

Supplementary Material to “Covariate-Adjusted Response-Adaptive Design with Delayed Outcomes”

Xinwei Ma

Jingshen Wang

Waverly Wei

Department of Economics

Division of Biostatistics

Department of Data Sciences
and Operations

University of California San Diego

University of California Berkeley

University of Southern California

December 9, 2025

Contents

1	Setup, Notation and Assumptions	1
2	Semiparametric Efficiency Bound	2
3	Treatment Effect Estimation and Statistical Inference	3
4	Numerical Algorithm	6
5	Allocation Consistency and Verification of Condition 1	11
6	Additional simulation results	14
7	Proofs	17
7.1	Proof of Theorem 1	17
7.2	Proof of Lemma 1	20
7.3	Proof of Lemma 2	21
7.4	Proof of Theorem 2	23
7.5	Proof of Theorem 3	23
7.6	Proof of Theorem 4	26
7.7	R code for Algorithm 1	26
7.8	Proof of Lemma 3	28
7.9	Proof of Theorem 5	28
7.10	Proof of Corollary 1	28
7.11	Proof of Theorem 6	28
7.12	Proof of Theorem 7	29

1 Setup, Notation and Assumptions

We begin by recalling our adaptive experiment setting: There are T stages, which are labeled by $t = 1, 2, \dots, T$. In each stage t , n_t participants are enrolled. Participants are labeled by $i = 1, 2, \dots, N$, with $N = \sum_{t=1}^T n_t$ being the total sample size. We employ the notation, $S_i = t$ to indicate that the i th individual is enrolled in stage t .

For each individual, we observe some baseline covariates, $X_i \in \mathcal{X}$. The observed outcome variable is denoted by Y_i . As individuals are randomized into control and treatment arms, we use $A_i \in \mathcal{A} = \{0, 1\}$ to represent the actual treatment status. Lastly, since the outcome information may not be immediately available after the treatment is assigned, we use D_i to represent the number of stages after which the experimenter can observe Y_i . As a result of our definition, Y_i is observed at the end of stage t if and only if $D_i + S_i \leq t$. Following the Neyman-Rubin framework, the potential outcomes are denoted by $Y_i(a)$ for different treatments $a \in \mathcal{A} = \{0, 1\}$.

In what follows, we collect some notation and the main assumptions adopted in this supplementary material as well as in the main paper.

Assumption 1.

- (i) *The covariates and potential outcomes, $\{X_i, \{Y_i(a) : a \in \mathcal{A}\}\}$, are independent and identically distributed across $i = 1, 2, \dots, N$.*
- (ii) *The covariates have a finite support, that is, $|\mathcal{X}| < \infty$. Let*

$$p(x) = \mathbb{P}[X_i = x],$$

then $\min_{x \in \mathcal{X}} p(x) > 0$.

- (iii) *Let*

$$\mu_s(x, a) = \mathbb{E}[Y_i(a)^s | X_i = x].$$

Then $\max_{x \in \mathcal{X}} \max_{a \in \mathcal{A}} \mu_4(x, a) < \infty$. ■

For future reference, we also define the unconditional moments of the potential outcomes as

$$\mu_s(a) = \mathbb{E}[Y_i(a)^s].$$

Conditional and unconditional average treatment effects are

$$\tau(x) = \mu_1(x, 1) - \mu_1(x, 0), \quad \tau = \mu_1(1) - \mu_1(0).$$

Conditional variances of the potential outcomes are denoted by

$$\sigma^2(x, a) = \mathbb{V}[Y_i(a) | X_i = x, A_i = a] = \mu_2(x, a) - \mu_1(x, a)^2.$$

We remark that in the main paper, we drop the subscript whenever $s = 1$; that is, for the

mean of the potential outcomes, we use

$$\mu(x, a) = \mu_1(x, a) = \mathbb{E}[Y_i(a)|X_i = x] \text{ and } \mu(a) = \mu_1(a) = \mathbb{E}[Y_i(a)].$$

The next assumption concerns the treatment assignment and the delay mechanism. Recall that $\mathcal{H}_t$ represents the history up to time t , and $\mathcal{H}_0$ is the trivial sigma-algebra.

Assumption 2.

(i) Conditioning on $S_i = t$, X_i , and $\mathcal{H}_{t-1}$, treatment A_i is randomly assigned, and

$$\mathbb{P}[A_i = a|X_i = x, S_i = t, \mathcal{H}_{t-1}] = \hat{e}_t^*(a|x),$$

where $\hat{e}_t^*(\cdot|x) \in \mathcal{H}_{t-1}$.

(ii) The delay mechanism is independent of the potential outcomes after conditioning on the covariates and the treatment assignment:

$$D_i \perp\!\!\!\perp \{Y_i(a) : a \in \mathcal{A}\} \mid X_i, A_i, S_i.$$

In addition, the delay mechanism is time-homogeneous:

$$\rho(d|x, a) = \mathbb{P}[D_i \leq d|X_i = x, A_i = a] = \mathbb{P}[D_i \leq d|X_i = x, A_i = a, S_i = t]. \quad \blacksquare$$

We also assume the treatment assignment probability is bounded away from zero, and that there is a strictly positive probability that the outcome information can be immediately observed.

Assumption 3.

(i) With probability 1, $\min_{x \in \mathcal{X}} \min_{a \in \mathcal{A}} \min_{t \leq T} \hat{e}_t^*(a|x) \geq \delta$ for some $\delta > 0$.

(ii) $\min_{x \in \mathcal{X}} \min_{a \in \mathcal{A}} \rho(0|x, a) > 0$. \blacksquare

Finally, we consider an asymptotic regime in which T is fixed, and $n_t \rightarrow \infty$. More precisely, we assume

Assumption 4.

The total number of stages, T , is fixed, and for all t ,

$$\frac{n_t}{N} \rightarrow r_t > 0.$$

Here, $N = \sum_{t=1}^T n_t$ is the total sample size. \blacksquare

2 Semiparametric Efficiency Bound

In this section, we establish the semiparametric efficiency bound for treatment effect estimation in the presence of delayed outcomes.

Theorem 1.

Let Assumptions 1–4 hold. In addition, assume (i) S_i are independently and identically distributed with $\mathbb{P}[S_i = t] = r_t$; (ii) $\hat{e}_t^*(a|x) = e_t(a|x)$ for some $e_t(\cdot|\cdot)$ that does not depend on $\mathcal{H}_{t-1}$. Then the efficient influence function for estimating $\mu_1(a)$ is

$$\psi(X_i, A_i, D_i, Y_i, S_i|a) = \frac{\mathbf{1}_{(A_i=a, D_i+S_i \leq T)}}{\sum_{t=1}^T r_t e_t(a|X_i) \rho(T-t|X_i, a)} \left(Y_i(a) - \mu_1(X_i, a) \right) + \mu_1(X_i, a) - \mu_1(a).$$

In addition, the efficient influence function for estimating τ is

$$\begin{aligned} \psi(X_i, A_i, D_i, Y_i, S_i) &= \psi(X_i, A_i, D_i, Y_i, S_i|1) - \psi(X_i, A_i, D_i, Y_i, S_i|0) \\ &= \frac{\mathbf{1}_{(A_i=1, D_i+S_i \leq T)}}{\sum_{t=1}^T r_t e_t(1|X_i) \rho(T-t|X_i, 1)} \left(Y_i(1) - \mu_1(X_i, 1) \right) \\ &\quad - \frac{\mathbf{1}_{(A_i=0, D_i+S_i \leq T)}}{\sum_{t=1}^T r_t e_t(0|X_i) \rho(T-t|X_i, 0)} \left(Y_i(0) - \mu_1(X_i, 0) \right) + \tau(X_i) - \tau. \end{aligned}$$

As a result, the semiparametric efficiency bound for estimating τ is

$$\mathbf{V} = \sum_{x \in \mathcal{X}} p(x) \left[\frac{\sigma^2(x, 1)}{\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} + \frac{\sigma^2(x, 0)}{\sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0)} + (\tau(x) - \tau)^2 \right]. \quad \blacksquare$$

Condition (ii) assumes that the treatment allocation rule, e_t , is not random. That is, we treat e_t as “pre-determined” or “fixed” in our semiparametric efficiency calculation. Given the semiparametric efficiency bound, we develop our adaptive experimental design algorithm to minimize the asymptotic variance of the treatment effect estimators.

3 Treatment Effect Estimation and Statistical Inference

In this section, we show that the main treatment effect estimator is consistent following our proposed adaptive experimental design, and admits an asymptotically normal distribution. We also provide a valid variance estimator. Main conclusions of this section build on a “high-level” condition on the consistency of the optimized treatment allocation rule. We verify this high-level condition in the next subsection.

To start, we recall that the estimated mean potential outcome for a specific subgroup is

$$\hat{\mu}_{s,T}(x, a) = \frac{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} Y_i^s}{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)}}.$$

Notice that outcome information may not be observed for all observations. In our notation above,

this is captured by the indicator function $\mathbf{1}_{(D_i+S_i \leq T)}$, which equals 1 if and only if outcome information for this individual is available at the end of the experiment.

Whenever necessary, we also use the notation $\hat{\mu}_{s,t}(x, a)$ to represent the estimator formed with observations in the first t stages; that is,

$$\hat{\mu}_{s,t}(x, a) = \frac{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq t)} Y_i^s}{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq t)}}.$$

Consistency of the estimator follows from the lemma below.

Lemma 1.

Let Assumptions 1–4 hold. Then for all $s \leq 2$,

$$\hat{\mu}_{s,T}(x, a) = \mu_s(x, a) + O_p\left(\sqrt{\frac{1}{N}}\right). \quad \blacksquare$$

Next, we show that $\hat{\mu}_{s,T}(x, a)$ admits an asymptotic normal distribution. To this end, we employ the following high-level condition, which we verify for our proposed design in the next section.

Condition 1.

There exists nonrandom $e_t(\cdot|\cdot)$, such that $\hat{e}_t^*(a|x) = e_t(\cdot|\cdot) + o_p(1)$ for all $a \in \mathcal{A}$ and all $x \in \mathcal{X}$. $\blacksquare$

Lemma 2.

Let Assumptions 1–4 hold. Then under Condition 1,

$$\begin{aligned} & \sqrt{N}(\hat{\mu}_{1,T}(x, a) - \mu_1(x, a)) \\ &= \frac{1}{p(x) \sum_{t=1}^T r_t e_t(a|x) \rho(T-t|x, a)} \frac{1}{\sqrt{N}} \sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} (Y_i - \mu_1(x, a)) + o_p(1) \\ &\stackrel{d}{\rightarrow} \mathcal{N}\left(0, \frac{\sigma^2(x, a)}{p(x) \sum_{t=1}^T r_t e_t(a|x) \rho(T-t|x, a)}\right) \end{aligned}$$

for $a \in \mathcal{A} = \{0, 1\}$. $\blacksquare$

Define estimated subgroup treatment effect:

$$\hat{\tau}_T(x) = \hat{\mu}_{1,T}(x, 1) - \hat{\mu}_{1,T}(x, 0),$$

then it is immediate from Lemma 2 that $\hat{\tau}_T(x)$ also admits an asymptotic normal distribution.

Theorem 2.

Let Assumptions 1–4 hold. Then under Condition 1,

$$\sqrt{N}(\hat{\tau}_T(x) - \tau(x)) \stackrel{d}{\rightarrow} \mathcal{N}\left(0, \frac{1}{p(x)} \left(\frac{\sigma^2(x, 1)}{\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} + \frac{\sigma^2(x, 0)}{\sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0)} \right) \right). \quad \blacksquare$$

Finally, we consider the estimated treatment effect:

$$\hat{\tau}_T = \sum_{x \in \mathcal{X}} \hat{p}_T(x) \hat{\tau}_T(x),$$

where

$$\hat{p}_T(x) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(X_i=x)}.$$

Theorem 3.

Let Assumptions 1–4 hold. Then under Condition 1,

$$\sqrt{N}(\hat{\tau}_T - \tau) \xrightarrow{d} \mathcal{N}(0, \mathbf{V}),$$

where

$$\mathbf{V} = \sum_{x \in \mathcal{X}} p(x) \left[\frac{\sigma^2(x, 1)}{\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} + \frac{\sigma^2(x, 0)}{\sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0)} + (\tau(x) - \tau)^2 \right]. \quad \blacksquare$$

Following the asymptotic normality result established in the above theorem, we now introduce a variance estimator. Define

$$\hat{\sigma}_T(x, a) = \hat{\mu}_{2,T}(x, a) - \hat{\mu}_{1,T}(x, a)^2$$

for $x \in \mathcal{X}$ and $a \in \mathcal{A}$. Next, let

$$\hat{e}_T(x, a) = \frac{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)}}{\sum_{i=1}^N \mathbf{1}_{(X_i=x)}}.$$

Finally, we define

$$\hat{\mathbf{V}}_T = \sum_{x \in \mathcal{X}} \hat{p}_T(x) \left[\frac{\hat{\sigma}_T^2(x, 1)}{\hat{e}_T(x, 1)} + \frac{\hat{\sigma}_T^2(x, 0)}{\hat{e}_T(x, 0)} + (\hat{\tau}_T(x) - \hat{\tau}_T)^2 \right].$$

The result below establishes valid statistical inference.

Theorem 4.

Let Assumptions 1–4 hold. Then under Condition 1,

$$\hat{\mathbf{V}}_T = \mathbf{V}_T + o_p(1).$$

Therefore,

$$\frac{\hat{\tau}_T - \tau}{\sqrt{\hat{\mathbf{V}}_T/N}} \xrightarrow{d} \mathcal{N}(0, 1).$$

■

4 Numerical Algorithm

Minimizing the objective function can be challenging, especially when T is large. As part of our methodological development, we first discuss a dimension reduction technique that maps the original optimization into a one dimensional convex problem, thereby greatly allowing fast and scalable implementation. To convey the main idea, we first consider the oracle problem, for which the underlying data-generating process (the potential outcomes distribution and the delay mechanism) is assumed to be known. However, we remark that the same algorithm can be applied to compute the feasible allocations in our fully forward-looking method, to be discussed below. To present the algorithm in its full generality, consider the following objective function

$$\mathbf{V}(e_1, e_2, \dots, e_J) := \frac{\mathbf{c}_1}{\mathbf{a}_0 + \mathbf{a}_1 e_1 + \dots + \mathbf{a}_J e_J} + \frac{\mathbf{c}_0}{\mathbf{b}_0 + \mathbf{b}_1(1 - e_1) + \dots + \mathbf{b}_J(1 - e_J)},$$

with the constraints that $e_j \in [\delta, 1 - \delta]$ for some pre-specified $\delta \in [0, 1/2)$. To gain some intuition, we can easily map the above objective function to the semiparametric efficiency bound, as the example below illustrates.

Example 1.

We will fix a specific covariate value x . Set $J = T$, and

$$\mathbf{c}_1 = \sigma^2(x, 1), \quad \mathbf{c}_0 = \sigma^2(x, 0), \quad \mathbf{a}_j = r_j \rho(T - j|x, 1), \quad \mathbf{b}_j = r_j \rho(T - j|x, 0).$$

Then the optimization problem can be recast as $\min_{e_1, e_2, \dots, e_T} \mathbf{V}(e_1, \dots, e_T)$. We also introduce two additional constants $\mathbf{a}_0$ and $\mathbf{b}_0$ into the objective function to accommodate a pilot stage in which the treatment allocation is fixed (say, to 0.5 or $1/|\mathcal{A}|$). However, this is not required: our algorithm proposed below allows $\mathbf{a}_0 = \mathbf{b}_0 = 0$. ■

Our strategy is to first re-parameterize the problem as

$$\min_{\alpha, \beta} \mathbf{V}(\alpha, \beta) := \mathbf{c}_1/\alpha + \mathbf{c}_0/\beta, \quad \text{subject to } (\alpha, \beta) \in \mathcal{V} \subset \mathbb{R}^2.$$

The set $\mathcal{V}$ is determined by the linear mappings

$$(e_1, \dots, e_J) \mapsto \alpha(e_1, \dots, e_J) := \mathbf{a}_0 + \mathbf{a}_1 e_1 + \dots + \mathbf{a}_J e_J,$$

and

$$(e_1, \dots, e_J) \mapsto \beta(e_1, \dots, e_J) := \mathbf{b}_0 + \mathbf{b}_1(1 - e_1) + \dots + \mathbf{b}_J(1 - e_J).$$

Algorithm 1 Optimal treatment allocation with dimension reduction

See Section 7.7 for R implementation

Initialization:

- 1: Fix the constants $c_1, c_0, a_0, \dots, a_J, b_0, \dots, b_J$.
- 2: Let $\ell(1), \dots, \ell(J)$ be a permutation of $1, \dots, J$, such that

$$\frac{a_{\ell(1)}}{b_{\ell(1)}} \leq \frac{a_{\ell(2)}}{b_{\ell(2)}} \leq \dots \leq \frac{a_{\ell(J)}}{b_{\ell(J)}}.$$

- 3: Define a point in $\mathbb{R}^2$, (α_0, β_0) , as

$$\alpha_0 = a_0 + (1 - \delta) \sum_{j=1}^J a_j, \quad \beta_0 = b_0 + \delta J.$$

Characterizing boundaries of $\mathcal{V}$:

- 4: **for** j from 1 to J **do**
- 5: Define a point in $\mathbb{R}^2$, (α_j, β_j) , as

$$\alpha_j = a_0 + (1 - \delta) \sum_{s=j+1}^J a_{\ell(s)} + \delta j, \quad \beta_j = b_0 + (1 - \delta) \sum_{s=1}^j b_{\ell(s)} + \delta(J - j).$$

- 6: Define a point in $\mathbb{R}^2$, $(\underline{\alpha}_j, \underline{\beta}_j)$, as

$$\underline{\alpha}_j = a_0 + (1 - \delta) \sum_{s=1}^{J-j} a_{\ell(s)} + \delta j, \quad \underline{\beta}_j = b_0 + (1 - \delta) \sum_{s=J-j+1}^J b_{\ell(s)} + \delta(J - j).$$

- 7: **end for**

- 8: Upper boundary of $\mathcal{V}$: connect points $(\alpha_0, \beta_0) \cdots (\alpha_1, \beta_1) \cdots (\alpha_2, \beta_2) \cdots (\alpha_J, \beta_J)$.

- 9: Lower boundary of $\mathcal{V}$: connect points $(\underline{\alpha}_0, \underline{\beta}_0) \cdots (\underline{\alpha}_1, \underline{\beta}_1) \cdots (\underline{\alpha}_2, \underline{\beta}_2) \cdots (\underline{\alpha}_J, \underline{\beta}_J)$.

Optimization:

- 10: Solve the optimization problem, $\min_{\alpha, \beta} V(\alpha, \beta)$ subject to $(\alpha, \beta) \in \mathcal{V}$.
 - 11: With the solution (α^*, β^*) , find the corresponding $e_1^*, \dots, e_J^*$.
-

The two denominators, α and β , can also be interpreted as the probability of observing the outcome in the two treatment arms, respectively, given a specific propensity score profile (treatment assignment rule). As a key ingredient of our methodological development, Algorithm 1 characterizes $\mathcal{V}$ as a convex set with piecewise linear boundaries in $\mathbb{R}^2$, which makes the optimization problem fast and scalable. The main idea behind the algorithm is to compare the (estimated) delay mechanism across the two treatment arms. If it is anticipated that units in one arm tend to exhibit longer delays, then they are allocated to this arm in relatively early stages. Also see Section 7.7 for R implementation of the algorithm.

The last step in our algorithm finds the optimal treatment allocation from the solution $(\alpha^*, \beta^*) \in \mathbb{R}^2$. Before discussing this step in more detail, the following result formally justifies the validity of our algorithm.

Lemma 3.

The optimization problem of $\min_{\alpha, \beta} V(\alpha, \beta)$ subject to $(\alpha, \beta) \in \mathcal{V}$ admits a unique solution, and the solution lies on the upper boundary of $\mathcal{V}$, which consists of line segments connecting points $(\alpha_0, \beta_0) \cdots (\alpha_1, \beta_1) \cdots (\alpha_J, \beta_J)$ specified in the algorithm. ■

We highlight that our algorithm and the theoretical result above apply in broader settings concerning treatment allocations in adaptive experiments. In particular, our dimension reduction technique, which maps the original problem into a univariate optimization, is applicable provided that the objective function takes a “single-index structure” where the allocation probabilities enter via a linear index.

The uniqueness of the solution to the minimization of $V(\alpha, \beta)$ does not imply that the original minimization problem also admits a unique solution. We further characterize the solution to the original problem.

Theorem 5.

Assume the solution lies on a line segment connecting $(\alpha, \beta_j) \cdots (\alpha_{j+1}, \beta_{j+1})$, then

$$\frac{a_s}{b_s} < \frac{a_{\ell(j+1)}}{b_{\ell(j+1)}} \Rightarrow e_s^* = \delta, \quad \frac{a_s}{b_s} > \frac{a_{\ell(j+1)}}{b_{\ell(j+1)}} \Rightarrow e_s^* = 1 - \delta.$$

The solution $(e_1^*, \dots, e_T^*)$ is unique if and only if there does not exist s such that $\frac{a_s}{b_s} = \frac{a_{\ell(j+1)}}{b_{\ell(j+1)}}$. ■

In applications, the oracle treatment assignment is not feasible as it relies on unknown features of potential outcomes, such as their conditional second moments and the delay mechanism. Therefore, a feasible algorithm must estimate these features using accrued information and adjust the treatment allocation sequentially. A key challenge is that in any experimental stage, certain “tail features” of the delay mechanism are not estimable. To give a concrete example, assume the researcher has collected data from two stages and is revising the treatment assignment for Stage 3. Employing available data, it is possible to estimate the probability of outcome information becoming immediately available or delayed by one stage. However, it is not possible to estimate, say, the probability of delayed outcomes by three stages. This makes global optimization of the efficiency bound generally impossible. To address this challenge, we propose two feasible algorithms for allocating treatments based on different specifications of the delay mechanism. Notably, our dimension reduction technique can be readily extended to create feasible algorithms that are both numerically stable and fast.

One-step forward algorithm

We first consider a one-step forward algorithm that uses collected experimental data to determine the optimal treatment allocation for the next one stage. To start, we initialize stage 1 of the experiment following the completely randomized design in line 1 of Algorithm 2. Fix some $t = 1, 2, \dots, T - 1$, and assume we have collected data for the first t stages of the experiment. We then

compute the optimal treatment allocation for the next stage, $t + 1$. Define the objective function

$$\hat{V}_t^{\text{one-step}}(e_{t+1}(\cdot)) = \sum_{x \in \mathcal{X}} \hat{p}_t(x) \left[\frac{\hat{\sigma}_t^2(x, 1)}{\sum_{\ell=1}^t \hat{\mathbf{a}}_\ell(x) \hat{e}_\ell^*(1|x) + \hat{\mathbf{a}}_{t+1}(x) e_{t+1}(x)} + \frac{\hat{\sigma}_t^2(x, 0)}{\sum_{\ell=1}^t \hat{\mathbf{b}}_\ell(x) \hat{e}_\ell^*(0|x) + \hat{\mathbf{b}}_{t+1}(x) (1 - e_{t+1}(x))} \right].$$

$$\text{with } \hat{p}_t(x) = \frac{\sum_{i=1}^N \mathbf{1}_{(X_i=x, S_i \leq t)}}{\sum_{\ell=1}^t n_\ell}, \quad \hat{\sigma}_t^2(x, a) = \frac{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq t)} (Y_i - \hat{\mu}_{1,t}(x, a))^2}{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq t)}},$$

where we recall that $\hat{\mu}_{1,t}(x, a)$ is simply the sample average of the outcome variable in subgroup x and treatment arm a . Other constants in the algorithm are defined analogously:

$$\begin{aligned} \hat{\mathbf{a}}_\ell(x) &= \hat{r}_\ell \hat{\rho}_t(t - \ell|x, 1), & \hat{\mathbf{a}}_{t+1}(x) &= \hat{r}_{t+1} \hat{\rho}_t(0|x, 1), \\ \hat{\mathbf{b}}_\ell(x) &= \hat{r}_\ell \hat{\rho}_t(t - \ell|x, 0), & \hat{\mathbf{b}}_{t+1}(x) &= \hat{r}_{t+1} \hat{\rho}_t(0|x, 0), \end{aligned}$$

with $\hat{r}_\ell = n_\ell/N$.

A key challenge is to estimate the delay mechanism. At the end of stage t , we are able to estimate part of the delay mechanism as

$$\hat{\rho}_t(d|x, a) = \sum_{\ell=0}^d \hat{\mathbb{P}}[D_i = \ell | X_i = x, A_i = a] = \sum_{\ell=0}^d \frac{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i=\ell, S_i \leq t-\ell)}}{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, S_i \leq t-\ell)}},$$

for $d = 0, 1, \dots, t-1$. However, no estimate for $\rho(t|x, a)$ will be available until the end of stage $t+1$. Therefore, to operationalize the one-step forward objective function, we propose two approaches: (1) **conservative**: $\hat{\rho}_t(t|x, a) = \hat{\rho}_t(t-1|x, a)$, assuming that the delayed outcomes will continue to be delayed in the next stage, and (2) **optimistic**: $\hat{\rho}_t(t|x, a) = 1$, which assumes all the delayed outcomes will be observed in the next stage.

We remark that the conservative extrapolation can be understood from a “minimax” perspective, where we sequentially minimize the asymptotic variance under the worst-case scenario of the delay mechanism. For this reason, we adopt this approach in most of our numerical experiments, as well as in the following section of this Supplementary Material.

At the end of stage T when the experiment concludes, we obtain the subgroup treatment effect estimates $\hat{\tau}_T(x)$ as well as the estimated ATE $\hat{\tau}_T$.

Algorithm 2 Feasible adaptive experimental design with one-step forward objective function

Initialization:

1: Enroll n_1 participants, and assign treatments with $\hat{e}_1^*(1|x) = \hat{e}_1^*(0|x) = \frac{1}{2}$;

Adaptive treatment allocation:

2: **for** $t \rightarrow 1$ to $T - 1$ **do**

3: Using $\mathcal{H}_t$, obtain $\hat{\sigma}_t(x, 1)$, $\hat{\sigma}_t(x, 0)$, $\hat{\rho}_t(d|x, 1)$, and $\hat{\rho}_t(d|x, 0)$ for $x \in \mathcal{X}$ and $d \leq t - 1$.

4: Estimate $\hat{\rho}_t(t|x, a)$ with either the **conservative** or the **optimistic** approach.

5: Construct the constants $\hat{a}_\ell(x)$ and $\hat{b}_\ell(x)$ for $\ell = 1, 2, \dots, t + 1$.

6: Minimize $\hat{V}_t^{\text{one-step}}$ to obtain $\hat{e}_{t+1}^*(x)$ for $x \in \mathcal{X}$.

7: Set $\hat{e}_{t+1}^*(1|x) = \hat{e}_{t+1}^*(x)$, and $\hat{e}_{t+1}^*(0|x) = 1 - \hat{e}_{t+1}^*(x)$.

8: Enroll a new set of n_{t+1} subjects and assign treatments.

9: **end for**

Fully forward-looking algorithm

Our previous algorithm optimizes an objective function that involves only one additional period, which can be quite different from the semiparametric efficiency bound. While this might be innocuous when it gets closer to the last stage of the experiment, the one-step method may not perform well in the early stages. A more suitable and more sophisticated approach will require optimizing all future treatment allocations in each design stage. Specifically, we define the following fully forward-looking objective function

$$\begin{aligned} & \hat{V}_t^{\text{forward}}(e_{t+1}(\cdot), \dots, e_T(\cdot)) \\ &= \sum_{x \in \mathcal{X}} \hat{p}_t(x) \left[\frac{\hat{\sigma}_t^2(x, 1)}{\sum_{\ell=1}^t \hat{a}_\ell(x) \hat{e}_\ell^*(1|x) + \sum_{\ell=t+1}^T \hat{a}_\ell(x) e_\ell(x)} + \frac{\hat{\sigma}_t^2(x, 0)}{\sum_{\ell=1}^t \hat{b}_\ell(x) \hat{e}_\ell^*(0|x) + \sum_{\ell=t+1}^T \hat{b}_\ell(x) (1 - e_\ell(x))} \right]. \end{aligned}$$

As the above algorithm is fully forward-looking at each stage, we expect it to be better aligned with the objective of variance minimization. It is also worth mentioning that our earlier Algorithm 1 can be employed to solve the forward-looking optimization problem: one simply needs to replace the unknown quantities in 1 by their estimates. It offers the same benefit of dimension reduction, which can be valuable in the early stages of the experiments. A feasible design algorithm employing the fully forward-looking algorithm is given below.

Similar to the one-step algorithm, the delay mechanism is not fully estimable at each stage. For example, at the end of stage t , one can only estimate $\rho(0|x, a), \dots, \rho(t-1|x, a)$ while the fully forward-looking algorithm requires estimates for $\rho(t|x, a), \dots, \rho(T-1|x, a)$. Several candidates are: (1) **conservative**: $\hat{\rho}_t(d|x, a) = \hat{\rho}_t(t-1|x, a)$ for all $d \geq t$; (2) **optimistic**: $\hat{\rho}_t(d|x, a) = 1$; and (3) **neutral**: $\hat{\rho}_t(d|x, a)$ linearly interpolates $\hat{\rho}_t(t-1|x, a)$ and 1 for $d = t, \dots, T-1$.

Algorithm 3 Feasible adaptive experimental design with fully forward-looking objective function

Initialization:

1: Enroll n_1 participants, and assign treatments with $\hat{e}_1^*(1|x) = \hat{e}_1^*(0|x) = \frac{1}{2}$;

Adaptive treatment allocation:

2: **for** $t \rightarrow 1$ to $T - 1$ **do**

3: Using $\mathcal{H}_t$, obtain $\hat{\sigma}_t(x, 1)$, $\hat{\sigma}_t(x, 0)$, $\hat{\rho}_t(d|x, 1)$, and $\hat{\rho}_t(d|x, 0)$ for $x \in \mathcal{X}$ and $d \leq t - 1$.

4: Estimate $\hat{\rho}_t(d|x, a)$ for $d \geq t$ with one of the three approaches discussed below.

5: Construct the constants $\hat{a}_\ell(x)$ and $\hat{b}_\ell(x)$ for $\ell = 1, 2, \dots, T$.

6: Minimize $\hat{V}_t^{\text{forward}}$ to obtain $\hat{e}_{t+1}^*(x), \dots, \hat{e}_T^*(x)$ for $x \in \mathcal{X}$.

7: Set $\hat{e}_{t+1}^*(1|x) = \hat{e}_{t+1}^*(x)$, and $\hat{e}_{t+1}^*(0|x) = 1 - \hat{e}_{t+1}^*(x)$.

8: Enroll a new set of n_{t+1} subjects and assign treatments.

9: **end for**

5 Allocation Consistency and Verification of Condition 1

In this section we discuss in more details theoretical properties of the optimized treatment allocation. As part of this endeavor, we will provide sufficient conditions under which the optimized treatment allocations converge to a nonrandom limit, that is, we will verify Condition 1. Together with the regularity conditions laid out earlier, such design consistency helps establish both consistency and asymptotic normality of the estimated treatment effects. We will adopt the fully forward-looking algorithm with the conservative extrapolation approach.

In what follows, we will define recursively two types of treatment allocations, denoted by $e_t^\dagger(\cdot|\cdot)$ and $e_t^*(\cdot|\cdot)$, respectively. Loosely speaking, $e_t^\dagger(\cdot|\cdot)$ corresponds to the optimized treatment allocation rule if one has access to an infinite sample but imperfect knowledge about the delay mechanism, while $e_t^*(\cdot|\cdot)$ is the optimized allocation if one has both an infinite sample and perfect knowledge about the delay mechanism. To start, we set

$$e_1^\dagger(1|x) = e_1^*(1|x) = \frac{1}{2}, \quad e_1^\dagger(0|x) = e_1^*(0|x) = \frac{1}{2}.$$

Now fix some $t = 1, 2, \dots, T - 1$: the first t stages of the experiment has concluded and we are optimizing treatment allocation for the next stage, $t + 1$. Define

$$\rho_t^\dagger(d|x, 1) = \begin{cases} \rho(d|x, 1) & \text{if } d < t \\ \rho(t-1|x, 1) & \text{if } d \geq t \end{cases}, \quad \rho_t^\dagger(d|x, 0) = \begin{cases} \rho(d|x, 0) & \text{if } d < t \\ \rho(t-1|x, 0) & \text{if } d \geq t \end{cases}.$$

Then define

$$\mathbf{V}_t^\dagger(e_{t+1}(\cdot), \dots, e_T(\cdot)) = \sum_{x \in \mathcal{X}} p(x) \left[\frac{\sigma^2(x, 1)}{\sum_{\ell=1}^t r_\ell \rho_t^\dagger(T - \ell|x, 1) e_\ell^\dagger(1|x) + \sum_{\ell=t+1}^T r_\ell \rho_t^\dagger(T - \ell|x, 1) e_\ell(x)} \right]$$

$$+ \frac{\sigma^2(x, 0)}{\sum_{\ell=1}^t r_\ell \rho_t^\dagger(T - \ell|x, 0) e_\ell^\dagger(0|x) + \sum_{\ell=t+1}^T r_\ell \rho_t^\dagger(T - \ell|x, 0)(1 - e_\ell(x))} \Big],$$

and

$$\begin{aligned} \mathbf{V}_t(e_{t+1}(\cdot), \dots, e_T(\cdot)) = \sum_{x \in \mathcal{X}} p(x) \Big[& \frac{\sigma^2(x, 1)}{\sum_{\ell=1}^t r_\ell \rho(T - \ell|x, 1) e_\ell^*(1|x) + \sum_{\ell=t+1}^T r_\ell \rho(T - \ell|x, 1) e_\ell(x)} \\ & + \frac{\sigma^2(x, 0)}{\sum_{\ell=1}^t r_\ell \rho(T - \ell|x, 0) e_\ell^*(0|x) + \sum_{\ell=t+1}^T r_\ell \rho(T - \ell|x, 0)(1 - e_\ell(x))} \Big]. \end{aligned}$$

Intuitively, $\mathbf{V}_t^\dagger$ can be understood as the large-sample analogue of the feasible objective function $\hat{\mathbf{V}}_t^{\text{forward}}$, while $\mathbf{V}_t$ is the oracle objective function that requires full knowledge of the data generating process and the delay mechanism.

The optimized treatment allocations for the next stage $t + 1$ are

$$e_{t+1}^\dagger(1|x) = 1 - e_{t+1}^\dagger(0|x) = \operatorname{argmin}_{e_{t+1}(\cdot)} \min_{e_{t+2}(\cdot), \dots, e_T(\cdot)} \mathbf{V}_t^\dagger(e_{t+1}(\cdot), \dots, e_T(\cdot)),$$

and

$$e_{t+1}^*(1|x) = 1 - e_{t+1}^*(0|x) = \operatorname{argmin}_{e_{t+1}(\cdot)} \min_{e_{t+2}(\cdot), \dots, e_T(\cdot)} \mathbf{V}_t(e_{t+1}(\cdot), \dots, e_T(\cdot)).$$

The following assumption will greatly help simplify some of the presentations. We remark that this assumption can be dropped with the expense of much lengthier proof and more cumbersome notation.

Assumption 5.

For any $d \neq d'$,

$$\frac{\rho(d|x, 1)}{\rho(d|x, 0)} \neq \frac{\rho(d'|x, 1)}{\rho(d'|x, 0)}.$$

■

We remark that it is still possible to have “ties” in $\rho_t^\dagger$ due to the conservative extrapolation we employed. Specifically, for $t \leq T/2$, the following always hold:

$$\frac{\rho_t^\dagger(T - t - 1|x, 1)}{\rho_t^\dagger(T - t - 1|x, 0)} = \frac{\rho_t^\dagger(T - t - 2|x, 1)}{\rho_t^\dagger(T - t - 2|x, 0)} = \dots = \frac{\rho_t^\dagger(t - 1|x, 1)}{\rho_t^\dagger(t - 1|x, 0)}.$$

Such ties will also arise when solving Algorithm 3, that is, for all $t \leq T/2$,

$$\frac{\hat{\mathbf{a}}_{t+1}(x)}{\hat{\mathbf{b}}_{t+1}(x)} = \frac{\hat{\mathbf{a}}_{t+2}(x)}{\hat{\mathbf{b}}_{t+2}(x)} = \dots = \frac{\hat{\mathbf{a}}_{T-t+1}(x)}{\hat{\mathbf{b}}_{T-t+1}(x)}$$

by construction. Therefore, we adopt the following tie-breaking rule:

Condition 2.

For all $t = 1, 2, \dots, T/2$,

$$\begin{aligned} \max \left\{ e_{t+2}^\dagger(1|x), \dots, e_{T-t+1}^\dagger(1|x) \right\} &> \delta \quad \text{only if} \quad e_{t+1}^\dagger(1|x) = 1 - \delta, \\ \max \left\{ \hat{e}_{t+2}^*(1|x), \dots, \hat{e}_{T-t+1}^*(1|x) \right\} &> \delta \quad \text{only if} \quad \hat{e}_{t+1}^*(1|x) = 1 - \delta. \end{aligned} \quad \blacksquare$$

While this condition may seem complicated, it is quite innocuous and is also straightforward to implement. Essentially, whenever the design algorithm (either the feasible allocation from minimizing the fully forward-looking objective function or its asymptotic analogue) dictates assigning some participants to the treatment arm in periods $t + 1, t + 2, \dots, T - t + 1$, we will prioritize period $t + 1$.

As a direct consequence of Theorem 5, the optimized allocations are unique. We formally state this result in the Corollary below.

Corollary 1.

Let Assumptions 1–5 and Condition 2 hold. Then both $e_t^\dagger(\cdot|\cdot)$ and $e_t^*(\cdot|\cdot)$ are unique for $t = 1, 2, \dots, T$. ■

The next theorem is our main design consistency result, which demonstrates that the empirically optimized treatment allocation converges in large samples.

Theorem 6.

Let Assumptions 1–5 and Condition 2 hold. Then Condition 1 holds with $e_t(\cdot|\cdot) = e_t^\dagger(\cdot|\cdot)$. ■

Before closing this section, we provide two insights regarding the variance minimization problem. The first result suggests that as more information about the delay mechanism becomes available in the adaptive experiment, the optimized asymptotic variance never increases. While this result may seem natural, we note that it relies on the use of conservative extrapolation. As discussed earlier, conservative extrapolation can be understood from a “minimax” perspective, where the asymptotic variance is sequentially minimized under the worst-case scenario of the delay mechanism.

The second result provides a bound on efficiency loss due to imperfect knowledge about the delay mechanism. Collectively, these two results provide theoretical guarantee to our proposed adaptive experimental design.

Theorem 7.

Let Assumptions 1–5 and Condition 2 hold. Define V^* as the minimized oracle asymptotic variance, and V to be the asymptotic variance of the estimated average treatment effect. Then (i) for all $1 \leq s < t \leq T - 1$,

$$V^* \leq V \leq \min_{e_{t+1}(\cdot), \dots, e_T(\cdot)} V_t^\dagger(e_{t+1}(\cdot), \dots, e_T(\cdot)) \leq \min_{e_{s+1}(\cdot), \dots, e_T(\cdot)} V_s^\dagger(e_{s+1}(\cdot), \dots, e_T(\cdot));$$

and (ii) for any $\varepsilon > 0$ and d_ε with $\min_{x \in \mathcal{X}, a \in \mathcal{A}} \rho(d_\varepsilon | x, a) \geq 1 - \varepsilon$, $V \leq V^* + C(\frac{d_\varepsilon}{T} + \varepsilon)$, where C does not depend on T , ε or d_ε . ■

6 Additional simulation results

We provide additional simulation evidence in this section on the finite-sample performance of the proposed fully forward-looking CARA design, focusing on the optimized variance under the power maximization objective across three extrapolation methods: neutral, optimistic, and conservative. The results are presented in Figures 1 (for $T = 4$) and 2 (for $T = 6$) where covariate stratification is based biological sex only, and in Figures 3 and 4 where we use both biological sex and WHO clinical stage.

We highlight a few key observations. First, under the conservative extrapolation method, the optimized variance tends to be large at the beginning and gradually decreases as the experiment progresses. This pattern aligns with Theorem 7 and reflects the “minimax” nature of the conservative approach. Specifically, the conservative method initially assumes that none of the missing outcomes will become available, leading to high variance early on. However, as more data are collected, the sequentially learned delay mechanism helps reduce the variance objective function.

In contrast, under the optimistic extrapolation method, the variance bound exhibits the opposite trend. Because the optimistic method assumes that missing outcomes will become available after one period, the initial variance tends to be small. Yet as information about the actual delay mechanism accumulates, the experimenter must adjust their expectation, leading to an increase in the variance bound. Finally, the neutral method, which uses linear extrapolation, shows relatively stable variance bounds over time. Encouragingly, we find that the final performance of our procedure, measured by the variance at the end of the experiment, is relatively robust across extrapolation methods. For comparison, the optimized oracle variance is also shown in the figures as a horizontal dashed line.

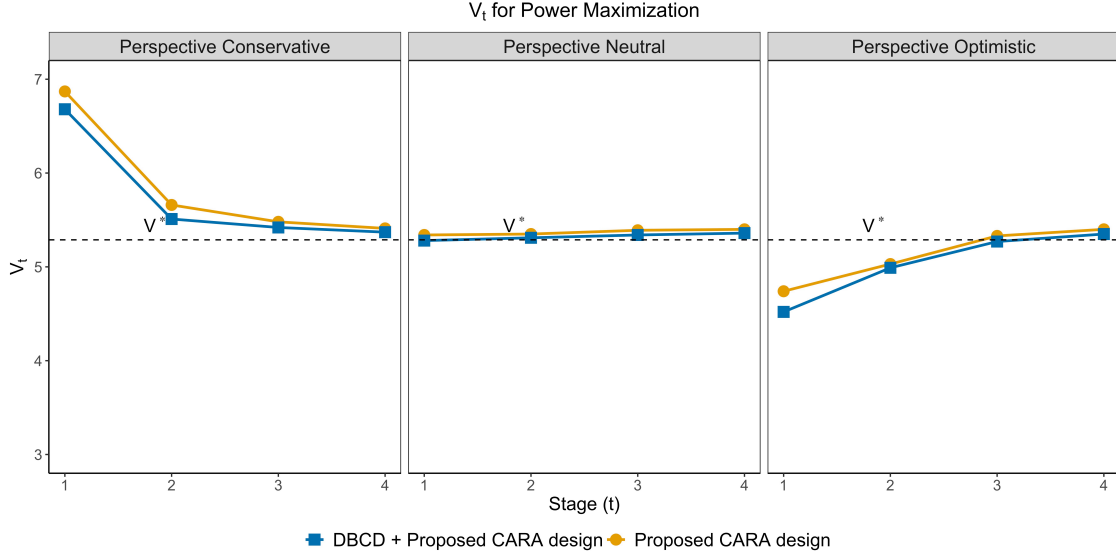


Figure 1. Optimized variance for the power maximization objective. The three panels correspond to the conservative, the neutral, and the optimistic approach of extrapolating the estimated delay mechanism. Horizon: $T = 4$. Covariate stratification is based on biological sex.

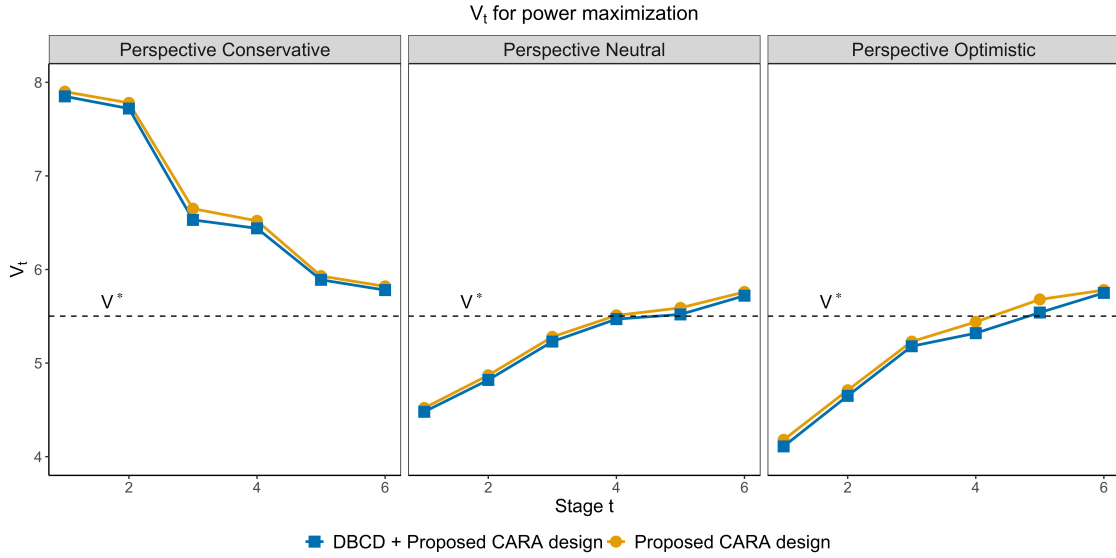


Figure 2. Optimized variance for the power maximization objective. The three panels correspond to the conservative, the neutral, and the optimistic approach of extrapolating the estimated delay mechanism. Horizon: $T = 6$. Covariate stratification is based on biological sex.

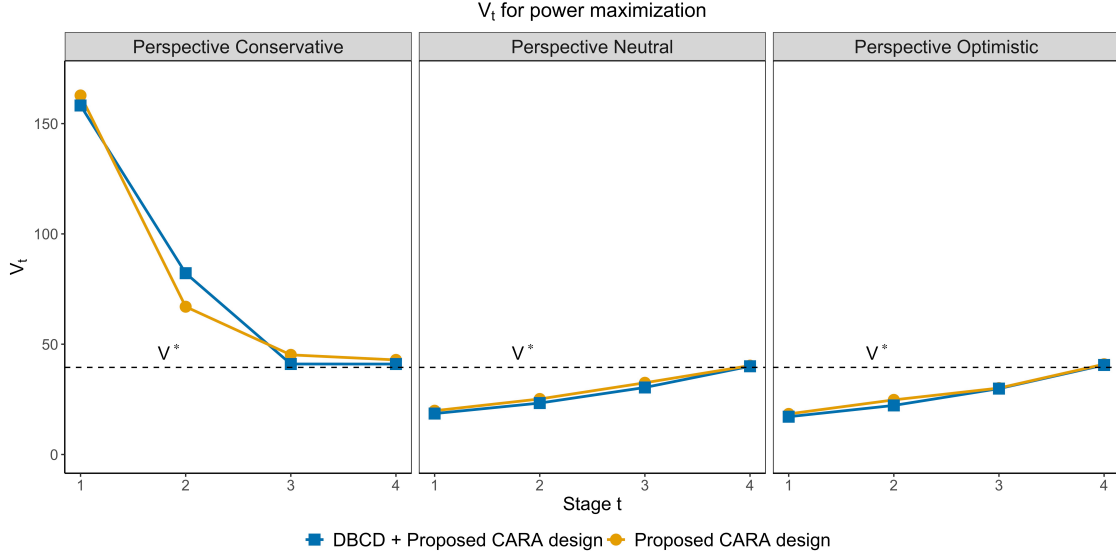


Figure 3. Optimized variance for the power maximization objective. The three panels correspond to the conservative, the neutral, and the optimistic approach of extrapolating the estimated delay mechanism. Horizon: $T = 4$. Covariate stratification is based on biological sex and WHO stage.

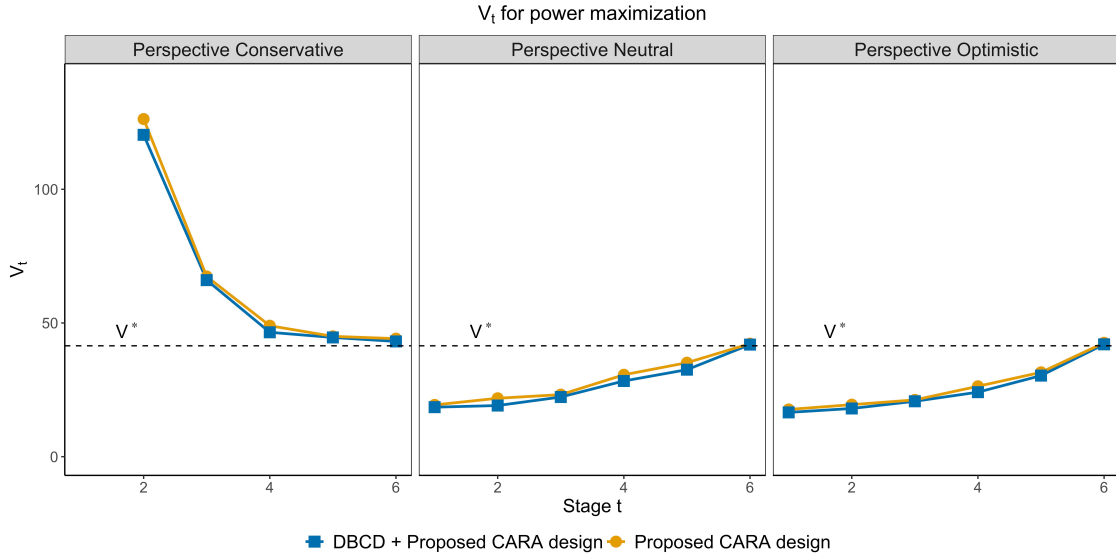


Figure 4. Optimized variance for the power maximization objective. The three panels correspond to the conservative, the neutral, and the optimistic approach of extrapolating the estimated delay mechanism. Horizon: $T = 6$. Covariate stratification is based on biological sex and WHO stage.

7 Proofs

7.1 Proof of Theorem 1

We omit the subscript i in the following derivation to save notation. We break down our proof into three steps: (1) characterizing the tangent space, (2) expressing the causal parameter τ and computing the pathwise derivative, and (3) projecting the resulting influence function onto the tangent space to obtain the efficient influence function. See [van der Vaart \(2000\)](#) and [Tsiatis \(2006\)](#) for textbook treatments of semiparametric efficiency theory.

Step 1. We first find the tangent space. To this end, we denote the joint density of the observed data as $f_\eta(X, A, D, Y, S)$, where the subscript η is adopted to represent and parameterize a specific one dimensional parametric submodel passing through the truth:

$$\begin{aligned} f_\eta(X, A, D, Y, S) \\ = p_\eta(X) \prod_{t=1}^T \left\{ r_{t,\eta} \prod_{a \in \mathcal{A}} \left[e_{t,\eta}(a|X) \left(\rho_\eta(T-t|X, a) f_{Y(a)|X,\eta}(Y|X) \right)^{\mathbb{1}_{(D \leq T-t)}} \right. \right. \\ \left. \left. \left(1 - \rho_\eta(T-t|X, a) \right)^{\mathbb{1}_{(D > T-t)}} \right]^{\mathbb{1}_{(A=a)}} \right\}^{\mathbb{1}_{(S=t)}}. \end{aligned}$$

Here, we recall that $r_{t,\eta}$ is the enrollment frequency of stage t , $p_\eta(\cdot)$ is the unconditional distribution of the covariates, $e_{t,\eta}(\cdot|\cdot)$ is the treatment assignment probability, $\rho_\eta(\cdot|\cdot, \cdot)$ is the delay mechanism, and $f_{Y(a)|X,\eta}(\cdot|\cdot)$ is the conditional distribution of the potential outcomes. We next write the above joint density as

$$\begin{aligned} f_\eta(X, A, D, Y, S) \\ = p_\eta(X) \times \prod_{a \in \mathcal{A}} \left(f_{Y(a)|X,\eta}(Y|X) \right)^{\mathbb{1}_{(A=a)} \sum_{t=1}^T \mathbb{1}_{(S=t)} \mathbb{1}_{(D \leq T-t)}} \\ \times \left(\prod_{t=1}^T r_{t,\eta}^{\mathbb{1}_{(S=t)}} \right) \\ \times \prod_{t=1}^T \prod_{a \in \mathcal{A}} \left(e_{t,\eta}(a|X) \right)^{\mathbb{1}_{(S=t)} \mathbb{1}_{(A=a)}} \\ \times \prod_{t=1}^T \prod_{a \in \mathcal{A}} \left(\rho_\eta(T-t|X, a) \right)^{\mathbb{1}_{(S=t)} \mathbb{1}_{(A=a)} \mathbb{1}_{(D \leq T-t)}} \left(1 - \rho_\eta(T-t|X, a) \right)^{\mathbb{1}_{(S=t)} \mathbb{1}_{(A=a)} \mathbb{1}_{(D > T-t)}}. \end{aligned}$$

The tangent space consists of scores for different submodels. Since each component of the joint distribution can independently vary across submodels, the tangent space takes the form of a direct sum of the following tangent subspaces; each tangent subspace corresponds to scores computed for

different components in the above joint density:

$$\begin{aligned}
\mathcal{T}_X &= \left\{ \mathfrak{f}_X(X) : \mathbb{E}[\mathfrak{f}_X(X)] = 0 \right\}; \\
\mathcal{T}_{Y(a)} &= \left\{ \left(\mathbb{1}_{(A=a)} \sum_{t=1}^T \mathbb{1}_{(S=t)} \mathbb{1}_{(D \leq T-t)} \right) \mathfrak{g}_a(Y|X) : \mathbb{E}[\mathfrak{g}_a(Y(a)|X)|X] = 0 \right\} \quad \text{for } a \in \mathcal{A} = \{0, 1\}; \\
\mathcal{T}_S &= \left\{ \sum_{t=1}^T \frac{\mathbb{1}_{(S=t)}}{r_t} \mathfrak{a}_t : \sum_{t=1}^T \mathfrak{a}_t = 0 \right\}; \\
\mathcal{T}_{A,t} &= \left\{ \mathbb{1}_{(S=t)} \left(\frac{\mathbb{1}_{(A=1)}}{e_t(1|X)} - \frac{\mathbb{1}_{(A=0)}}{e_t(0|X)} \right) \mathfrak{e}_t(X) : \mathfrak{e}_t(\cdot) \text{ general function} \right\}; \\
\mathcal{T}_{D,a,t} &= \left\{ \mathbb{1}_{(S=t)} \mathbb{1}_{(A=a)} \left(\frac{\mathbb{1}_{(D \leq T-t)}}{\rho(T-t|X, a)} - \frac{\mathbb{1}_{(D > T-t)}}{1 - \rho(T-t|X, a)} \right) \mathfrak{p}_{a,T-t}(X) : \mathfrak{p}_{a,T-t}(\cdot) \text{ general function} \right\} \\
&\quad \text{for } a \in \mathcal{A} = \{0, 1\}.
\end{aligned}$$

In the above, $\mathfrak{f}_X(\cdot)$, $\mathfrak{g}_a(\cdot|X)$, $\mathfrak{a}_t$, $\mathfrak{e}_t(\cdot)$, and $\mathfrak{p}_{a,T-t}(\cdot)$ denote general functions satisfying certain restrictions specified in each of the tangent subspaces. For example, $\mathcal{T}_X$ consists of functions of X that are mean zero.¹

It is routine to verify that the above spaces are mutually orthogonal (i.e., elements in the tangent subspaces are uncorrelated), which justifies that the whole tangent space can be written as their direct sum.

Step 2. In this step, consider the estimand $\mu_1(a) = \mathbb{E}[Y(a)]$ (for example, set $a = 1$ to convey the main idea). We will take pathwise derivative of $\mu_1(a)$ and then find an influence function. We start from the following

$$\mu_{1,\eta}(a) = \mathbb{E}_\eta \left[\frac{\mathbb{1}_{(S=1)} \mathbb{1}_{(A=a)} \mathbb{1}_{(D \leq T-1)} Y}{r_{1,\eta} e_{1,\eta}(a|X) \rho_\eta(T-1|X, a)} \right],$$

and again, η parameterizes a particular one-dimensional submodel. The pathwise derivative is

$$\begin{aligned}
\left. \frac{\partial \mu_{1,\eta}(a)}{\partial \eta} \right|_{\eta=0} &= \mathbb{E} \left[\left(\frac{\mathbb{1}_{(S=1)} \mathbb{1}_{(A=a)} \mathbb{1}_{(D \leq T-1)} Y}{r_1 e_1(a|X) \rho(T-1|X, a)} - \mu_1(a) \right) \mathfrak{s}(X, A, D, Y, S) \right] \\
&\quad - \mathbb{E} \left[\frac{\mathbb{1}_{(S=1)} \mathbb{1}_{(A=a)} \mathbb{1}_{(D \leq T-1)} Y}{\left(r_1 e_1(a|X) \rho(T-1|X, a) \right)^2} e_1(a|X) \rho(T-1|X, a) \mathfrak{a}_1 \right] \tag{I}
\end{aligned}$$

$$\begin{aligned}
&\quad - \mathbb{E} \left[\frac{\mathbb{1}_{(S=1)} \mathbb{1}_{(A=a)} \mathbb{1}_{(D \leq T-1)} Y}{\left(r_1 e_1(a|X) \rho(T-1|X, a) \right)^2} r_1 \rho(T-1|X, a) \mathfrak{e}_1(X) \right] \tag{II}
\end{aligned}$$

¹Strictly speaking, elements of a tangent space should also have finite second moment. We dropped such qualifications in the definition of the subspaces to save notation and conserve space. See [van der Vaart \(2000\)](#) for additional technical details.

$$- \mathbb{E} \left[\frac{\mathbf{1}_{(S=1)} \mathbf{1}_{(A=a)} \mathbf{1}_{(D \leq T-1)} Y}{\left(r_1 e_1(a|X) \rho(T-1|X, a) \right)^2} r_1 e_1(a|X) \mathfrak{p}_{a, T-1}(X) \right], \quad (\text{III})$$

and $\mathfrak{s}(\cdot, \cdot, \cdot, \cdot, \cdot)$ is the score for this particular parametric submodel. Then

$$\begin{aligned} (\text{I}) &= \mathbb{E} \left[- \frac{\mathbf{1}_{(S=1)} \mu_1(X, a)}{r_1^2} \mathfrak{a} \right] = - \mathbb{E} \left[\frac{\mathbf{1}_{(S=1)} \mu_1(X, a)}{r_1} \left(\sum_{t=1}^T \frac{\mathbf{1}_{(S=t)}}{r_t} \mathfrak{a}_t \right) \right] \\ &= \mathbb{E} \left[- \frac{(\mathbf{1}_{(S=1)} - r_1) \mu_1(X, a)}{r_1} \left(\sum_{t=1}^T \frac{\mathbf{1}_{(S=t)}}{r_t} \mathfrak{a}_t \right) \right] = \mathbb{E} \left[- \frac{(\mathbf{1}_{(S=1)} - r_1) \mu_1(X, a)}{r_1} \mathfrak{s}(X, A, D, Y, S) \right]. \end{aligned}$$

Similarly

$$\begin{aligned} (\text{II}) &= \mathbb{E} \left[- \frac{\mathbf{1}_{(S=1)} \mathbf{1}_{(A=a)} \mu_1(X, a)}{r_1 e_1(a|X)^2} \mathfrak{e}_1(X) \right] \\ &= \mathbb{E} \left[- \frac{\mathbf{1}_{(S=1)} \mathbf{1}_{(A=a)} \mu_1(X, a)}{r_1 e_1(a|X)} \mathbf{1}_{(S=1)} \left(\frac{\mathbf{1}_{(A=1)}}{e_1(1|X)} - \frac{\mathbf{1}_{(A=0)}}{e_1(0|X)} \right) \mathfrak{e}_1(X) \right] \\ &= \mathbb{E} \left[- \frac{\mathbf{1}_{(S=1)} (\mathbf{1}_{(A=a)} - e_1(a|X)) \mu_1(X, a)}{r_1 e_1(a|X)} \mathbf{1}_{(S=1)} \left(\frac{\mathbf{1}_{(A=1)}}{e_1(1|X)} - \frac{\mathbf{1}_{(A=0)}}{e_1(0|X)} \right) \mathfrak{e}_1(X) \right] \\ &= \mathbb{E} \left[- \frac{\mathbf{1}_{(S=1)} (\mathbf{1}_{(A=a)} - e_1(a|X)) \mu_1(X, a)}{r_1 e_1(a|X)} \mathfrak{s}(X, A, D, Y, S) \right]. \end{aligned}$$

Finally,

$$(\text{III}) = \mathbb{E} \left[- \frac{\mathbf{1}_{(S=1)} \mathbf{1}_{(A=a)} (\mathbf{1}_{(D \leq T-1)} - \rho(T-1|X, a)) \mu_1(X, a)}{r_1 e_1(a|X) \rho(T-1|X, a)} \mathfrak{s}(X, A, D, Y, S) \right].$$

Therefore by collecting terms,

$$\left. \frac{\partial \mu_{1, \eta}(a)}{\partial \eta} \right|_{\eta=0} = \mathbb{E} \left[\left(\phi_1(X, A, D, Y, S) + \phi_2(X) \right) \mathfrak{s}(X, A, D, Y, S) \right],$$

where

$$\phi_1(X, A, D, Y, S) = \frac{\mathbf{1}_{(S=1)} \mathbf{1}_{(A=a)} \mathbf{1}_{(D \leq T-t)} (Y - \mu_1(X, a))}{r_1 e_1(a|X) \rho(T-t|X, a)}, \quad \phi_2(X) = \mu_1(X, a) - \mu_1(a).$$

The above shows that the pathwise derivative can be written as an inner product with the score. As a result, $\phi_1(X, A, D, Y, S) + \phi_2(X)$ gives an influence function. The problem, however, is that ϕ_1 is not in the tangent space (note that ϕ_2 is in $\mathcal{T}_X$).

Step 3. In this step, we project ϕ_1 onto the tangent space to obtain the efficient influence function. To complete the efficient influence function, we need to find the projection of the first term to $\mathcal{T}_{Y(a)}$. (Note that the other tangent spaces are irrelevant, as ϕ_1 is already orthogonal to them.)

The projection is uniquely determined by

$$\mathbb{E} \left[\left(\phi_1(X, A, D, Y, S) - \left(\mathbf{1}_{(A=a)} \sum_{t=1}^T \mathbf{1}_{(S=t)} \mathbf{1}_{(D \leq T-t)} \right) \mathfrak{g}_a(Y|X) \right) \cdot \left(\mathbf{1}_{(A=a)} \sum_{t=1}^T \mathbf{1}_{(S=t)} \mathbf{1}_{(D \leq T-t)} \right) \mathfrak{g}_a(Y|X) \right] = 0.$$

First, we have

$$\mathbb{E} \left[\phi_1(X, A, D, Y, S) \left(\mathbf{1}_{(A=a)} \sum_{t=1}^T \mathbf{1}_{(S=t)} \mathbf{1}_{(D \leq T-t)} \right) \mathfrak{g}_a(Y|X) \right] = \mathbb{E}[(Y(a) - \mu_1(X, a)) \mathfrak{g}_a(Y|X)].$$

We also have

$$\mathbb{E} \left[\left(\left(\mathbf{1}_{(A=a)} \sum_{t=1}^T \mathbf{1}_{(S=t)} \mathbf{1}_{(D \leq T-t)} \right) \mathfrak{g}_a(Y|X) \right)^2 \right] = \mathbb{E} \left[\left(\sum_{t=1}^T r_t e_t(a|X) \rho(T-t|X, a) \right) \mathbb{E}[\mathfrak{g}_a(Y(a)|X)^2|X] \right].$$

Then we can simply set

$$\mathfrak{g}_a(Y(a)|X) = \frac{Y(a) - \mu_1(X, a)}{\sum_{t=1}^T r_t e_t(a|X) \rho(T-t|X, a)}.$$

7.2 Proof of Lemma 1

We start by writing

$$\begin{aligned} \hat{\mu}_{s,T}(x, a) - \mu_s(x, a) &= \frac{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} (Y_i^s - \mu_s(x, a))}{\sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)}} \\ &= \left(\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} \right)^{-1} \left(\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} (Y_i^s - \mu_s(x, a)) \right). \end{aligned}$$

Next, recall that individuals are ordered by the stage in which they are enrolled. Letting $N_t = \sum_{s=1}^t n_s$, we have

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} &= \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} \\ &= \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \left(\mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} - p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \right) \\ &\quad + \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a). \end{aligned}$$

Given our assumption that $p(x)$, $\rho(0|x, a)$, and the treatment probabilities are bounded from below,

$$\frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \gtrsim 1.$$

Next, we notice the variance satisfies

$$\begin{aligned} & \mathbb{V} \left[\frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \left(\mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} - p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \right) \right] \\ &= \frac{1}{N^2} \sum_{t=1}^T \mathbb{E} \left[\mathbb{V} \left[\sum_{i=N_{t-1}+1}^{N_t} \left(\mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} - p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \right) \middle| \mathcal{H}_{t-1} \right] \right] = O\left(\frac{1}{N}\right). \end{aligned}$$

As a result,

$$\left(\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} \right)^{-1} = O_p(1).$$

To conclude the proof, we again compute the variance

$$\begin{aligned} & \mathbb{V} \left[\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} \left(Y_i^s - \mu_s(x, a) \right) \right] \\ &= \mathbb{V} \left[\frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} \left(Y_i^s - \mu_s(x, a) \right) \right] \\ &= \frac{1}{N^2} \sum_{t=1}^T \mathbb{E} \left[\mathbb{V} \left[\sum_{i=N_{t-1}+1}^{N_t} \mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} \left(Y_i^s - \mu_s(x, a) \right) \middle| \mathcal{H}_{t-1} \right] \right] = O\left(\frac{1}{N}\right). \end{aligned}$$

The bound following the last inequality is due to our assumption on the finite fourth moment of the potential outcomes.

7.3 Proof of Lemma 2

Recall that individuals are ordered by the stage in which they are enrolled, and therefore we let $N_t = \sum_{s=1}^t n_s$. We start with the denominator, which can be rewritten as

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(X_i=x, A_i=a, D_i+S_i \leq T)} &= \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} \\ &= \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \left(\mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} - p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \right) \end{aligned}$$

$$+ p(x) \sum_{t=1}^T \frac{n_t}{N} \hat{e}_t^*(a|x) \rho(T-t|x, a).$$

Clearly the second term satisfies

$$p(x) \sum_{t=1}^T \frac{n_t}{N} \hat{e}_t^*(a|x) \rho(T-t|x, a) = p(x) \sum_{t=1}^T r_t e_t(a|x) \rho(T-t|x, a) + o_p(1)$$

by our condition on the consistency of treatment allocation.

The first term is easily shown to be mean zero by a standard martingale calculation. Its variance is

$$\begin{aligned} & \mathbb{V} \left[\frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \left(\mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} - p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \right) \right] \\ &= \frac{1}{N^2} \sum_{t=1}^T n_t \mathbb{E} \left[p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \left(1 - p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \right) \right] = O \left(\frac{1}{N} \right). \end{aligned}$$

As a result, we showed that

$$\frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} = p(x) \sum_{t=1}^T r_t e_t(a|x) \rho(T-t|x, a) + o_p(1).$$

We now analyze the numerator. It can be rewritten as

$$\frac{1}{\sqrt{N}} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} (Y_i - \mu_1(x, a)).$$

To establish asymptotic normality, we apply the martingale central limit theorem, which requires computing the conditional variance and verify a Lindeberg-Feller condition ([Hall and Heyde, 1980](#)).

The conditional variance is given by

$$\frac{1}{\sqrt{N}} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{V} \left[\mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} (Y_i - \mu_1(x, a)) \middle| \{(X_j, A_j, D_j, Y_j) : 1 \leq j < i\} \right].$$

Next, we notice that individuals enrolled in the same stage are independent of each other, and the treatment assignment only depends on accrued information up to the previous stage. As a result, the above reduces to

$$\frac{1}{\sqrt{N}} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{V} \left[\mathbf{1}_{(X_i=x, A_i=a, D_i \leq T-t)} (Y_i - \mu_1(x, a)) \middle| \{(X_j, A_j, D_j, Y_j) : 1 \leq j \leq N_{t-1}\} \right]$$

$$\begin{aligned}
&= \frac{1}{N} \sum_{t=1}^T n_t \mathbb{V} \left[\mathbb{1}_{(X_i=x, A_i=a, D_i \leq T-t)} (Y_i - \mu_1(x, a)) \middle| \mathcal{H}_{t-1} \right] \\
&= \sum_{t=1}^T \frac{n_t}{N} p(x) \hat{e}_t^*(a|x) \rho(T-t|x, a) \sigma^2(x, a) = p(x) \sigma^2(x, a) \sum_{t=1}^T r_t e_t(a|x) \rho(T-t|x, a) + o_p(1).
\end{aligned}$$

Finally, we verify the Lindeberg-Feller condition for the numerator with fourth conditional moments:

$$\begin{aligned}
&\frac{1}{N^2} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{E} \left[\left| \mathbb{1}_{(X_i=x, A_i=a, D_i \leq T-t)} (Y_i - \mu_1(x, a)) \right|^4 \middle| \{(X_j, A_j, D_j, Y_j) : 1 \leq j < i\} \right] \\
&= \frac{1}{N^2} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{E} \left[\left| \mathbb{1}_{(X_i=x, A_i=a, D_i \leq T-t)} (Y_i - \mu_1(x, a)) \right|^4 \middle| \mathcal{H}_{t-1} \right] \lesssim \frac{1}{N^2} \sum_{t=1}^T n_t = o(1).
\end{aligned}$$

7.4 Proof of Theorem 2

In the proof of Lemma 2, we notice that the term for the control units,

$$\mathbb{1}_{(X_i=x, A_i=0, D_i \leq T-t)} (Y_i - \mu_1(x, 0)),$$

and that for the treated units,

$$\mathbb{1}_{(X_i=x, A_i=1, D_i \leq T-t)} (Y_i - \mu_1(x, 1)),$$

have zero conditional covariance due to the indicators $\mathbb{1}_{(A_i=0)}$ and $\mathbb{1}_{(A_i=1)}$. As a result, $\hat{\mu}_{1,T}(x, 1)$ and $\hat{\mu}_{1,T}(x, 0)$ are jointly asymptotically normal with a zero covariance, meaning that they are asymptotically independent:

$$\begin{bmatrix} \sqrt{N}(\hat{\mu}_{1,T}(x, 0) - \mu_1(x, 0)) \\ \sqrt{N}(\hat{\mu}_{1,T}(x, 1) - \mu_1(x, 1)) \end{bmatrix} \xrightarrow{d} \mathcal{N} \left(0, \begin{bmatrix} \frac{\sigma^2(x, 0)}{p(x) \sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0)} & 0 \\ 0 & \frac{\sigma^2(x, 1)}{p(x) \sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} \end{bmatrix} \right).$$

The conclusion of the theorem then follows from recognizing $\hat{\tau}_T(x) = \hat{\mu}_{1,T}(x, 1) - \hat{\mu}_{1,T}(x, 0)$, whose asymptotic distribution is normal with a variance being the sum of the two variance terms in the previous display.

7.5 Proof of Theorem 3

We start by writing the centered and scaled estimator as

$$\sqrt{N}(\hat{\tau}_T - \tau) = \sum_{x \in \mathcal{X}} \hat{p}_T(x) \sqrt{N}(\hat{\tau}_T(x) - \tau(x)) + \sum_{x \in \mathcal{X}} \sqrt{N} \hat{p}_T(x) (\tau(x) - \tau).$$

Next, we recall that $\sum_{x \in \mathcal{X}} p(x) \tau(x) = \tau$, which implies

$$\sqrt{N}(\hat{\tau}_T - \tau) = \underbrace{\sum_{x \in \mathcal{X}} \hat{p}_T(x) \sqrt{N}(\hat{\tau}_T(x) - \tau(x))}_{\text{(I)}} + \underbrace{\sum_{x \in \mathcal{X}} \sqrt{N}(\hat{p}_T(x) - p(x))(\tau(x) - \tau)}_{\text{(II)}},$$

where we labeled the two terms above by (I) and (II), respectively. By Theorem 2 and its proof, term (I) can be written as

$$\begin{aligned} \text{(I)} &= \frac{1}{\sqrt{N}} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \sum_{x \in \mathcal{X}} \frac{\mathbf{1}_{(X_i=x, A_i=1, D_i \leq T-t)}}{\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} (Y_i - \mu_1(x, 1)) \\ &\quad - \frac{1}{\sqrt{N}} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \sum_{x \in \mathcal{X}} \frac{\mathbf{1}_{(X_i=x, A_i=0, D_i \leq T-t)}}{\sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0)} (Y_i - \mu_1(x, 0)) + o_p(1). \end{aligned}$$

Term (II) expands into

$$\text{(II)} = \frac{1}{\sqrt{N}} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \sum_{x \in \mathcal{X}} (\tau(x) - \tau) \mathbf{1}_{(X_i=x)}.$$

We notice that it is already mean zero. To summarize, we have

$$\sqrt{N}(\hat{\tau}_T - \tau) = \frac{1}{\sqrt{N}} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} (u_i - v_i + w_i) + o_p(1),$$

where

$$\begin{aligned} u_i &= \sum_{x \in \mathcal{X}} \frac{\mathbf{1}_{(X_i=x, A_i=1, D_i \leq T-t)}}{\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} (Y_i - \mu_1(x, 1)) \\ v_i &= \sum_{x \in \mathcal{X}} \frac{\mathbf{1}_{(X_i=x, A_i=0, D_i \leq T-t)}}{\sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0)} (Y_i - \mu_1(x, 0)) \\ w_i &= \sum_{x \in \mathcal{X}} (\tau(x) - \tau) \mathbf{1}_{(X_i=x)}. \end{aligned}$$

We will compute the conditional variance for each term. For the term involving u_i , one has

$$\begin{aligned} &\frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{V}[u_i | \{(X_j, A_j, D_j, Y_j) : 1 \leq j < i\}] = \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{V}[u_i | \mathcal{H}_{t-1}] \\ &= \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \sum_{x \in \mathcal{X}} \mathbb{V}\left[\frac{\mathbf{1}_{(X_i=x, A_i=1, D_i \leq T-t)}}{\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} (Y_i - \mu_1(x, 1)) \middle| \mathcal{H}_{t-1}\right] \\ &= \sum_{t=1}^T \sum_{x \in \mathcal{X}} \frac{n_t}{N} \frac{p(x) \hat{e}_t^*(1|x) \rho(T-t|x, 1)}{(\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1))^2} \sigma^2(x, 1) = \sum_{x \in \mathcal{X}} p(x) \frac{\sigma^2(x, 1)}{\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} + o_p(1), \end{aligned}$$

and for the term with v_i , one has

$$\begin{aligned}
& \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{V}[v_i | \{(X_j, A_j, D_j, Y_j) : 1 \leq j < i\}] = \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{V}[v_i | \mathcal{H}_{t-1}] \\
&= \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \sum_{x \in \mathcal{X}} \mathbb{V}\left[\frac{\mathbf{1}_{(X_i=x, A_i=0, D_i \leq T-t)}}{\sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0)} (Y_i - \mu_1(x, 0)) \middle| \mathcal{H}_{t-1}\right] \\
&= \sum_{t=1}^T \sum_{x \in \mathcal{X}} \frac{n_t}{N} \frac{p(x) \hat{e}_t^*(0|x) \rho(T-t|x, 0)}{(\sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0))^2} \sigma^2(x, 0) \\
&= \sum_{x \in \mathcal{X}} p(x) \frac{\sigma^2(x, 0)}{\sum_{t=1}^T r_t e_t(0|x) \rho(T-t|x, 0)} + o_p(1),
\end{aligned}$$

and finally,

$$\begin{aligned}
& \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{V}[w_i | \{(X_j, A_j, D_j, Y_j) : 1 \leq j < i\}] = \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \mathbb{V}[w_i | \mathcal{H}_{t-1}] \\
&= \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \sum_{x \in \mathcal{X}} \mathbb{V}[(\tau(x) - \tau) \mathbf{1}_{(X_i=x)} | \mathcal{H}_{t-1}] \\
&= \frac{1}{N} \sum_{t=1}^T \sum_{i=N_{t-1}+1}^{N_t} \sum_{x \in \mathcal{X}} (\tau(x) - \tau)^2 p(x) = \sum_{x \in \mathcal{X}} (\tau(x) - \tau)^2 p(x).
\end{aligned}$$

Next, we notice that u_i , v_i , and w_i are (conditionally) uncorrelated. Specifically,

$$\mathbb{E}[u_i v_i | \{(X_j, A_j, D_j, Y_j) : 1 \leq j < i\}] = 0$$

due to the presence of the indicator $\mathbf{1}_{(A_i=1)}$ in u_i and the indicator $\mathbf{1}_{(A_i=0)}$ in v_i . The zero correlation between u_i and w_i follows from

$$\begin{aligned}
& \mathbb{E}[u_i w_i | \{(X_j, A_j, D_j, Y_j) : 1 \leq j < i\}] = \mathbb{E}[u_i w_i | \mathcal{H}_{t-1}] \\
&= \sum_{x \in \mathcal{X}} \mathbb{E}\left[\frac{\mathbf{1}_{(X_i=x, A_i=1, D_i \leq T-t)}}{\sum_{t=1}^T r_t e_t(1|x) \rho(T-t|x, 1)} (Y_i - \mu_1(x, 1)) (\tau(x) - \tau) \middle| \mathcal{H}_{t-1}\right] = 0
\end{aligned}$$

by the law of iterated expectation, since $Y_i - \mu_1(x, 1)$ is mean zero by conditioning on $\mathcal{H}_{t-1}$ and $\mathbf{1}_{(X_i=x, A_i=1, D_i \leq T-t)}$. Similarly, $\mathbb{E}[v_i w_i | \{(X_j, A_j, D_j, Y_j) : 1 \leq j < i\}] = 0$.

To complete the proof of asymptotic normality, one needs to verify a conditional Lindeberg-Feller condition. See, for example, the proof of Lemma 2 for an illustration.

7.6 Proof of Theorem 4

By the standard law of large numbers, we have

$$\hat{p}_T(x) = p(x) + O_p\left(\sqrt{\frac{1}{N}}\right).$$

Then from the proof of Lemma 2,

$$\hat{e}_T(x, a) = \sum_{t=1}^T r_t e_t(a|x) \rho(T-t|x, a) + O_p\left(\sqrt{\frac{1}{N}}\right).$$

Finally, consistency of the estimated conditional variance, $\hat{\sigma}_T^2(x, a)$, directly follows from Lemma 1. This establishes the consistency of the variance estimator.

7.7 R code for Algorithm 1

```
#####
# The program minimizes the following objective function
#
#               c1
#  -----
#   a0 + a1*e1 + a2*e2 + ... + aJ*eJ
#
#               c0
#  +  -----
#   b0 + b1*(1-e1) + b2*(1-e2) + ... + bJ*(1-eJ)
#
# subject to the constraints
#
#   low <= e1, e_2, ..., eJ <= up
#####
# c1 and c0 correspond to conditional variances
# e1 represents the treatment allocation for the next period to be optimized
# a1, ..., aJ and b1, ..., bJ consist of the estimated delay mechanism and the
#   stage-specific enrollment frequencies
# a0 and b0 are introduced to allow either a pilot, or to capture allocations
#   in previous stages
#####
# Parameters
#   c1 and c0, as described above
#   a0 and b0: as described above
#   a and b, vectors of a1, ..., aJ and b1, ..., bJ
#   low and up: as described above
# Returned values:
```

```

# allocation: optimized allocations
# objective: minimized objective function
#####

delayedAssign <- function(c1, c0, a0 = 0, b0 = 0, a, b, low, up) {

  # number of stages
  Nstage <- length(a)
  # sort stages according to ratio a/b
  stageIndex <- sort(a/b, index.return = TRUE)$ix
  a <- a[stageIndex]; b <- b[stageIndex]

  # initialization
  allocation <- matrix(NA, nrow = Nstage, ncol = Nstage)
  for (i in 1:Nstage) for (j in 1:Nstage) {
    allocation[i, j] <- up * (j > i) + low * (j < i)
  }
  objective <- rep(0, Nstage)

  # optimization
  for (i in 1:Nstage) {
    if (i == 1 & i == Nstage) { # only one stage to optimize
      objFunc <- function(x) {
        (c1 / (a0 + a[i] * x)) + (c0 / (b0 + b[i] * (1-x)))
      }
    } else {
      objFunc <- function(x) {
        (c1 / (a0 + sum(allocation[i, -1*i] * a[-1*i]) + a[i] * x)) +
          (c0 / (b0 + sum((1-allocation[i, -1*i]) * b[-1*i]) + b[i] * (1-x)))
      }
    }

    temp <- optimize(objFunc, lower = low, upper = up, maximum = FALSE)
    objective[i] <- temp$objective; allocation[i, i] <- temp$minimum
  }
  allocation[, stageIndex] <- allocation

  # compute optimized allocation
  allocation <- allocation[which.min(objective), ]
  objective <- min(objective)
  return(list(allocation = allocation, objective = objective))
}

```

7.8 Proof of Lemma 3

The objective function $V(\alpha, \beta)$ is strictly convex, and the domain of the optimization problem $\mathcal{V}$ is convex. Therefore, the solution is unique. We also notice that V is strictly decreasing with respect to the partial ordering of $\mathbb{R}^2$, which means the minimizer must lie on the upper boundary of $\mathcal{V}$.

7.9 Proof of Theorem 5

The first claim of the theorem follows from the construction of $\mathcal{V}$. To show the second claim, we notice that the line segments connecting (α_j, β_j) all have different slopes. Therefore, the solution of the minimization problem either will be a vertex or will lie on a line segment, and in the latter scenario, the line segment is unique.

7.10 Proof of Corollary 1

This result follows directly from Lemma 3 and Theorem 5.

7.11 Proof of Theorem 6

We first note that minimizing the objective function $\hat{V}_t^{\text{forward}}$ is equivalent to minimizing

$$\hat{V}_t^{\text{forward}}(e_{t+1}, \dots, e_T | x) = \frac{\hat{\sigma}_t^2(x, 1)}{\sum_{\ell=1}^t \hat{b}'_{\ell}(x) \hat{e}_{\ell}^*(1|x) + \sum_{\ell=t+1}^T \hat{b}'_{\ell}(x) e_{\ell}} + \frac{\hat{\sigma}_t^2(x, 0)}{\sum_{\ell=1}^t \hat{b}_{\ell}(x) \hat{e}_{\ell}^*(0|x) + \sum_{\ell=t+1}^T \hat{b}_{\ell}(x) (1 - e_{\ell})}.$$

That is, the optimization can be done separately for each subgroup $x \in \mathcal{X}$. Similarly, we define

$$\begin{aligned} V_t^{\dagger}(e_{t+1}, \dots, e_T | x) &= \frac{\sigma^2(x, 1)}{\sum_{\ell=1}^t r_{\ell} \rho_t^{\dagger}(T - \ell | x, 1) e_{\ell}^{\dagger}(1|x) + \sum_{\ell=t+1}^T r_{\ell} \rho_t^{\dagger}(T - \ell | x, 1) e_{\ell}} \\ &+ \frac{\sigma^2(x, 0)}{\sum_{\ell=1}^t r_{\ell} \rho_t^{\dagger}(T - \ell | x, 0) e_{\ell}^{\dagger}(0|x) + \sum_{\ell=t+1}^T r_{\ell} \rho_t^{\dagger}(T - \ell | x, 0) (1 - e_{\ell})}, \end{aligned}$$

which is the large-sample analogue of $\hat{V}_t^{\text{forward}}(e_{t+1}, \dots, e_T | x)$. In the remaining of this proof, we will fix some $x \in \mathcal{X}$ and show the consistency of $\hat{e}_{t+1}^*(1|x)$ for $t = 1, 2, \dots, T$.

To start, the conclusion holds for $t = 1$ by design: $\hat{e}_1^*(1|x) = e_1^{\dagger}(1|x) = 1/2$. Now assume the conclusion holds for $\ell = 1, 2, \dots, t$. Then by employing the same technique used in the proof of Lemma 1 (that is, mean and conditional variance calculation), it is standard to verify that for each t ,

$$\max_{d \geq 0, x \in \mathcal{X}, a \in \mathcal{A}} \left| \hat{\rho}_t(d|x, a) - \rho_t^{\dagger}(d|x, a) \right| = o_p(1),$$

which further implies that

$$\sup_{e_{t+1}, \dots, e_T} \left| \hat{V}_t^{\text{forward}}(e_{t+1}, \dots, e_T | x) - V_t^\dagger(e_{t+1}, \dots, e_T | x) \right| = o_p(1).$$

Consistency of the treatment allocation then follows from standard M-estimation argument. See for example [van der Vaart \(2000\)](#).

7.12 Proof of Theorem 7

To show the first claim, fix some t and consider the objective function

$$V_t^\dagger(e_{t+1}(\cdot), \dots, e_T(\cdot)) = V_t^\dagger(e_{t+1}(\cdot), \dots, e_T(\cdot) ; \rho_t^\dagger(\cdot | \cdot, \cdot)).$$

Here, we have augmented the notation to emphasize that the objective function depends on the delay mechanism learned at stage t . In addition, we remark that the objective function is monotonically decreasing in the delay mechanism. At stage t , the optimization process implies that

$$\min_{e_{t+1}(\cdot), \dots, e_T(\cdot)} V_t^\dagger(e_{t+1}(\cdot), \dots, e_T(\cdot) ; \rho_t^\dagger(\cdot | \cdot, \cdot)) = \min_{e_{t+2}(\cdot), \dots, e_T(\cdot)} V_t^\dagger(e_{t+1}^\dagger(\cdot), e_{t+2}(\cdot), \dots, e_T(\cdot) ; \rho_t^\dagger(\cdot | \cdot, \cdot)).$$

Here, the meaning of the right-hand side is that we first plug in the optimal allocation for stage $t + 1$, and then optimize with respect to all future stages. To make progress, we use the fact that conservative extrapolation is employed. Under the conservative extrapolation, the learned delay mechanism never decreases; that is,

$$\rho_t^\dagger(d | x, a) \leq \rho_{t+1}^\dagger(d | x, a)$$

for all d , x , and a . As a result, we have

$$V_t^\dagger(e_{t+1}^\dagger(\cdot), e_{t+2}(\cdot), \dots, e_T(\cdot) ; \rho_t^\dagger(\cdot | \cdot, \cdot)) \geq V_t^\dagger(e_{t+1}^\dagger(\cdot), e_{t+2}(\cdot), \dots, e_T(\cdot) ; \rho_{t+1}^\dagger(\cdot | \cdot, \cdot)).$$

The right-hand side coincides with the objective function at stage $t + 1$:

$$V_{t+1}^\dagger(e_{t+2}(\cdot), \dots, e_T(\cdot)) = V_{t+1}^\dagger(e_{t+2}(\cdot), \dots, e_T(\cdot) ; \rho_{t+1}^\dagger(\cdot | \cdot, \cdot)) = V_t^\dagger(e_{t+1}^\dagger(\cdot), e_{t+2}(\cdot), \dots, e_T(\cdot) ; \rho_{t+1}^\dagger(\cdot | \cdot, \cdot)).$$

This confirms the inequalities

$$\min_{e_T(\cdot)} V_{T-1}^\dagger(e_T(\cdot)) \leq \min_{e_{T-1}(\cdot), e_T(\cdot)} V_{T-2}^\dagger(e_{T-1}(\cdot), e_T(\cdot)) \leq \min_{e_2(\cdot), \dots, e_T(\cdot)} V_1^\dagger(e_2(\cdot), \dots, e_T(\cdot)).$$

To conclude the proof of this claim, we notice that $\rho_T^\dagger(d | \cdot, \cdot) = \rho(d | \cdot, \cdot)$ for all $d \leq T - 1$. In other words, by the time the experiment concludes, all relevant information about the delay mechanism

has been acquired. As a result,

$$\begin{aligned} \min_{e_T(\cdot)} V_{T-1}^\dagger(e_T(\cdot)) &= V_{T-1}^\dagger(e_T^\dagger(\cdot)) = V_{T-1}^\dagger\left(e_T^\dagger(\cdot) ; \rho_{T-1}^\dagger(\cdot|\cdot, \cdot)\right) \\ &\geq V_{T-1}^\dagger\left(e_T^\dagger(\cdot) ; \rho_T^\dagger(\cdot|\cdot, \cdot)\right) = V_{T-1}^\dagger\left(e_T^\dagger(\cdot) ; \rho(\cdot|\cdot, \cdot)\right). \end{aligned}$$

In the above, the first equality follows from the definition that $e_T^\dagger$ is the optimized allocation; the second equality is simply a notational augmentation to emphasize the dependence on the delay mechanism; the inequality follows from the monotonicity of the objective function with respect to the learned delay mechanism; and the final equality follows from our earlier discussion. It is worth mentioning that $V_{T-1}^\dagger(e_T^\dagger(\cdot) ; \rho(\cdot|\cdot, \cdot)) = V$. Clearly,

$$V_{T-1}^\dagger\left(e_T^\dagger(\cdot) ; \rho(\cdot|\cdot, \cdot)\right) \geq \min_{e_2(\cdot), \dots, e_T(\cdot)} V_1^\dagger\left(e_2(\cdot), \dots, e_T(\cdot) ; \rho(\cdot|\cdot, \cdot)\right) \geq V^*.$$

To show the second claim, we first fix some $\varepsilon > 0$ and d_ε such that

$$\rho(d|x, a) - \rho(d_\varepsilon|x, a) \leq \varepsilon$$

for all x, a , and all $d \geq d_\varepsilon$. Then for all $t \geq d_\varepsilon$,

$$\begin{aligned} &\max_{d \geq 0, x \in \mathcal{X}, a \in \mathcal{A}} \left| \rho(d|x, a) - \rho_t^\dagger(d|x, a) \right| = \max_{d \geq 0, x \in \mathcal{X}, a \in \mathcal{A}} \left(\rho(d|x, a) - \rho_t^\dagger(d|x, a) \right) \\ &\leq \max_{d \leq t-1, x \in \mathcal{X}, a \in \mathcal{A}} \left(\rho(d|x, a) - \rho_t^\dagger(d|x, a) \right) + \max_{d \geq t, x \in \mathcal{X}, a \in \mathcal{A}} \left(\rho(d|x, a) - \rho_t^\dagger(d|x, a) \right) \\ &= 0 + \max_{d \geq t, x \in \mathcal{X}, a \in \mathcal{A}} \left(\rho(d|x, a) - \rho_t^\dagger(d|x, a) \right) \\ &\leq \max_{d \geq t, x \in \mathcal{X}, a \in \mathcal{A}} \left(\rho(d|x, a) - \rho_t^\dagger(d_\varepsilon|x, a) \right) \leq \varepsilon. \end{aligned}$$

Fix $t = d_\varepsilon$. Then since the objective function is Lipschitz continuous in the learned delay mechanism (as the denominators are bounded away from zero), one has

$$V_t^\dagger\left(e_{t+1}(\cdot), \dots, e_T(\cdot) ; \rho(\cdot|\cdot, \cdot)\right) \leq V_t^\dagger\left(e_{t+1}(\cdot), \dots, e_T(\cdot) ; \rho_t^\dagger(\cdot|\cdot, \cdot)\right) \leq V_t^\dagger\left(e_{t+1}(\cdot), \dots, e_T(\cdot) ; \rho(\cdot|\cdot, \cdot)\right) + C\varepsilon$$

for some constant C . To complete the proof, we further augment the notation by considering

$$V_t^\dagger\left(e_{t+1}(\cdot), \dots, e_T(\cdot) ; \rho(\cdot|\cdot, \cdot)\right) = V_t^\dagger\left(e_{t+1}(\cdot), \dots, e_T(\cdot) ; e_1^\dagger(\cdot), \dots, e_t^\dagger(\cdot), \rho(\cdot|\cdot, \cdot)\right),$$

to emphasize that the objective function at the end of stage t depends on past allocations, $e_1^\dagger(\cdot), \dots, e_t^\dagger(\cdot)$. Again, due to the Lipschitz continuity of the objective function, we have

$$V_t^\dagger\left(e_{t+1}(\cdot), \dots, e_T(\cdot) ; e_1^\dagger(\cdot), \dots, e_t^\dagger(\cdot), \rho(\cdot|\cdot, \cdot)\right) \leq V_t^\dagger\left(e_{t+1}(\cdot), \dots, e_T(\cdot) ; e_1^*(\cdot), \dots, e_t^*(\cdot), \rho(\cdot|\cdot, \cdot)\right) + C' \frac{d_\varepsilon}{T}$$

for some possibly different constant C' . The second claim of the theorem then follows from the

observation that

$$\min_{e_{t+1}(\cdot), \dots, e_T(\cdot)} \mathbf{V}_t^\dagger \left(e_{t+1}(\cdot), \dots, e_T(\cdot) ; e_1^*(\cdot), \dots, e_t^*(\cdot), \rho(\cdot|\cdot, \cdot) \right) = \mathbf{V}^*.$$

References

- Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Application*, New York, NY: Academic Press.
- Tsiatis, A. A. (2006). *Semiparametric Theory and Missing Data*, New York, NY: Springer.
- van der Vaart, A. W. (2000). *Asymptotic Statistics*, Cambridge, UK: Cambridge University Press.