriate steps. In particular, all random variables are still measurable with respect to $\sigma(X_1, \dots, X_{n-1})$. Lastly, although this theorem uses the actual variance (not an upper bound), using an upper bound only makes the confidence sequence width strictly wider and hence the validity still holds.

Finally, utilizing Theorem 4.4 Equation (159) in (Strassen, 1967) we have that

$$\frac{1}{t}\sum_{j=1}^{t} u_j = \frac{1}{t}\sum_{j=1}^{t} \sigma_{j|j-1} Z_j + o\left(\frac{\tilde{S}^{3/8}_{t|t-1}\log(\tilde{S}_{t|t-1})}{t}\right) \quad a.s., \tag{15}$$

Combining Equation (14) and Equation (15) implies that with probability at least $(1-\alpha)$,

$$\forall t \geq 1, \left|\frac{1}{t}\sum_{i=1}^{t} u_i\right| < \sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2+1}}{\alpha}\right)} + o\left(\frac{\tilde{S}^{3/8}_{t|t-1}\log(\tilde{S}_{t|t-1})}{t}\right). \tag{16}$$

Using Assumption 5, we have that

$$\frac{1}{t}\sum_{j=1}^{t}\hat{\tau}_{j|j-1} \pm \sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2+1}}{\alpha}\right)} \tag{17}$$

forms an $(1-\alpha)-$asymptotic confidence sequence for $\bar{\tau}_t(w^{obs}_{1:(t-1)})$, where we used Assumption 5 so that the $\tilde{V}_t/V_t \xrightarrow{a.s.} 1$ holds where $\tilde{V}_t$ is the non-asymptotic confidence width in Equation (16) (with the little o term) and V_t is defined in Equation (17) (without the little o term).

Step 3: Using empirical variance Unfortunately, the confidence sequence in Equation (17) can not be directly used because $\bar{\sigma}_t$ is based off the true variance and hence not obtainable from the data. The last step is to replace Equation (17) with our estimated variance $\tilde{\sigma}_t^2 := S_{t|t-1}/t$.

More precisely, we now show that if we further have $\tilde{\sigma}_t^2 \xrightarrow{a.s.} \bar{\sigma}_t^2$, then we have

$$\frac{1}{t}\sum_{j=1}^{t}\hat{\tau}_{j|j-1} \pm \sqrt{\frac{2(t\tilde{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\tilde{\sigma}_t^2\eta^2+1}}{\alpha}\right)}$$

forms a $(1-\alpha)$-asymptotic confidence sequence for $\tau_t(w^{obs}_{1:(t-1)})$, giving us the desired result. For completeness, we replicate this part of the proof under our setting. First we rewrite the assumption of $\tilde{\sigma}_t^2 \xrightarrow{a.s.} \bar{\sigma}_t^2$ as $\tilde{\sigma}_t^2 - \bar{\sigma}_t^2 = o(\bar{\sigma}_t^2)$. Then Equation (17) gives us

$$\begin{aligned}
\sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2+1}}{\alpha}\right)} &= \sqrt{\frac{2(t(\tilde{\sigma}_t^2+o(\bar{\sigma}_t^2))\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t(\tilde{\sigma}_t^2+o(\bar{\sigma}_t^2))\eta^2+1}}{\alpha}\right)} \\
&= \sqrt{\frac{t(\tilde{\sigma}_t^2+o(\bar{\sigma}_t^2))\eta^2+1}{t^2\eta^2}\log\left(\frac{t(\tilde{\sigma}_t^2+o(\bar{\sigma}_t^2))\eta^2+1}{\alpha^2}\right)} \\
&= \sqrt{\frac{t\tilde{\sigma}_t^2\eta^2+o(t\bar{\sigma}_t^2)+1}{t^2\eta^2}\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+o(t\bar{\sigma}_t^2)+1}{\alpha^2}\right)} \\
&= \sqrt{\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2}+o(\bar{\sigma}_t^2/t)\right)\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+o(t\bar{\sigma}_t^2)+1}{\alpha^2}\right)}
\end{aligned}$$

Focusing on the second logarithmic term, we have

$$\begin{aligned}
\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+o(t\bar{\sigma}_t^2)+1}{\alpha^2}\right) &= \log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}+o(t\bar{\sigma}_t^2)\right) \\
&= \log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\left[1+o(1)\right]\right) \\
&= \log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right)+\log(1+o(1)) \\
&= \log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right)+o(1),
\end{aligned}$$

where the last line follows because $\log(1+x) = x + o(1)$ for $|x| < 1$. Returning back to the main expression we have

$$\begin{aligned}
\sqrt{\frac{2(t\bar{\sigma}_t^2\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\bar{\sigma}_t^2\eta^2+1}}{\alpha}\right)} &= \sqrt{\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2} + o(V_t/t^2)\right)\left[\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right) + o(1)\right]} \\
&= \sqrt{\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2}\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right) + o(V_t/t^2) + o(V_t\log V_t/t^2) + o(V_t/t^2)} \\
&= \sqrt{\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2}\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right) + o(V_t\log V_t/t^2)} \\
&= \sqrt{\frac{t\tilde{\sigma}_t^2\eta^2+1}{t^2\eta^2}\log\left(\frac{t\tilde{\sigma}_t^2\eta^2+1}{\alpha^2}\right)} + o(\sqrt{V_t\log V_t}/t),
\end{aligned}$$

where the last line follows because $\sqrt{a+b} \le \sqrt{a}+\sqrt{b}$ for $a, b \ge 0$. This formally shows how our confidence sequence in Theorem 5.2 is a valid $(1-\alpha)-$asymptotic confidence sequence for μ_t with approximation rate $o(\sqrt{V_t\log V_t}/t)$ given that our variance estimator is strongly consistent.

However, Lemma 5.1 only tells us that $\tilde{\sigma}_t^2$ is conditionally unbiased for $\bar{\sigma}_t^2$. To establish the consistency result, we again use a version of strong law of large numbers for martingale sequence difference. We denote $U_t := \tilde{\sigma}_t^2 - \bar{\sigma}_t^2$ and U_t is a martingale sequence difference with respect to the filtration $\mathcal{F}_{t-1}$. Using classical results in (Chow, 1971), we have that $U_t \xrightarrow{a.s.} 0$ since Assumption 6 immediately satisfies the needed uniformly integrability condition. Since all the convergence statements above are almost-sure convergence, steps 1-3 give the desired claim.

E Optimizing and choosing η parameter

In this section, we show in detail how an analyst can choose η to optimize the confidence sequence width for a desired specific time t^*. The derivations are nearly identical to those presented in (Waudby-Smith et al., 2021), but we repeat them here for completeness.

Our confidence width presented in all the theorems have the following structure

$$B_t(\alpha) := \sqrt{\frac{2(t\eta^2+1)}{t^2\eta^2}\log\left(\frac{\sqrt{t\eta^2+1}}{\alpha}\right)},$$

where we have omitted the variance terms and instead substituted each $\hat{\sigma}_i = 1$ since we want η to be data-independent.[8] Then,

$$\begin{aligned}
\operatorname*{argmin}_{\eta>0} B_t(\alpha) &= \sqrt{\operatorname*{argmin}_{x>0} f(x)}, \\
\text{where} \quad f(x) &:= \frac{tx+1}{t^2x}\log\left(\frac{tx+1}{\alpha^2}\right), \quad x := \eta^2.
\end{aligned}$$

[8]Consequently, we are not formally optimizing η for the actual confidence width, but η can still be conceptually interpreted as minimizing the confidence sequence width at a desired time t^* (See Appendix C.3 in (Waudby-Smith et al., 2021) for more details).

Furthermore, $\lim_{x\to 0} f(x) = \lim_{x\to\infty} f(x)$ and thus if we can find the critical point by finding a solution for $\partial f/\partial x = 0$, then this must be the unique minimum.

Therefore, we have that

$$\frac{\partial f}{\partial x} = -\frac{1}{t^2x^2}\log\left(\frac{tx+1}{\alpha^2}\right) + \frac{1}{tx}.$$

Setting the above to zero, we obtain

$$-\alpha^2\exp(1) = -(tx+1)\exp(-(tx+1)).$$

Therefore, we have that the solution is $-(tx+1) = W_{-1}(-\alpha^2\exp(1))$, where W_{-1} is the lower branch of the Lambert W function. The solution only exists if

$$-\alpha^2\exp(1) \geq -\exp(1),$$

or equivalently if $\alpha^2 \leq 1$, which is always true for any $\alpha \in [0,1]$. Therefore, we have that

$$\operatorname*{argmin}_{\eta>0} B_t(\alpha) = \sqrt{\frac{-W_{-1}(-\alpha^2\exp(1)) - 1}{t^*}}.$$