The Labor Hub

Discussion Paper No. 20

INFERENCE FOR BATCHED ADAPTIVE EXPERIMENTS

Jan Kemper and Davud Rostam-Afschar

Inference for Batched Adaptive Experiments*

Jan Kemper[a], Davud Rostam-Afschar[b,*]

[a] *University of Mannheim, ZEW*
[b] *University of Mannheim, IZA@LISER, GLO, Labor Hub, NeSt*

Abstract

The advantages of adaptive experiments have led to their rapid adoption in economics, other fields, and among practitioners. However, adaptive experiments pose challenges for causal inference. This note suggests a BOLS (batched ordinary least squares) test statistic for inference on treatment effects in adaptive experiments. The statistic provides a precision-equalizing aggregation of per-period treatment-control differences under heteroskedasticity. The combined test statistic is a linear combination of heteroskedastic per-period z-statistics and can be used to construct asymptotically valid confidence intervals. We provide simulation results comparing rejection rates in the typical case with few treatment periods and few (or many) observations per batch.

Keywords: Adaptive experiments, Heteroskedasticity, Causal inference, Randomized controlled trial
JEL: C12, C13, C9, D83

1. Introduction

Adaptive experiments have become increasingly common because they allow for *earning while learning*. Such designs have been applied, for example, by Kasy and Sautmann (2021), Caria et al. (2024), Offer-Westort et al. (2021), Avivi et al. (2021), Tabord-Meehan (2023), Hoffmann et al. (2023), Gaul et al. (2025), Kemper and Rostam-Afschar (2026). They combine exploration and exploitation by updating treatment probabilities based on accumulated evidence. However, the dependence of assignment on past outcomes breaks the usual assumptions of random sampling and independent treatment assignment, complicating statistical inference. This is particularly problematic when there is no clear difference between outcomes under different treatments. For example, usual confidence intervals and

*We thank Leif Döring, Taro Fujita, Michael Knaus, Kei Hirano, Jack Porter, David Preinerstorfer, Christoph Rothe, Richard Winter, Kelly W. Zhang for insightful comments and discussions, as well as participants at various seminars and conferences for helpful comments. Davud Rostam-Afschar thanks the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for financial support through CRC *TRR 266 Accounting for Transparency* (Project-ID 403041268). We declare that we have no interests, financial or otherwise, that relate to the research described in this paper.

*Corresponding author
Email addresses: jan.kemper@zew.de (Jan Kemper), rostam-afschar@uni-mannheim.de (Davud Rostam-Afschar)

bootstrap methods may overreject null hypotheses. Hadad et al. (2021) use large-number-of-trials asymptotics to construct generally valid confidence intervals. Zhang et al. (2020) note that the number of trials is typically limited, but treatment assignment is adapted after each batch of observations arrives.

This note extends their argument to the more general and empirically relevant case of heteroskedastic outcomes, deriving the corresponding BOLS (batched ordinary least squares) test statistic and exploring its asymptotic distribution. The heteroskedastic case is relevant because researchers usually design experiments such that not only the outcome means but also their variances may differ by treatment arm. Often the outcome is binary (success/failure), which results in heteroskedasticity by construction.

The failure of standard procedures stems from the behavior of the treatment assignment probabilities, π_t. Under a zero-margin null hypothesis where treatments perform similarly, adaptive algorithms do not settle on a single optimal arm. Consequently, π_t does not concentrate on a deterministic constant but instead converges in distribution to a random variable. Under homoskedasticity, the weighting scheme proposed by Zhang et al. (2020) relies on the cancellation of this random limit. However, we show that under heteroskedasticity ($\sigma_1^2 \neq \sigma_0^2$), this cancellation fails. The asymptotic variance of the standard BOLS estimator remains a random variable, causing the estimator to converge to a non-normal limit.

We solve this by proposing a *locally studentized* BOLS statistic. By standardizing the treatment effect by its heteroskedastic standard error within each batch *before* aggregation, we mechanically force the predictable variance contribution of each batch to be equal. This eliminates the random limits of the assignment probabilities from the asymptotic distribution, restoring asymptotic normality and ensuring valid inference. The results are in line with Hirano and Porter (2025), who show that any limit distribution generated by joint choices of adaptive assignment rules and statistics can be represented within a unified Gaussian limit experiment (see also Bibaut and Kallus, 2025). However, Hirano and Porter (2025) do not study BOLS estimators, derive precision-equalizing weights, or analyze finite-sample performance in adaptive bandit designs. We fill this gap by deriving a heteroskedastic extension of the BOLS test statistic with precision-equalizing weights. Our approach yields a simple, implementable inference procedure and shows when heteroskedasticity matters in adaptive experiments.

2. Treatment Effects in Adaptive Experiments

2.1. Target

The main target of this analysis is to show that a test statistic Z_{het}, aggregated across batches, is asymptotically standard normal in batched adaptive experiments under heteroskedasticity. In applied settings, the most relevant application of this statistic is to test whether the treatment effect (e.g. treatment vs control arm) under homogeneous treatment effects is statistically different from zero across batches.

2.2. Setting and outcomes

Consider an experiment where in batches $t = 1, \ldots, T$, participants $i = 1, \ldots, n_t$ are assigned to treatment arm $a \in \{0, 1\}$. Their potential outcomes are

$$Y_{i,t}(a) = \mu_{a,t} + \varepsilon_{i,t}(a). \tag{1}$$

Let $\varepsilon_{i,t}(a)$ denote an arm-specific error term (heteroskedasticity) and $\Delta_t = \mu_{1,t} - \mu_{0,t}$ denote the batch-specific treatment effect. We test the joint null $H_0 : \Delta_t = \delta_{0,t}$ for all t. A relevant special case is that of homogeneous treatment effects $\Delta_t = \Delta$, which could be tested against $\delta_{0,t} = 0$ for all t.

2.3. Treatment assignment

Let $H_{t-1} = \{(Y_{i,s}, A_{i,s}) : i = 1, \ldots, n_s;\ s = 1, \ldots, t-1\}$ denote the history of outcomes and treatment assignments from batches $1, \ldots, t-1$. The adaptive assignment mechanism is characterized as follows:

(i) The treatment probability for batch t is $\pi_t = P(A_{i,t} = 1 \mid H_{t-1})$, which is H_{t-1}-measurable.

(ii) Within batch t, treatment assignments are conditionally i.i.d.: $A_{1,t}, \ldots, A_{n_t,t} \mid H_{t-1} \overset{i.i.d.}{\sim} \text{Bernoulli}(\pi_t)$.

(iii) Given $(H_{t-1}, A_{i,t})$, outcomes $Y_{i,t}$ are independent across individuals within batch t.

2.4. Treatment effects

In period t, there are $N_{1,t} = \sum_{i=1}^{n_t} A_{i,t}$ treated and $N_{0,t} = n_t - N_{1,t}$ control units. We distinguish the algorithmic assignment probability $\pi_t = P(A_{i,t} = 1 \mid H_{t-1})$ from the realized treatment share $\hat{\pi}_t = N_{1,t}/n_t$.

Let the per-period difference in sample means be

$$\hat{\Delta}_t = \frac{1}{N_{1,t}} \sum_{i=1}^{n_t} A_{i,t} Y_{i,t} - \frac{1}{N_{0,t}} \sum_{i=1}^{n_t} (1 - A_{i,t}) Y_{i,t} = \bar{Y}_{1,t} - \bar{Y}_{0,t}. \tag{2}$$

Let $\mathcal{G}_t = H_{t-1} \vee \sigma(A_{1,t}, \ldots, A_{n_t,t})$ denote the information available after the batch-t assignments are realized but before the batch-t outcomes are known. Conditional on $\mathcal{G}_t$, the realized treatment counts $N_{1,t}$ and $N_{0,t}$ are fixed. Under conditional independence and exogeneity to period-t potential outcome shocks,

$$\text{Var}(\bar{Y}_{1,t} \mid \mathcal{G}_t) = \frac{\sigma_{1,t}^2}{N_{1,t}}, \qquad \text{Var}(\bar{Y}_{0,t} \mid \mathcal{G}_t) = \frac{\sigma_{0,t}^2}{N_{0,t}}. \tag{3}$$

Hence, the conditional variance of the period difference is

$$v_t \ \equiv \ \text{Var}(\hat{\Delta}_t - \Delta_t \mid \mathcal{G}_t) = \frac{\sigma_{1,t}^2}{N_{1,t}} + \frac{\sigma_{0,t}^2}{N_{0,t}} = \frac{1}{n_t} \left(\frac{\sigma_{1,t}^2}{\hat{\pi}_t} + \frac{\sigma_{0,t}^2}{1 - \hat{\pi}_t} \right). \tag{4}$$

3. Inference

3.1. The batchwise and the BOLS test statistics

In adaptive experiments, the assignment probability π_t is random because it depends on the realized history. Consequently, the realized allocation shares $\hat{\pi}_t$ and the variances of batchwise estimators are random. This may result in asymptotic non-normality for non-studentized estimators. Intuitively, if outcomes under two treatments are hard to distinguish,

either treatment might get assigned more observations in repeated samples, and consequently the assignment probability does not concentrate. Zhang et al. (2020) show that the selection probability is fixed when conditioning on the history up to a given batch, and that the homoskedastic batchwise OLS, scaled by the selection probability, is asymptotically normal. Under heteroskedasticity, this scaling no longer removes the random assignment probabilities from the limiting variance. We therefore base inference not on an estimator but on a z-statistic that accounts for heteroskedasticity.

The key idea is to studentize within each batch *before* aggregation. For each batch t, define the z-score

$$z_{t,\text{het}} = \frac{\hat{\Delta}_t - \delta_{0,t}}{\sqrt{v_t}}, \tag{5}$$

where $\delta_{0,t}$ is the null value of the treatment effect in batch t. The variance v_t is random, but the conditional variance of each score $\text{Var}(z_{t,\text{het}} \mid \mathcal{G}_t) = \text{Var}(\hat{\Delta}_t - \Delta_t \mid \mathcal{G}_t)/v_t = 1$ is not. This is the crucial property: local studentization eliminates the random assignment probabilities from the variance. Therefore, $\sum_{t=1}^{T} \text{Var}(z_{t,\text{het}} \mid \mathcal{G}_t) = T$.

Our objective is to derive a test statistic aggregated across batches Z_{het} that is asymptotically $\mathcal{N}(0,1)$ under heteroskedasticity. The BOLS test statistic is the sum of batchwise z-scores, normalized by the *deterministic* factor $1/\sqrt{T}$:

$$Z_{\text{het}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_{t,\text{het}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\hat{\Delta}_t - \delta_{0,t}}{\sqrt{v_t}}. \tag{6}$$

The martingale-difference property implies that scores from different batches are uncorrelated, such that Z_{het} has unit variance $\text{Var}(Z_{\text{het}}) = T^{-1} \sum_{t=1}^{T} E\big[\text{Var}(z_{t,\text{het}} \mid \mathcal{G}_t)\big] = 1$.

The statistic serves two purposes. It can be used to test the batchwise null $H_0 : \Delta_t = \delta_{0,t}$ for all $t = 1, \ldots, T$ (Section 3.2). Under the common-effect restriction $\Delta_t = \Delta$ for all t, this null is $H_0 : \Delta = \delta_0$. Additionally, we derive a point estimate and a confidence interval for Δ (Section 3.3).

3.2. Asymptotic theory

Assumption 1 (Conditional Independence and Exogeneity)**.** *Within batch t, conditional on H_{t-1}, the potential outcomes $\{Y_{i,t}(0), Y_{i,t}(1)\}_{i=1}^{n_t}$ are independent across individuals and independent of the randomization used to generate the treatment assignments in batch t.*

Assumption 2 (Moment Conditions)**.** *For all t and $a \in \{0,1\}$: (i) $E[\varepsilon_{i,t}(a) \mid H_{t-1}] = 0$ and $E[\varepsilon_{i,t}(a)^2 \mid H_{t-1}] = \sigma^2_{a,t}$; (ii) there exists some $\delta > 0$ such that $E[|\varepsilon_{i,t}(a)|^{2+\delta} \mid H_{t-1}] < M < \infty$ almost surely; and (iii) the variances are nondegenerate and uniformly bounded, $0 < \underline{\sigma}^2 \leq \sigma^2_{a,t} \leq \bar{\sigma}^2 < \infty$.*

Assumption 3 (Clipping Constraint)**.** *There exists a constant $\kappa \in (0, 0.5]$ such that $P(\pi_t \in [\kappa, 1-\kappa]) \to 1$ as $n_t \to \infty$ for all t.*

Assumption 4 (Consistent Variance Estimation)**.** *The variance estimators satisfy $\hat{\sigma}^2_{a,t} \xrightarrow{p} \sigma^2_{a,t}$ for all t and $a \in \{0,1\}$.*

Assumptions 1–3 correspond to Conditions 1, 3, and the clipping constraint (Definition 1) in Zhang et al. (2020), extended to allow $\sigma_{1,t}^2 \neq \sigma_{0,t}^2$.

Theorem 1 (Asymptotic Normality of the Heteroskedastic BOLS Statistic)**.** *Under Assumptions 1–3, consider the batchwise null hypothesis $H_0 : \Delta_t = \delta_{0,t}$ for all $t = 1, \ldots, T$, with null values $\delta_{0,1}, \ldots, \delta_{0,T}$. Define the infeasible period scores and combined statistic as*

$$z_{t,\text{het}} = \frac{\hat{\Delta}_t - \delta_{0,t}}{\sqrt{v_t}}, \qquad Z_{\text{het}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} z_{t,\text{het}}. \tag{7}$$

Then, under H_0, for fixed T and $\min_t n_t \to \infty$,

$$Z_{\text{het}} \xrightarrow{d} \mathcal{N}(0,1). \tag{8}$$

If, in addition, Assumption 4 holds and $\hat{v}_t = \hat{\sigma}_{1,t}^2/N_{1,t} + \hat{\sigma}_{0,t}^2/N_{0,t}$, then the feasible statistic

$$\hat{Z}_{\text{het}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\hat{\Delta}_t - \delta_{0,t}}{\sqrt{\hat{v}_t}} \tag{9}$$

also satisfies $\hat{Z}_{\text{het}} \xrightarrow{d} \mathcal{N}(0,1)$ under H_0. If, in addition, treatment effects are assumed common across batches, $\Delta_t = \Delta$ for all t, then the same statistic tests $H_0 : \Delta = \delta_0$. The proof is provided in Appendix B.

Intuition of the proof. We studentize the estimates *locally*. By dividing the batchwise effect by its heteroskedastic standard error *before* aggregation across batches, we normalize each batch contribution to have predictable variance $1/T$ in the final statistic. The proof treats the adaptively chosen assignment probabilities as random but predictable quantities and applies a triangular-array martingale central limit theorem directly to the individual-level observations across all batches.

3.3. From BOLS test statistic to confidence intervals

Inverting the feasible BOLS test statistic of Theorem 1 yields a point estimate and a confidence interval for a *weighted aggregate effect*. This is a weighted average of the batch-specific effects Δ_t with weights proportional to their inverse standard errors. Let $w_t = 1/\sqrt{v_t}$ denote the inverse standard error of Δ_t, let $S = \sum_{s=1}^{T} w_s$, and define the weights

$$\omega_t = \frac{w_t}{S}, \qquad \omega_t > 0, \qquad \sum_{t=1}^{T} \omega_t = 1. \tag{10}$$

The weighted aggregate effect and its estimator, which we call the *heteroskedastic BOLS estimator*, are

$$\theta = \sum_{t=1}^{T} \omega_t \, \Delta_t, \qquad \hat{\theta} = \sum_{t=1}^{T} \hat{\omega}_t \, \hat{\Delta}_t. \tag{11}$$

Without further restrictions, θ is a weighted average of the batchwise Δ_t. Since the Δ_t can differ and the weights ω_t depend on the realized allocations, θ is random and varies

across realizations of the experiment. The restriction to a homogeneous effect $\Delta_t = \Delta$ for all t ensures that the target is a fixed parameter. Under this restriction θ equals Δ regardless of the weights, and the heteroskedastic BOLS estimator $\hat{\theta}$ consistently estimates Δ. The test $H_0 : \Delta = \delta_0$ and the corresponding confidence interval refer to a single, well-defined parameter in this case. The following further results are all derived under the *homogeneous treatment effect assumption* ($\Delta_t = \Delta$). Remark 4 discusses the implications for $\hat{\theta}$ when we lift the homogeneous treatment effect assumption.

Evaluating the feasible statistic of Theorem 1 at $\delta_{0,t} = \Delta_t$ and writing $\hat{q}_T = \sqrt{T}/\hat{S}$, it follows from (11) that

$$\hat{Z}_{\text{het}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\hat{\Delta}_t - \Delta_t}{\sqrt{\hat{v}_t}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{w}_t (\hat{\Delta}_t - \Delta_t) = \frac{\hat{\theta} - \theta}{\hat{q}_T}. \tag{12}$$

Corollary 1 (Asymptotically Valid Confidence Interval)**.** *Suppose Assumptions 1–4 hold and $\Delta_t = \Delta$ for all t. Then*

$$\text{CI}_{1-\alpha} = \hat{\theta} \pm z_{1-\alpha/2}\, \hat{q}_T \tag{13}$$

satisfies $P\big(\Delta \in \text{CI}_{1-\alpha}\big) \to 1 - \alpha$, where $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standard normal distribution.

Coverage follows by inversion from (12) and Theorem 1, since $(\hat{\theta} - \theta)/\hat{q}_T \xrightarrow{d} \mathcal{N}(0, 1)$.

3.4. Properties of the heteroskedastic BOLS estimator

The heteroskedastic BOLS estimator is a weighted average of the batchwise estimates $\hat{\Delta}_t$ with data-dependent weights $\hat{\omega}_t$. Its estimation error is

$$\hat{\theta} - \theta = \sum_{t=1}^{T} \hat{\omega}_t \, (\hat{\Delta}_t - \Delta_t), \tag{14}$$

which follows from (11).

Remark 1 (Bias)**.** *Each batchwise estimator is conditionally unbiased, $E[\hat{\Delta}_t - \Delta_t \mid \mathcal{G}_t] = 0$. The heteroskedastic BOLS estimator nevertheless need not be unbiased, $E[\hat{\theta} - \theta] \neq 0$ in general, because future weights depend on past outcomes and thus the weights in (14) are correlated with the estimation errors.*

Remark 2 (Consistency)**.** *Despite its finite-sample bias, $\hat{\theta}$ is consistent under homogeneous effects. Since $\hat{\omega}_t \geq 0$ and $\sum_{t=1}^{T} \hat{\omega}_t = 1$ by (10), (14) gives*

$$|\hat{\theta} - \theta| \leq \sum_{t=1}^{T} \hat{\omega}_t \, |\hat{\Delta}_t - \Delta_t| \leq \max_t |\hat{\Delta}_t - \Delta_t|. \tag{15}$$

The first inequality comes from the triangle inequality that implies that the absolute aggregate error does not exceed the sum of the individual absolute errors; the second holds because a weighted average cannot exceed its largest element. These inequalities hold in finite samples, regardless of how the adaptive weights defined in (10) are determined. Hence, for fixed T and $n = \min_t n_t \to \infty$,

$$\hat{\theta} - \theta = O_p(n^{-1/2}), \qquad \hat{\theta} \xrightarrow{p} \Delta. \tag{16}$$

Remark 3 (Distribution of $\hat{\theta}$)**.** *In general, $\hat{\theta}$ is not asymptotically normal, because the adaptive weights $\hat{\omega}_t$ remain random in large samples. This does not affect the confidence interval in Corollary 1, which relies only on the asymptotic normality of $\hat{Z}_{\text{het}} = (\hat{\theta} - \theta)/\hat{q}_T$. If the weights are fixed, $\hat{\theta} - \theta = \sum_t \omega_t(\hat{\Delta}_t - \Delta_t)$ is a fixed linear combination of asymptotically independent normal batch estimates and hence asymptotically normal, with variance $\sum_t \omega_t^2 v_t = T/S^2$. When the treatment effect is large relative to the outcome noise, the algorithm quickly identifies the better arm, the assignment probabilities settle at a deterministic value, and $\hat{\theta}$ is approximately normal.*

Remark 4 (Heterogeneous treatment effects)**.** *Under heterogeneous treatment effects, $\hat{\theta}$ does not converge to a fixed parameter, because under adaptive assignment the weights remain random in large samples, so that $\hat{\theta}$ converges to different linear combinations of the Δ_t across realizations of the experiment. Nonetheless, Theorem 1 remains valid, and a confidence interval can be derived for the linear combination given by the realized weights. This interval is, however, difficult to interpret, since its target is not a fixed parameter but depends on the history of the experiment.*

3.5. Feasible implementation

The heteroskedastic BOLS estimator (11) and the confidence interval in Corollary 1 require only the batchwise estimates $\hat{\Delta}_t$ and the variance estimates

$$\hat{v}_t = \frac{\hat{\sigma}_{1,t}^2}{N_{1,t}} + \frac{\hat{\sigma}_{0,t}^2}{N_{0,t}}, \tag{17}$$

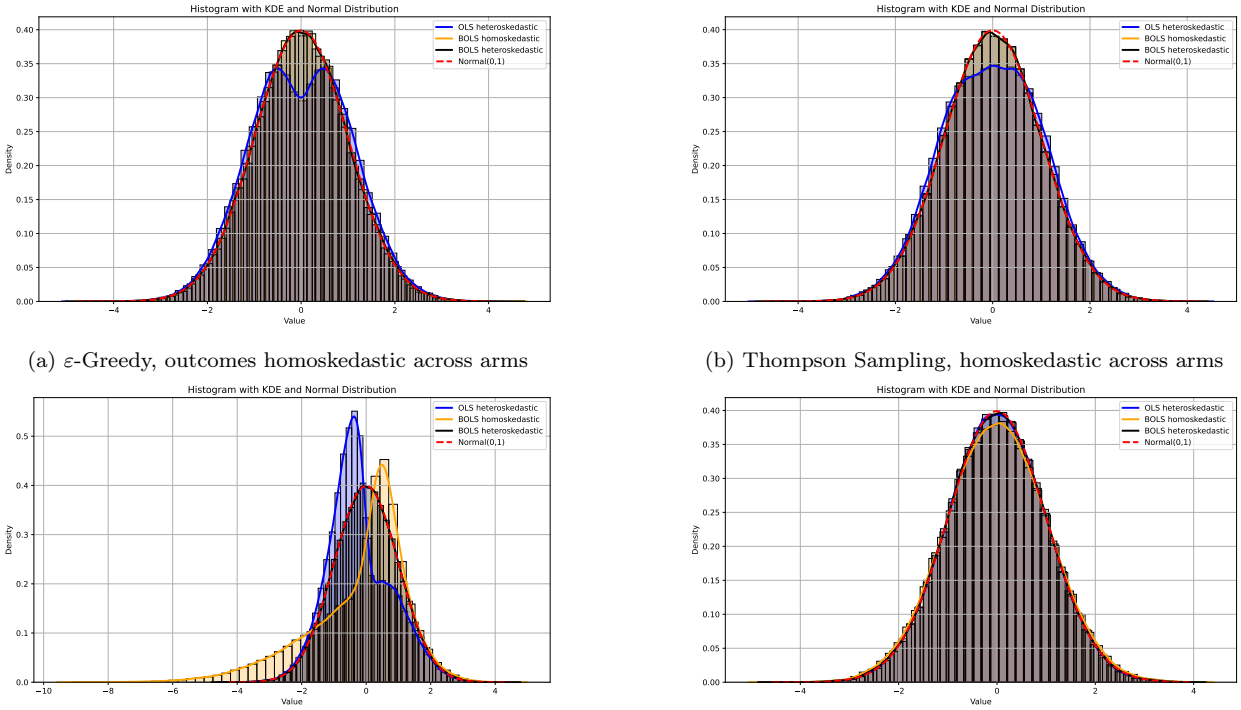
where $\hat{\sigma}_{a,t}^2$ is the sample variance of the outcomes in arm a of batch t. Table A.1 shows the resulting weights for the non-adaptive and adaptive cases under homoskedasticity and heteroskedasticity.

Remark 5 (Variance Estimation and Weights)**.** *(i) In small samples, finite-sample adjusted variance estimators (e.g., HC2/HC3) may be preferable. (ii) If the variances are constant across batches, $\sigma_{a,t}^2 = \sigma_a^2$, arm-specific variance estimators pooled across batches also satisfy Assumption 4 and are more precise (cf. Corollary 4 in Zhang et al., 2020). Pooling should use deviations from the batch-specific arm means, since $\mu_{a,t}$ may vary across batches. (iii) Under homoskedasticity ($\sigma_{1,t}^2 = \sigma_{0,t}^2 = \sigma^2$) and constant batch size, the weights reduce to $\omega_t = \sqrt{N_{1,t}N_{0,t}}/\sum_{s=1}^{T}\sqrt{N_{1,s}N_{0,s}}$, so batches with more balanced realized sizes receive larger weight. With a pooled variance estimate, the homoskedastic $\hat{\sigma}^2$ cancels and the feasible weights $\hat{\omega}_t$ equal this expression exactly. (iv) The extension to K treatment arms and contextual settings can be constructed following the same arguments applied in this analysis.*

4. Monte Carlo Simulations

We compare three test statistics: the heteroskedasticity-robust OLS statistic, the BOLS statistic derived under homoskedasticity, and our heteroskedasticity-robust BOLS statistic. We report *rejection rates*, i.e., the proportion of simulated samples in which the true null hypothesis is rejected. Under the zero-margin null $H_0 : \Delta = 0$, the algorithm struggles to identify the optimal arm, so that treatment probabilities π_t remain highly variable. The

rejection rate measures the empirical size of the test (nominal 5%). Data are generated using two common adaptive sampling algorithms, ε-Greedy and Bernoulli Thompson Sampling in two-arm settings.



(a) ε-Greedy, outcomes homoskedastic across arms

(b) Thompson Sampling, homoskedastic across arms

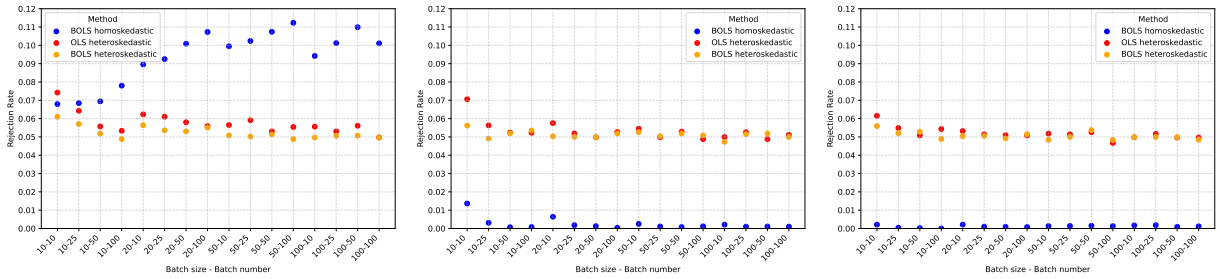
(c) ε-Greedy, outcomes heteroskedastic across arms

(d) Thompson Sampling, heteroskedastic across arms

Notes: The figure shows the results of Monte Carlo simulations with 100,000 repetitions. Panels a) and c) show the distribution of the heteroskedasticity-robust OLS test statistic, the homoskedastic BOLS test statistic, and the heteroskedasticity-robust BOLS test statistic for data generated from an ε-Greedy experiment. The batch size is 500, the number of batches 25, the experiment consists of two arms, each with an expected value of 1. In panel a) the standard deviation is 1 for arm one and 1 for arm two (homoskedasticity). In panel c) the standard deviation for arm one is 4 and for arm two is 1, everything else remains equal. The red dotted line indicates the density of the standard normal distribution. Panel b) shows data generated from a Bernoulli Thompson algorithm. The batch size is 500, the number of batches 25, the experiment consists of two arms with an expected value of 0.5 and 0.5 (homoskedasticity). In panel d) the experiment consists of two arms with an expected value of 0.7 and 0.4 (heteroskedasticity).

Figure 1: Simulation I: Non-normality of OLS and homoskedastic BOLS

Simulation I: Non-normality of test statistics based on OLS and homoskedastic BOLS. Figure 1 shows the empirical distributions of the test statistics with 25 batches each with a size of 500 observations. For ε-Greedy (left panels), both arms have Gaussian outcomes (think of log incomes) with mean 1, with variances $(1^2, 1^2)$ in panel a) and $(4^2, 1^2)$ in panel c). The exploration rate is $\varepsilon = 0.2$. Each design is repeated 100,000 times. Consistent with Zhang et al. (2020), panel a) shows that the test statistic from OLS under a zero-margin is non-normally distributed and homoskedastic BOLS recovers its normality. But in the same setting under heteroskedasticity (panel c), both OLS and the homoskedastic BOLS statistics deviate markedly from normality in the zero-margin case. The homoskedastic BOLS statistic severely overrejects (17% instead of 5%). OLS yields approximately correct rejection rates



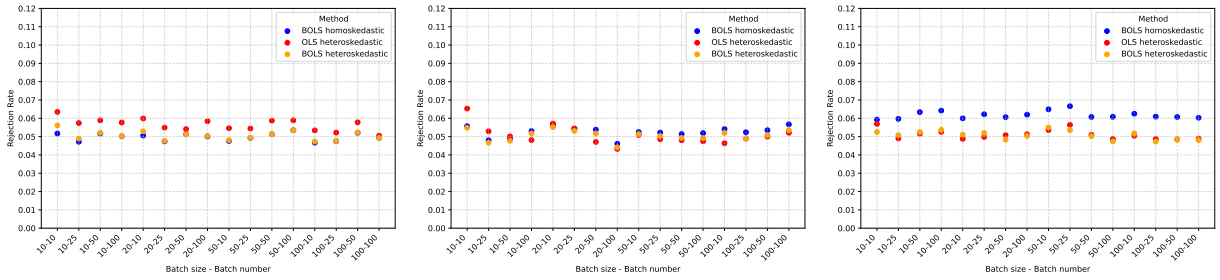
(a) $\Delta = 0$ (b) $\Delta = 1$ (c) $\Delta = 2$

Notes: This figure shows Monte Carlo simulation results for the ε-Greedy algorithm. Combinations of batch size and number of batches increase along the horizontal axis. Rejection rates are denoted on the vertical axis. The parameter Δ indicates the difference between the true expected values of both arms. Each circle shows the average rejection rate for the given test statistic which is indicated by color. For each batch size / number of batches combinations 10,000 repetitions were executed. In panel (a) the expected value μ_1 for arm 1 is 1 and μ_2 for arm 2 is also 1. The standard deviation in all panels for arm 1 is $\sigma_1 = 2$ and for arm 2 is $\sigma_2 = 1$. In panel (b) the expected value is $\mu_1 = 2$ and $\mu_2 = 1$. For panel (c) it is $\mu_1 = 3$ and $\mu_2 = 1$. In the batch size 10-10 scenario, on average, no test statistic could be calculated for one percent of the draws because one of the arms was never played.

Figure 2: ε-Greedy

but exhibits non-normal behavior. In contrast, our heteroskedasticity-robust BOLS statistic closely matches the standard normal distribution and delivers correct 5% rejection rates in both cases.

For Thompson Sampling (right panels), we set Bernoulli success probabilities to $(p_1, p_2) = (0.5, 0.5)$ in panel b) and to $(p_1, p_2) = (0.7, 0.4)$ in panel d). Panel b) confirms that the test statistic based on OLS is non-normally distributed and the homoskedastic BOLS test statistic fixes the problem. However, under mechanical heteroskedasticity (panel d), the homoskedastic BOLS statistic overrejects ($\approx 6\%$) because it ignores heteroskedasticity. As the success probabilities differ substantially, both OLS and our heteroskedasticity-robust BOLS statistic closely match the standard normal distribution and deliver correct 5% rejection rates.



(a) $\Delta = 0$ (b) $\Delta = 0.1$ (c) $\Delta = 0.2$

Notes: This figure shows Monte Carlo simulation results for the Bernoulli Thompson Sampling algorithm. Combinations of batch size and number of batches increase along the horizontal axis. Rejection rates are denoted on the vertical axis. The parameter Δ indicates the difference between the true expected values of both arms. Each circle shows the average rejection rate for the given test statistic which is indicated by color. For each batch size / number of batches combinations 10,000 repetitions were executed. In panel (a) the success probability for arm 1 is $p_1 = 0.5$ and arm 2 is also $p_2 = 0.5$. In panel (b) it is $p_1 = 0.6$ and $p_2 = 0.5$. For panel (c) it is $p_1 = 0.7$ and $p_2 = 0.5$.

Figure 3: Bernoulli Thompson Sampling

Simulation II: Rejection Rates at Small and Large Margins in Small and Large Samples. To study behavior in smaller samples, we vary the number of batches (10–100) and batch sizes (10–100). Each configuration is repeated 10,000 times. Figures 2 and 3 report rejection rates of a 5% significance level test of the true null $H_0 : \Delta = \delta_0$.

For ε-Greedy (Figure 2), the homoskedastic BOLS statistic is unreliable in all settings: it overrejects sharply at $\Delta = 0$ and underrejects for positive margins. The heteroskedasticity-robust OLS statistic performs reasonably well and improves as the margin grows. Across all designs, our heteroskedasticity-robust BOLS statistic maintains rejection rates close to 5%.

For Thompson Sampling (Figure 3), the zero-margin case exhibits no heteroskedasticity, so the heteroskedastic and the homoskedastic BOLS statistics perform well with rejection rates near 5%. The heteroskedasticity-robust OLS overrejects somewhat. When margins increase, inducing heteroskedasticity, the homoskedastic BOLS statistic begins to overreject, while both robust statistics remain close to nominal size.

5. Conclusion

Adaptive experiments have made precision-equalizing inference increasingly relevant in sequential settings where treatment assignment depends on past outcomes. The BOLS inverse-standard-error-weighted statistic provides a simple, asymptotically valid procedure for inference under heteroskedasticity, extending previous results derived for the homoskedastic case.

References

Avivi, H., P. Kline, E. Rose, and C. Walters (2021): "Adaptive Correspondence Experiments," *AEA Papers and Proceedings*, 111, 43–48.

Bibaut, A., and N. Kallus (2025): "Demystifying Inference After Adaptive Experiments," *Annual Review of Statistics and Its Application*, 12, 407–423.

Caria, A. S., G. Gordon, M. Kasy, S. Quinn, S. O. Shami, and A. Teytelboym (2024): "An Adaptive Targeted Field Experiment: Job Search Assistance for Refugees in Jordan," *Journal of the European Economic Association*, 22(2), 781–836.

Gaul, J. J., F. Keusch, D. Rostam-Afschar, and T. Simon (2025): "Invitation Messages for Business Surveys: A Multi-Armed Bandit Experiment," *Survey Research Methods*, 19(4), 409–429.

Hadad, V., D. A. Hirshberg, R. Zhan, S. Wager, and S. Athey (2021): "Confidence Intervals for Policy Evaluation in Adaptive Experiments," *Proceedings of the National Academy of Sciences*, 118(15), e2014602118.

Hall, P., and C. C. Heyde (1980): *Martingale Limit Theory and Its Application*. Academic Press, New York.

Hirano, K., and J. R. Porter (2025): "Asymptotic Representations for Sequential Decisions, Adaptive Experiments, and Batched Bandits," arXiv 2302.03117, econ.EM.

Hoffmann, M., B. Picard, C. Marie, and G. Bied (2023): "An Adaptive Experiment to Boost Online Skill Signaling and Visibility," Draft, Accessed: 2026-11-01.

Kasy, M., and A. Sautmann (2021): "Adaptive Treatment Assignment in Experiments for Policy Choice," *Econometrica*, 89(1), 113–132.

Kemper, J., and D. Rostam-Afschar (2026): "Earning While Learning: How to Run Batched Bandit Experiments," *Stata Journal*, 26(3), 461–493.

Offer-Westort, M., A. Coppock, and D. P. Green (2021): "Adaptive Experimental Design: Prospects and Applications in Political Science," *American Journal of Political Science*, 65(4), 826–844.

Tabord-Meehan, M. (2023): "Stratification Trees for Adaptive Randomization in Randomized Controlled Trials," *Review of Economic Studies*, 90(5), 2646–2673.

Zhang, K., L. Janson, and S. Murphy (2020): "Inference for Batched Bandits," *Advances in Neural Information Processing Systems*, 33, 9818–9829.

Appendix A. Weight structure under homoskedasticity and heteroskedasticity

Table A.1: Weight structure under homoskedasticity and heteroskedasticity for two-armed experiments

	Non-adaptive	**Adaptive**
	Homoskedastic case: $\sigma^2_{1,t} = \sigma^2_{0,t} = \sigma^2$	
Realized share $\hat{\pi}_t$	$\hat{\pi}_t = \pi_t$, constant across batches	π_t depends on $\mathcal{H}_{t-1}$, $\hat{\pi}_t$ random
Variance v_t	$v_t = \dfrac{\sigma^2}{n_t\,\pi_t(1-\pi_t)}$	$v_t = \dfrac{\sigma^2}{n_t\,\hat{\pi}_t(1-\hat{\pi}_t)}$
Weight $\omega_t = \dfrac{w_t}{S}$	$\dfrac{\sqrt{n_t}}{\sum_{s=1}^{T}\sqrt{n_s}}$	$\dfrac{\sqrt{n_t\,\hat{\pi}_t(1-\hat{\pi}_t)}}{\sum_{s=1}^{T}\sqrt{n_s\,\hat{\pi}_s(1-\hat{\pi}_s)}}$
	Heteroskedastic case: $\sigma^2_{1,t} \neq \sigma^2_{0,t}$	
Realized share $\hat{\pi}_t$	$\hat{\pi}_t = \pi_t$, constant across batches	π_t depends on $\mathcal{H}_{t-1}$, $\hat{\pi}_t$ random
Variance v_t	$v_t = \dfrac{1}{n_t}\left(\dfrac{\sigma^2_{1,t}}{\pi_t} + \dfrac{\sigma^2_{0,t}}{1-\pi_t}\right)$	$v_t = \dfrac{1}{n_t}\left(\dfrac{\sigma^2_{1,t}}{\hat{\pi}_t} + \dfrac{\sigma^2_{0,t}}{1-\hat{\pi}_t}\right)$
Weight $\omega_t = \dfrac{w_t}{S}$	$\dfrac{\sqrt{n_t}\left(\frac{\sigma^2_{1,t}}{\pi_t} + \frac{\sigma^2_{0,t}}{1-\pi_t}\right)^{-1/2}}{\sum_{s=1}^{T}\sqrt{n_s}\left(\frac{\sigma^2_{1,s}}{\pi_s} + \frac{\sigma^2_{0,s}}{1-\pi_s}\right)^{-1/2}}$	$\dfrac{\sqrt{n_t}\left(\frac{\sigma^2_{1,t}}{\hat{\pi}_t} + \frac{\sigma^2_{0,t}}{1-\hat{\pi}_t}\right)^{-1/2}}{\sum_{s=1}^{T}\sqrt{n_s}\left(\frac{\sigma^2_{1,s}}{\hat{\pi}_s} + \frac{\sigma^2_{0,s}}{1-\hat{\pi}_s}\right)^{-1/2}}$

Notes: $n_t = N_{1,t} + N_{0,t}$ is total batch size, and $\hat{\pi}_t = N_{1,t}/n_t$ is the realized treatment share. The corresponding algorithmic assignment probability is denoted by π_t. The non-adaptive column refers to fixed-count designs, in which $N_{1,t} = \pi_t n_t$ is set in advance (with $\pi_t n_t$ an integer) and π_t is the same in every batch, so that $\hat{\pi}_t = \pi_t$ exactly. In the adaptive case π_t is $\mathcal{H}_{t-1}$-measurable and the weights depend on the realized shares. Inverse-standard-error weights equalize the predictable variance of the batchwise contributions and are $w_t = 1/\sqrt{v_t}$, their sum $S = \sum_{s=1}^{T} w_s$. Under homoskedasticity, adaptive weights ω_t depend on both sample size and realized balance $\hat{\pi}_t(1-\hat{\pi}_t)$, whereas in the non-adaptive case $\pi_t(1-\pi_t)$ is common to all batches and cancels, leaving weights that depend on n_t alone. Under heteroskedasticity, weights additionally adjust for treatment-specific outcome variances $\sigma^2_{a,t}$.

Appendix B. Proof of Theorem 1

We prove asymptotic normality of Z_{het} using a triangular-array martingale central limit theorem applied directly to the individual-level observations across all batches. Throughout, T is fixed and $n = \min_t n_t \to \infty$. All quantities may depend on n; this dependence is suppressed for readability.

B.1 Setup

We first condition on the realized assignment pattern in each batch. Recall that H_{t-1} denotes the history before batch t, and define

$$\mathcal{G}_t = H_{t-1} \vee \sigma(A_{1,t}, \ldots, A_{n_t,t}).$$

Thus $\mathcal{G}_t$ contains the pre-batch history and the full assignment vector in batch t, but not the batch-t outcomes. Conditional on $\mathcal{G}_t$, the realized treatment counts

$$N_{1,t} = \sum_{i=1}^{n_t} A_{i,t}, \qquad N_{0,t} = \sum_{i=1}^{n_t} (1 - A_{i,t})$$

are known. The variance parameters $\sigma^2_{a,t}$ are fixed arm- and batch-specific quantities satisfying, by Assumption 2,

$$E[\varepsilon_{i,t}(a)^2 \mid H_{t-1}] = \sigma^2_{a,t}.$$

Hence

$$v_t = \frac{\sigma^2_{1,t}}{N_{1,t}} + \frac{\sigma^2_{0,t}}{N_{0,t}}$$

is $\mathcal{G}_t$-measurable and can be treated as fixed in conditional variance calculations.

By the potential-outcome representation in equation (1), the observed outcome is

$$Y_{i,t} = A_{i,t} Y_{i,t}(1) + (1 - A_{i,t}) Y_{i,t}(0) = A_{i,t}\{\mu_{1,t} + \varepsilon_{i,t}(1)\} + (1 - A_{i,t})\{\mu_{0,t} + \varepsilon_{i,t}(0)\}.$$

Substituting this into the means of equation (2) gives

$$\frac{1}{N_{1,t}} \sum_{i=1}^{n_t} A_{i,t} Y_{i,t} = \frac{1}{N_{1,t}} \sum_{i=1}^{n_t} A_{i,t}\{\mu_{1,t} + \varepsilon_{i,t}(1)\} = \mu_{1,t} + \frac{1}{N_{1,t}} \sum_{i=1}^{n_t} A_{i,t} \varepsilon_{i,t}(1),$$

because $\sum_{i=1}^{n_t} A_{i,t} = N_{1,t}$. Similarly,

$$\frac{1}{N_{0,t}} \sum_{i=1}^{n_t} (1 - A_{i,t}) Y_{i,t} = \mu_{0,t} + \frac{1}{N_{0,t}} \sum_{i=1}^{n_t} (1 - A_{i,t}) \varepsilon_{i,t}(0).$$

Hence

$$\widehat{\Delta}_t = \mu_{1,t} - \mu_{0,t} + \frac{1}{N_{1,t}} \sum_{i=1}^{n_t} A_{i,t} \varepsilon_{i,t}(1) - \frac{1}{N_{0,t}} \sum_{i=1}^{n_t} (1 - A_{i,t}) \varepsilon_{i,t}(0).$$

Since $\Delta_t = \mu_{1,t} - \mu_{0,t}$, subtracting Δ_t yields

$$\widehat{\Delta}_t - \Delta_t = \frac{1}{N_{1,t}} \sum_{i=1}^{n_t} A_{i,t} \varepsilon_{i,t}(1) - \frac{1}{N_{0,t}} \sum_{i=1}^{n_t} (1 - A_{i,t}) \varepsilon_{i,t}(0).$$

B.2 Martingale difference representation

For each batch t, define the within-batch filtration

$$\mathcal{F}_{t,0} = \mathcal{G}_t, \qquad \mathcal{F}_{t,i} = \mathcal{G}_t \vee \sigma(Y_{1,t}, \ldots, Y_{i,t}), \quad i = 1, \ldots, n_t.$$

The filtrations are concatenated across batches in lexicographic order. Thus $\mathcal{F}_{t,i-1}$ contains the history before batch t, the full assignment vector in batch t, and the first $i-1$ outcomes $Y_{1,t}, \ldots, Y_{i-1,t}$, but not the current outcome $Y_{i,t}$. Therefore the current centered shock still has conditional mean zero.

Define the realized centered outcome shock

$$\widetilde{\varepsilon}_{i,t} = (1 - A_{i,t})\varepsilon_{i,t}(0) + A_{i,t}\varepsilon_{i,t}(1).$$

Because $A_{i,t} \in \{0, 1\}$, we have $A_{i,t}^2 = A_{i,t}$, $(1 - A_{i,t})^2 = 1 - A_{i,t}$, and $A_{i,t}(1 - A_{i,t}) = 0$, so the batchwise estimation error can equivalently be expressed as

$$\widehat{\Delta}_t - \Delta_t = \sum_{i=1}^{n_t} \left(\frac{A_{i,t}}{N_{1,t}} - \frac{1 - A_{i,t}}{N_{0,t}} \right) \widetilde{\varepsilon}_{i,t}.$$

The statistic Z_{het} is the $1/\sqrt{T}$-normalized sum of the batchwise studentized estimation errors. We therefore define the individual contribution

$$D_{t,i} = \frac{1}{\sqrt{T}} \frac{1}{\sqrt{v_t}} \left(\frac{A_{i,t}}{N_{1,t}} - \frac{1 - A_{i,t}}{N_{0,t}} \right) \widetilde{\varepsilon}_{i,t}.$$

Since $\Delta_t = \delta_{0,t}$ under the batchwise null,

$$Z_{\text{het}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{\widehat{\Delta}_t - \Delta_t}{\sqrt{v_t}} = \sum_{t=1}^{T} \sum_{i=1}^{n_t} D_{t,i}.$$

Since $A_{i,t}$, $N_{1,t}$, $N_{0,t}$, and v_t are $\mathcal{F}_{t,i-1}$-measurable, only the centered outcome shock $\widetilde{\varepsilon}_{i,t}$ remains random at step i. By Assumption 1, the potential-outcome shocks are independent of the batch-t randomization and independent across individuals within the batch, conditional on H_{t-1}. By Assumption 2, these shocks have conditional mean zero:

$$E[\widetilde{\varepsilon}_{i,t} \mid \mathcal{F}_{t,i-1}] = 0.$$

Therefore,

$$E[D_{t,i} \mid \mathcal{F}_{t,i-1}] = 0,$$

so $\{D_{t,i}\}$ is a triangular array of martingale differences.

B.3 Predictable quadratic variation

The conditional second moment of $D_{t,i}$ is

$$E[D_{t,i}^2 \mid \mathcal{F}_{t,i-1}] = \frac{1}{T} \frac{1}{v_t} \left[\frac{A_{i,t}\, \sigma_{1,t}^2}{N_{1,t}^2} + \frac{(1 - A_{i,t})\, \sigma_{0,t}^2}{N_{0,t}^2} \right].$$

Summing over i within batch t:

$$\sum_{i=1}^{n_t} E[D_{t,i}^2 \mid \mathcal{F}_{t,i-1}] = \frac{1}{T} \frac{1}{v_t} \left[\frac{\sigma_{1,t}^2}{N_{1,t}} + \frac{\sigma_{0,t}^2}{N_{0,t}} \right] = \frac{1}{T}.$$

Therefore,

$$\sum_{t=1}^{T} \sum_{i=1}^{n_t} E[D_{t,i}^2 \mid \mathcal{F}_{t,i-1}] = 1.$$

The predictable quadratic variation is exactly normalized to one. Because of the local studentization no random quantity remains in the variance accumulation.

B.4 Lindeberg condition

To rule out large individual observations, the clipping condition is relevant. It ensures that both treatment arms receive a number of observations proportional to n_t, so no single observation receives too much weight. By Assumption 3 and the conditional law of large numbers, with probability approaching one there exists $\underline{\kappa} > 0$ such that $N_{1,t} \geq \underline{\kappa} n_t$, $N_{0,t} \geq \underline{\kappa} n_t$ for all t. On this event, together with the uniform variance bounds in Assumption 2, $v_t = \frac{\sigma_{1,t}^2}{N_{1,t}} + \frac{\sigma_{0,t}^2}{N_{0,t}} \asymp n_t^{-1}$. This means that v_t is bounded above and below by constant multiples of $1/n_t$. While $N_{1,t}^{-1}$ and $N_{0,t}^{-1}$ are both of order n_t^{-1}, $v_t^{-1/2} \asymp n_t^{1/2}$. Therefore each coefficient multiplying a centered outcome in $D_{t,i}$ is of order $n_t^{-1/2}$.

Using the bounded $(2+\delta)$-th moment condition in Assumption 2, there exists $C < \infty$ such that

$$\sum_{t=1}^{T} \sum_{i=1}^{n_t} E\big[|D_{t,i}|^{2+\delta} \mid \mathcal{F}_{t,i-1}\big] \leq C \sum_{t=1}^{T} n_t^{-\delta/2} \longrightarrow 0.$$

This Lyapunov condition rules out large individual contributions and, by Markov's inequality, implies the martingale Lindeberg condition: for every $\varepsilon > 0$,

$$\sum_{t=1}^{T} \sum_{i=1}^{n_t} E\big[D_{t,i}^2 \, \mathbb{1}\{|D_{t,i}| > \varepsilon\} \mid \mathcal{F}_{t,i-1}\big] \xrightarrow{p} 0.$$

The martingale Lindeberg condition rules out asymptotically relevant outliers.

B.5 Martingale central limit theorem

The triangular-array martingale difference sequence satisfies: (i) the predictable quadratic variation equals 1 exactly, and (ii) the Lindeberg condition holds. By the triangular-array martingale CLT (Hall and Heyde, 1980, Theorem 3.2),

$$Z_{\text{het}} = \sum_{t=1}^{T} \sum_{i=1}^{n_t} D_{t,i} \xrightarrow{d} \mathcal{N}(0,1).$$

This proves the infeasible part of Theorem 1. □

Remark (Stacked scores and other projections). The same argument also gives asymptotic normality of the stacked batchwise test statistic

$$\mathbf{z}_{\text{het}} = \left(\frac{\widehat{\Delta}_1 - \Delta_1}{\sqrt{v_1}}, \ldots, \frac{\widehat{\Delta}_T - \Delta_T}{\sqrt{v_T}} \right)' \xrightarrow{d} N(0, I_T).$$

Indeed, replacing the scalar weight $1/\sqrt{T}$ in $D_{t,i}$ by any fixed coefficient c_t gives

$$c'\mathbf{z}_{\text{het}} = \sum_{t=1}^{T} c_t \frac{\widehat{\Delta}_t - \Delta_t}{\sqrt{v_t}} \xrightarrow{d} N(0, c'c),$$

so the claim follows by Cramér–Wold. The statistic Z_{het} is the special case $c = T^{-1/2}\mathbf{1}_T$ under the batchwise null $\Delta_t = \delta_{0,t}$ for all t. For a common treatment effect, $\Delta_t = \Delta$, this is the natural projection. Other pre-specified projections are mainly useful for testing time variation in Δ_t, such as linear trends or early-versus-late contrasts.

B.6 Feasible statistic

By Assumption 4, $\hat{\sigma}^2_{a,t} \xrightarrow{p} \sigma^2_{a,t}$. Since T is fixed and, with probability approaching one, the treatment counts are bounded away from zero at rate n_t, it follows that $\max_{1\le t\le T} |\hat{v}_t/v_t - 1| \xrightarrow{p} 0$. Since

$$\frac{\hat{\Delta}_t - \Delta_t}{\sqrt{\hat{v}_t}} = \frac{\hat{\Delta}_t - \Delta_t}{\sqrt{v_t}} \left(\frac{v_t}{\hat{v}_t}\right)^{1/2},$$

the infeasible batch scores are tight and the multiplicative factor converges to one uniformly over fixed T. Hence $\hat{Z}_{\text{het}} - Z_{\text{het}} = o_p(1)$, and by Slutsky's theorem, $\hat{Z}_{\text{het}} \xrightarrow{d} \mathcal{N}(0,1)$. □