Regret Analysis in Threshold Policy Design*

Federico Crippa
Department of Economics
Northwestern University
federicocrippa2025@u.northwestern.edu

April 17, 2024

Abstract

Threshold policies are targeting mechanisms that assign treatments based on whether an observable characteristic exceeds a certain threshold. They are widespread across multiple domains, such as welfare programs, taxation, and clinical medicine. This paper addresses the problem of designing threshold policies using experimental data, when the goal is to maximize the population welfare. First, I characterize the regret—a measure of policy optimality—of the Empirical Welfare Maximizer (EWM) policy, popular in the literature. Next, I introduce the Smoothed Welfare Maximizer (SWM) policy, which improves the EWM's regret convergence rate under an additional smoothness condition. The two policies are compared studying how differently their regrets depend on the population distribution, and investigating their finite sample performances through Monte Carlo simulations. In many contexts, the welfare guaranteed by the novel SWM policy is larger than with the EWM. An empirical illustration demonstrates how the treatment recommendation of the two policies may in practice notably differ.

Keywords: Threshold policies, heterogeneous treatment effects, statistical decision theory, randomized experiments.

JEL classification codes: C14, C44

*I am grateful to Ivan Canay, Joel Horowitz and Charles Manski for their guidance in this project. I am also thankful to Anastasiia Evdokimova, Amilcar Velez and all the participants of the Econometric Reading Group at Northwestern University for their comments and suggestions.

1 Introduction

Treatments are rarely universally assigned. When their effects are heterogeneous across individuals, policymakers aim to target those who would benefit the most from specific interventions. Scholarships, for example, are awarded to students with high academic performance or financial need; tax credits are provided to companies engaged in research and development activities; medical treatments are prescribed to sick patients. Despite the potential complexity and multidimensionality of heterogeneous treatment effects, treatment eligibility criteria are often kept quite simple. This paper studies one of the most common of these simple assignment mechanisms: threshold policies, where the decision to assign the treatment is based on whether a scalar observable characteristic — referred to as the *index* — exceeds a specified *threshold*.

Threshold policies are ubiquitous, ranging across multiple domains. In welfare policies, they regulate the qualification for public health insurance programs through age (Card et al., 2008; Shigeoka, 2014) and anti-poverty programs through income (Crost et al., 2014). In taxation, they determine marginal rates through income brackets (Taylor, 2003). In clinical medicine, the referral for liver transplantation depends on whether a composite of laboratory values obtained from blood tests is beyond a certain threshold (Kamath and Kim, 2007). Even criminal offenses are defined through threshold policies: sanctions for Driving Under the Influence are based on whether the Blood Alcohol Content exceeds specific values.

The regression discontinuity design has been developed to study the treatment effect at the point of discontinuity of threshold policies: its popularity in econometrics and applied economics is a further indication of how widespread threshold policies are. Regression discontinuity design focuses on an ex-post evaluation. In this paper, my perspective is different: I consider the ex-ante problem faced by a policymaker wanting to implement a threshold policy and interested in maximizing the average social welfare, targeting individuals who would benefit from the treatment. Experimental data are available: how should they be used to implement the threshold policy in the population?

Answering this question requires defining a criterion by which policies are evaluated. Since the performance of a policy depends on the unknown data distribution, the policy maker searches for a policy that behaves uniformly well across a specified family of data distribution (the state space). The regret of a policy is the (possibly) random difference between the maximum achievable welfare and the welfare it generates in the population. Policies can be evaluated considering their maximum expected regret (Manski, 2004; Hirano and Porter, 2009; Kitagawa and Tetenov, 2018; Athey and Wager, 2021; Mbakop and Tabord-Meehan, 2021; Manski, 2023), or other worst-case statistics of the regret distribution (Manski and Tetenov, 2023; Kitagawa et al., 2022). Once the criterion has been established, optimal policy learning aims to pinpoint the policy that minimizes it. Rather than directly tackling the functional minimization problem, following the literature, I consider candidate threshold policy functions and characterize some properties of their regret.

The first contribution of this paper is to show how to derive the asymptotic distribution of the regret for a given threshold policy. The underlying intuition is simple: threshold policies use sample data to choose the threshold, which is hence a random variable with a certain asymptotic behavior. A Taylor expansion establishes a map between the regret of a policy and its threshold, allowing one to characterize the asymptotic distribution of the regret through the asymptotic behavior of the threshold. This shifts the problem to characterizing the asymptotic of the threshold estimator, simplifying the analysis as threshold estimators can be studied with common econometric tools.

I start considering the Empirical Welfare Maximizer (EWM) policy studied by Kitagawa and Tetenov (2018). They derive uniform bounds for the expected regret of the policy for various policy classes, where the policy class impacts the findings only in terms of its VC dimensionality. My approach is more specific, considering only threshold policies, but also more informative: leveraging the knowledge of the policy class, I characterize the entire asymptotic distribution of the regret. As mentioned above, this requires the derivation of the asymptotic distribution of the threshold for the EWM policy, which is non-standard:

it exhibits the "cube root asymptotics" behavior studied in Kim and Pollard (1990). The convergence rate is $n^{1/3}$, and the asymptotic distribution is of Chernoff (1964) type. The non-standard behavior and the unusual convergence rate are due to the discontinuity in the objective function and are reflected in the asymptotic distribution of the regret, and in its $n^{2/3}$ convergence rate.

My second contribution is hence to propose a novel threshold policy, the Smoothed Welfare Maximizer (SWM) policy. This policy replaces the indicator function in the EWM policy's objective function with a smooth kernel. Under certain regularity assumptions, the threshold estimator for the SWM policy is asymptotically normal and its regret achieves a $n^{4/5}$ convergence rate. This implies that the regret's convergence rate with the SWM is faster than with the EWM policy.

Building on these asymptotic results, I extend the comparison of the regrets with the EWM and the SWM policies beyond their convergence rates. My findings allow to compare the asymptotic distributions and investigate how differently they depend on the data distribution; theoretical results are helpful to inform and guide the Monte Carlo simulations, which confirm that the asymptotic results approximate the actual finite sample behaviors. Notably, the simulations confirm that the SWM policy may guarantee lower expected regret in finite samples.

To demonstrate the practical differences between the two policies, I present an empirical illustration considering a job-training treatment. In that context, the SWM threshold policy would recommend treating 79% of unemployed workers, as opposed to 83% with the EWM policy. This difference of 4 percentage points is economically non-negligible.

1.1 Related Literature

This paper relates to the statistical decision theory literature studying the problem of policy assignment with covariates (Manski, 2004; Stoye, 2012; Kitagawa and Tetenov, 2018; Athey and Wager, 2021; Mbakop and Tabord-Meehan, 2021; Sun et al., 2021; Sun, 2021; Viviano

and Bradic, 2023). My setting is mainly related to Kitagawa and Tetenov (2018) and Athey and Wager (2021), with some notable differences. Kitagawa and Tetenov (2018) study the EWM policy for policy classes with finite VC dimension. They derive finite sample bounds for regret without relying on smoothness assumptions. Athey and Wager (2021) consider a double robust version of the EWM and allow for observational data. Under smoothness assumptions analogous to mine, they derive asymptotic bounds for the expected regret for policy classes with finite VC dimensions. Conversely, results in this paper apply exclusively to threshold policies, relying on a combination of the assumptions in Kitagawa and Tetenov (2018) and Athey and Wager (2021). The narrower focus allows for more comprehensive results: the entire asymptotic distribution of the regret is derived rather than providing some bounds for the expected regret.

Another critical distinction lies in the different nature of the uniform convergence rates: my state space is a subset of theirs, which is why the rates I derive for the EWM and the SWM policies are faster than the $\sqrt{n}$ rate reported as the optimal by Kitagawa and Tetenov (2018) and Athey and Wager (2021). Their $\sqrt{n}$ rate is, in fact, uniformly valid for a family of data distributions that, at least for threshold policies, include extreme cases (e.g. discontinuous conditional ATE), which determine the slower rate of convergence. Their uniform results may be viewed as a benchmark: when more structure is added to the problem and some distribution in the family excluded, they can be improved.

Optimal policy learning finds its empirical counterpart in the literature dealing with targeting, especially common in development economics. Recent studies rely on experimental evidence to decide who to treat in a data-driven way (Hussam et al., 2022; Aiken et al., 2022), even if Haushofer et al. (2022) pointed out the need for a more formalized approach to the targeting decision problem. The availability for the policy maker of appropriate tools to use the data in the decision process is probably necessary to guarantee a broader adoption of data-driven targeting strategies. Focusing on threshold policies, this paper explicitly formulates the decision problem, introduces implementable policies (the EWM and the SWM

policy), and compares their asymptotic properties.

Turning to the threshold estimators, I already mentioned that the EWM policy exhibits the cube root of n asymptotics studied by Kim and Pollard (1990), distinctive of several estimators in different contexts. Noteworthy examples are the maximum score estimator in choice models (Manski, 1975), the split point estimator in decision trees (Banerjee and McKeague, 2007), and the risk minimizer in classification problems (Mohammadi et al., 2005), among others. Specific to my analysis is the emergence of the cube root asymptotic within a causal inference problem relying on the potential outcomes model, which is then mirrored in the regret's asymptotic distribution.

Addressing the cube root problem by smoothing the indicator in the objective function aligns closely with the strategy proposed by Horowitz (1992) for studying the asymptotic behavior of the maximum score estimator. Objective functions are nonetheless different, and in my context, I derive the asymptotic distribution for both the unsmoothed (EWM) and the smoothed (SWM) policies. This is convenient, as it allows me to compare not only the convergence rates but also the entire asymptotic distributions of the estimators and their regrets and study the asymptotic approximations in Monte Carlo simulations.

The rest of the paper is structured as follows. Section 2 introduces the problem and outlines my analytical approach. Section 3 derives formal results for the asymptotic of the EWM and SWM policies and their regrets. In Section 4, I investigate finite sample performance of the EWM and SWM policies through Monte Carlo simulations, while in Section 5 I consider the analysis of experimental data from the National Job Training Partnership Act (JTPA) Study to compare the practical implications of the policies. Section 6 concludes.

2 Overview of the Problem

I consider the problem of a policymaker who wants to implement a binary treatment in a population of interest. Individuals are characterized by a vector of observable characteristics

$\mathbf{X} \in \mathbb{R}^d$, on which the policy maker bases the treatment assignment choice. A policy is hence a map $\pi(\mathbf{x}) : \mathbb{R}^d \to \{0, 1\}$, from observable characteristics to the binary treatment status. The policy maker is utilitarian: its goal is to maximize the average welfare in the population. Indicating by Y_1 and Y_0 the potential outcomes with and without the treatment, population average welfare generated by a policy π can be written as

$$W(\pi) = \mathbb{E}[Y_1\pi(X) + Y_0(1 - \pi(X))].$$

When treatment effects are heterogeneous, the same treatment can have opposite average effects across individuals with different $\mathbf{X}$. For this reason, the policy assignment may vary with $\mathbf{X}$: the policymaker wants to target only those who benefit from being treated, to maximize the average welfare.

The policy learning literature has considered several classes Π of policy functions, such for example linear eligibility indexes, decision trees, or monotone rules, discussed in Kitagawa and Tetenov (2018); Athey and Wager (2021); Mbakop and Tabord-Meehan (2021). This paper focuses on threshold policies, a specific class of policy functions that can be represented as

$$\pi(\mathbf{X}) = \pi(X, t) = \mathbf{1}\{X > t\}.$$

Treatment is assigned whenever the scalar observable index $X \in \mathbb{R}$ exceeds threshold t, a parameter that must be chosen. These threshold policies are widespread: they regulate, beyond others, organ transplants Kamath and Kim (2007), taxation (Taylor, 2003), and access to social welfare programs (Card et al., 2008; Crost et al., 2014). This paper aims not to justify or advocate for the use of threshold policies, and the restriction to this policy class is taken as exogenous.

I will focus on the case when the index X is chosen before the experiment. Population

welfare depends only on threshold t, and can be written as

$$W(\pi) = W(t) = \mathbb{E}[Y_1 \mathbf{1}\{X > t\} + Y_0 \mathbf{1}\{X \leq t\}].$$

Choosing the policy is equivalent to choosing the threshold. If the joint distribution of Y_1, Y_0, and X were known, the policy maker would implement the policy with threshold t^* defined as:

$$t^* \in \arg\max_t \mathbb{E}[Y_1 \mathbf{1}\{X > t\} + Y_0 \mathbf{1}\{X \leq t\}] \tag{1}$$

which would guarantee the maximum achievable welfare $W(t^*)$.

The problem described in equation (2) is unfeasible since the joint distribution of Y_1, Y_0, and X is unknown. The policy maker observes an experimental sample $Z = \{Z_i\}_{i=1}^n = \{Y_i, D_i, X_i\}$, where Y is the outcome of interest, D the randomly assigned treatment status, and X the policy index. Experimental data, which allows to identify the conditional average treatment effect, are used to learn the threshold policy $\hat{t}_n = \hat{t}_n(Z)$, function of the observed sample.

Statistical decision theory deals with the problem of choosing the map $\hat{t}_n$. First, it is necessary to specify the decision problem the policymaker faces. For any threshold policy $\hat{t}_n$, define the regret $\mathcal{R}(\hat{t}_n)$:

$$\mathcal{R}(\hat{t}_n) = W(t^*) - W(\hat{t}_n),$$

a measure of welfare loss indicating the suboptimality of policy $\hat{t}_n$. The regret depends on the unknown data distribution: the policymaker specifies a state space, and searches for a policy that does well uniformly for all the data distributions in the state space. Following Manski (2004), statistical decision theory has mainly focused on the problem of minimizing the maximum expected regret, looking for a policy $\hat{t}_n$ that does uniformly well on average

across repeated samples.

Directly solving the constrained minimization problem of the functional $\sup \mathbb{E}[\mathcal{R}(\hat{t}_n)]$ is impractical: the literature instead focuses on considering a specific policy map and studying its properties, for example showing its rate optimality, through finite sample valid (Kitagawa and Tetenov, 2018) or asymptotic (Athey and Wager, 2021) arguments. Following this approach, I characterize and compare some properties for the regret of two different threshold policies, the Empirical Welfare Maximizer (EWM) policy, commonly studied in the literature, and the novel Smoothed Welfare Maximizer (SWM) policy.

Kitagawa and Tetenov (2018) derive bounds for the expected regret of the EWM policy for a wide range of policy function classes. In their results, the policy class Π affects the bounds only through its VC dimension, and the knowledge of Π is not further exploited. Conversely, I leverage the additional structure from the knowledge of the policy class and characterize the entire asymptotic distribution of the regret for the EWM and the SWM threshold policies, comparing how their regrets depend on the data distribution. My results could hence be of interest also when decision problems not involving the expected regret are considered, as in Manski and Tetenov (2023) and Kitagawa et al. (2022): I characterize the asymptotic behavior of regret quantiles, and the asymptotic distributions can be used to simulate expectations of their non-linear functions.

To derive my results, I take advantage of the link between a threshold policy function $\hat{t}_n$ and its regret $\mathcal{R}(\hat{t}_n)$. Let $\{r_n\}$ be a sequence such that $r_n \to \infty$ for $n \to \infty$, and suppose that $r_n(\hat{t}_n - t^*)$ converges to a non degenerate limiting distribution, i.e $(\hat{t}_n - t^*) = O_p(r_n^{-1})$.

Assume function $W(t)$ to be twice continuously differentiable, and consider its second-order Taylor expansion around t^*:

$$W(\hat{t}_n) = W(t^*) + \underbrace{W'(t^*)}_{=0} \left(\hat{t}_n - t^*\right) + \frac{1}{2} W''(\tilde{t}) \left(\hat{t}_n - t^*\right)^2$$

where $|\tilde{t} - t^*| \leq |\hat{t}_n - t^*|$, and $W'(t^*) = 0$ by optimality of t^*. The previous equation can be

written as

$$r_n^2 \mathcal{R}(\hat{t}_n) = \frac{1}{2} W''(\tilde{t}) \left[r_n \left(\hat{t}_n - t^* \right) \right]^2, \tag{2}$$

establishing a relationship between the convergence rates of $\hat{t}_n$ and $\mathcal{R}(\hat{t}_n)$, and between their asymptotic distributions. Equation (2) therefore shows how the rate of convergence and the asymptotic distribution of regret $\mathcal{R}(\hat{t}_n)$ can be studied through the rate of convergence and the asymptotic distribution of policy $\hat{t}_n$. In the next section, I consider the EWM policy $\hat{t}_n^e$ and the SWM policy $\hat{t}_n^s$: through their asymptotic behaviors, I characterize the asymptotic distributions of their regrets $\mathcal{R}(\hat{t}_n^e)$ and $\mathcal{R}(\hat{t}_n^s)$.

3 Formal Results

Let Y_0 and Y_1 be scalar potential outcomes, D the binary treatment assignment in the experiment, $\mathbf{X} \in \mathbb{R}^d$ a vector of d observable characteristics, and X the observable index. $\{Y_0, Y_1, D, \mathbf{X}\}$ are random variables distributed according to the distribution P. They satisfy the following assumptions, which guarantee the identification of the optimal threshold:

Assumption 1. (Identification)

1.1 (No interference) *Observed outcome Y is related with potential outcomes by the expression $Y = DY_1 + (1-D)Y_0$.*

1.2 (Unconfoundedness) *Distribution P satisfies $D \perp\!\!\!\perp (Y_0, Y_1) | \boldsymbol{X}$.*

1.3 (Overlap) *Propensity score $p(\boldsymbol{x}) = \mathbb{E}[D|\boldsymbol{X} = \boldsymbol{x}]$ is assumed to be known and such that $p(\boldsymbol{x}) \in (\eta, 1-\eta)$, for some $\eta \in (0, 0.5)$.*

1.4 (Joint distribution) *Potential outcomes (Y_0, Y_1) and index X are continuous random variables with joint probability density function $\varphi(y_0, y_1, x)$, and marginal densities φ_0,*

φ_1, and f_x respectively. Expectations $\mathbb{E}[Y_0|x]$ and $\mathbb{E}[Y_1|x]$, for x in the support of X, exist.

Assumptions 1.1, 1.2, and 1.3 are standard assumptions in many causal models. Assumption 1.1 requires the outcome of each unit to depend only on their treatment status, excluding spillover effects. Assumption 1.2 requires random assignment of the treatment, conditionally on $\mathbf{X}$. Assumption 1.3 requires that, for any value of $\mathbf{X}$, there is a positive probability of observing both treated and untreated units. Probabilities of being assigned to the treatment may vary with $\mathbf{X}$, allowing for stratified experiments.

Assumption 1.4 specifies the focus on continuous outcome and index. While it would be possible to accommodate discrete Y_0 and Y_1, maintaining the continuity of X remains essential. The arguments developed in this paper, in fact, are not valid for a discrete index: my focus is on studying optimal threshold policies in contexts where the probability of observing any value on the support of the index X is zero, and the threshold must be chosen from a continuum of possibilities.

Under Assumption 1, optimal policy t^* defined in (2) can be written as

$$\begin{aligned} t^* &\in \arg\max_t \mathbb{E}_P[Y_1\mathbf{1}\{X > t\} + Y_0\mathbf{1}\{X \leq t\}] \\ &= \arg\max_t \mathbb{E}_P\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\mathbf{1}\{X > t\}\right] \end{aligned}$$

and is hence identified. This standard result specifies under which conditions an experiment allows to identify t^*.

3.1 Empirical Welfare Maximizer Policy

Policymaker observes an i.i.d. random sample $Z = \{Y_i, D_i, X_i\}$ of size n from P, and considers the Empirical Welfare Maximizer policy $\hat{t}_n^e$, the sample analog of t^* in equation (2)[1]:

$$\hat{t}_n^e = \arg\max_t \frac{1}{n} \sum_{i=1}^{n} \left(\frac{D_i Y_i}{p(\mathbf{X}_i)} \mathbf{1}\{X_i > t\} + \frac{(1 - D_i) Y_i}{(1 - p(\mathbf{X}_i))} \mathbf{1}\{X_i \leq t\} \right). \tag{3}$$

Policy $\hat{t}_n^e$ can be seen as an extremum estimator, maximizer of a function not continuous in t.

3.1.1 Consistency of $\hat{t}_n^e$

First, I will prove that $\hat{t}_n^e$ consistently estimates the optimal threshold t^*, implying that $\mathcal{R}(\hat{t}_n^e) \to^p 0$. To prove this result, I need the following assumptions on the data distribution.

Assumption 2. (Consistency)

2.1 (Maximizer t^*) *Maximizer $t^* \in \mathcal{T}$ of $\mathbb{E}[(Y_1 - Y_0)\mathbf{1}\{X > t\}]$ exists and is unique. It is an interior point of the compact parameter space $\mathcal{T} \subseteq \mathbb{R}$.*

2.2 (Square integrability) *Conditional expectations $\mathbb{E}[Y_0^2|X]$ and $\mathbb{E}[Y_1^2|X]$ exist.*

2.3 (Smoothness) *In a neighbourhood of t^*, density $f_x(x)$ is positive, and function $\mathbb{E}[(Y_1 - Y_0)\mathbf{1}\{X > t\}]$ is at least s-times continuously differentiable in t.*

By requiring the existence of the optimal threshold in the interior of the parameter space, Assumption 2.1 is assuming heterogeneity in the sign of the conditional average treatment effect $\mathbb{E}[Y_1 - Y_0|X]$. It is because of this heterogeneity that the policymaker implements the threshold policy, targeting groups that would benefit from being treated. The assumption neither excludes the multiplicity of local maxima, as long as the global one is unique, nor excludes unbounded support for X, but requires the parameter space to be compact. A sufficient condition for Assumption 2.1, easy to interpret and plausible in many applications,

[1] Computationally, the problem requires to evaluate the function inside the argmax $n + 1$ times; the solution is the convex set of points in $\mathbb{R}$ that give the maximum of these values.

is that the conditional average treatment effect has negative and positive values, and crosses zero exactly once.

Assumption 2.2 requires that the conditional potential outcomes have finite second moments and is satisfied when Y is assumed to be bounded (as in Kitagawa and Tetenov (2018)).

Assumption 2.3 will be used with increasing values of s to prove different results. To prove consistency, it needs to hold for $s = 0$, requiring the continuity of the objective function $W(t)$ in a neighborhood of t^*. The derivative of $\mathbb{E}[(Y_1 - Y_0)\mathbf{1}\{X > t\}]$ with respect to t is equal to $f_x(t)\tau(t)$, where $\tau(x) = \mathbb{E}[Y_1 - Y_0|X = x]$ is the conditional average treatment effect. Assumption 2.3 with $s \geq 1$ hence requires smoothness of $f_x(x)$ and $\tau(x)$, in a neighborhood of t^*.

The following theorem proves the consistency of $\hat{t}_n^e$ for t^*.

Theorem 1. *Consider the EWM policy $\hat{t}_n^e$ defined in equation* (3.1) *and the optimal policy t^* defined in equation* (2)*. Under Assumptions 1 and 2 with $s = 0$,*

$$\hat{t}_n^e \to^{a.s.} t^*$$

i.e. $\hat{t}_n^e$ is a consistent estimator for t^*.

Note that, since $\hat{t}_n^e$ is not continuous in t, the proof of Theorem 1 does not resort to the standard arguments for consistency of extremum estimators. Conversely, it relies on the approach used to prove consistency for estimators exhibiting the "cube root asymptotics" behavior, discussed in the next section.

3.1.2 Asymptotic Distribution for $\hat{t}_n^e$

The fact that $\hat{t}_n^e$ is not continuous in t directly affects the convergence rate and the asymptotic distribution. The EWM policy $\hat{t}_n^e$ exhibits the "cube root asymptotics" behavior studied in Kim and Pollard (1990), the same as, beyond others, the maximum score estimator (Manski,

1975), and the split point estimator in decision trees (Banerjee and McKeague, 2007).

The limiting distribution is not Gaussian, and its derivation requires an additional regularity condition on the tails of the probability density distribution of Y_1 and Y_0:

Assumption 3. (Tail condition) *Let φ_1 and φ_0 be the probability density functions of Y_1 and Y_0. Assume that, as $|y| \to \infty$, $\varphi_1(y) = o(|y|^{-(4+\delta)})$ and $\varphi_0(y) = o(|y|^{-(4+\delta)})$, for $\delta > 0$.*

The following theorem gives the asymptotic distribution of $\hat{t}_n^e$.

Theorem 2. *Consider the EWM policy $\hat{t}_n^e$ defined in equation* (3.1) *and the optimal policy t^* defined in equation* (2)*. Under Assumptions 1, 2 with $s = 2$ and 3, as $n \to \infty$,*

$$n^{1/3}\left(\hat{t}_n^e - t^*\right) \to^d (2\sqrt{K}/H)^{2/3} \arg\max_r \left(B(r) - r^2\right)$$

where $B(r)$ is the two-sided standard Brownian motion process, and K and H are

$$\begin{aligned} K =& f_x(t^*)\left(\frac{1}{p(\boldsymbol{X})}\mathbb{E}[Y_1^2|X = t^*] + \frac{1}{1-p(\boldsymbol{X})}\mathbb{E}[Y_0^2|X = t^*]\right) \\ H =& f_x(t^*)\left(\frac{\partial \mathbb{E}\left[Y_1 - Y_0|X = t^*\right]}{\partial X}\right). \end{aligned}$$

The limiting distribution of $n^{1/3}\left(\hat{t}_n^e - t^*\right)$ is a of Chernoff type (Chernoff, 1964). The Chernoff's distribution is the probability distribution of the random variable $\arg\max_r B(r) - r^2$, where $B(r)$ is the two-sided standard Brownian motion process. The process $B(r) - r^2$ can be simulated, and the distribution of $\arg\max_r B(r) - r^2$ numerically studied. Groeneboom and Wellner (2001) report values for selected quantiles.

It's worth noticing how the variance of $\hat{t}_n^e$ depends on the data distribution. K and H are functions of the density of X, the variance of the Conditional Average Treatment Effect, and the derivative of the CATE at t^*. The optimal threshold is estimated with more precision when more data around the optimal threshold are available (larger density), when the treatment effect changes more rapidly (larger derivative of CATE), and when the outcomes

have less variability (smaller variance of CATE). Exponents on K and H, determined by the cube root behavior, will be crucial for comparing $\hat{t}_n^e$ and $\hat{t}_n^s$.

Results in Theorem 2 can be used to derive asymptotic valid confidence intervals for $\hat{t}_n^e$, as discussed in appendix A. More interestingly, they can be combined with equation 2 to characterize the asymptotic distribution of the regret $\mathcal{R}(\hat{t}_n^e)$, as derived in the following corollary.

Corollary 2.1. *The asymptotic distribution of regret $\mathcal{R}(\hat{t}_n^e)$ is:*

$$n^{\frac{2}{3}}\mathcal{R}(\hat{t}_n^e) \to^d \left(\frac{2K^2}{H}\right)^{\frac{1}{3}} \left(\arg\max_r B(r) - r^2\right)^2.$$

The expected value of the asymptotic distribution is $K^{\frac{2}{3}}H^{-\frac{1}{3}}C^e$, where

$$C^e = \sqrt[3]{2}\mathbb{E}\left[\left(\arg\max_r B(r) - r^2\right)^2\right]$$

is a constant not dependent on P.

For the regret of the EWM policy, Corollary 2.1 establishes a $n^{\frac{2}{3}}$ rate, faster than the $\sqrt{n}$ rate found to be the optimal for the EWM expected regret (Kitagawa and Tetenov, 2018). It is essential to highlight how the two results are different: the rate derived by Kitagawa and Tetenov (2018) for the expected regret is valid uniformly over a family of distributions that may violate the assumptions in Theorem 4, for example including distributions where the CATE is discontinuous at the threshold. On the other hand, Corollary 2.1 is derived for distributions satisfying Assumptions 1, 2 and 3. They imply that K is bounded and H is bounded away from zero: the state space can be further characterized setting an upper bound $\overline{K} < \infty$ for K and a lower bound $\underline{H} > 0$ for H, implying a maximum for the expectation of the asymptotic regret equal to $\overline{K}^{\frac{2}{3}}\underline{H}^{-\frac{1}{3}}C^e$.

3.2 Smoothed Welfare Maximizer Policy

Corollary 2.1 shows how the cube root of n convergence rate of the EWM policy directly impacts the convergence rate of its regret. In this section, I propose an alternative threshold policy, the Smoothed Welfare Maximizer policy, that achieves a faster rate of convergence and hence guarantees a faster rate of convergence for its regret. My approach exploits some additional smoothness assumptions on the distribution P: Corollary 2.1 holds when $f_x(x)$ and $\tau(x)$ are assumed to be at least once differentiable; if they are at least twice differentiable, the SWM policy guarantees a $n^{\frac{4}{5}}$ convergence rate for the regret. Note that asking the density of the index and the conditional average treatment effect to be twice differentiable seems plausible for many applications. In the context of policy learning, it is, for example, assumed by Athey and Wager (2021) to derive their results[2].

My approach involves smoothing the objective function in (3.1), in the same spirit as the smoothed maximum score estimator proposed by Horowitz (1992) to deal with inference for the maximum score estimator (Manski, 1975). The Smoothed Welfare Maximizer (SWM) policy $\hat{t}_n^s$ is defined as:

$$\hat{t}_n^s = \arg\max_t \frac{1}{n} \sum_{i=1}^{n} \left[\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))} \right) k\left(\frac{X_i - t}{\sigma_n} \right) \right] \tag{4}$$

where σ_n is a sequence of positive real numbers such that $\lim_{n\to\infty} \sigma_n = 0$, and the function $k(\cdot)$ satisfies:

Assumption 4. (Kernel function) *Kernel function $k(\cdot) : \mathbb{R} \to \mathbb{R}$ is continuous, bounded, and with limits $\lim_{x\to-\infty} k(x) = 0$ and $\lim_{x\to\infty} k(x) = 1$.*

In practice, the indicator function found in $\hat{t}_n^e$ is here substituted by a smooth function $k(\cdot)$ with the same limiting behavior, which guarantees the differentiability of the expression in the argmax. The bandwidth σ_n, decreasing with the sample size, ensures that when $n \to \infty$

[2]This assumption is implied by the high-level assumptions made in the paper, as the authors discuss in footnote 15.

the policy converges to the optimal one, as proved in the next section.

3.2.1 Consistency of $\hat{t}_n^s$

I start showing consistency of $\hat{t}_n^s$ for t^*, which implies $\mathcal{R}(\hat{t}_n^s) \to^p 0$.

Theorem 3. *Consider the SWM policy $\hat{t}_n^s$ defined in equation* (3.2) *and the optimal policy t^* defined in equation* (2)*. Under Assumptions 1, 2 with $s = 0$ and 4, as $n \to \infty$,*

$$\hat{t}_n^s \to^{a.s.} t^*$$

i.e. $\hat{t}_n^s$ is a consistent estimator for t^*.

Theorems 3 and 1 are analogous: they rely on the same assumptions on the data (Assumptions 1, 2) to prove the consistency of $\hat{t}_n^e$ and $\hat{t}_n^s$. Where the two policies differ is in the asymptotic distributions: smoothness in the objective function for $\hat{t}_n^s$ guarantees asymptotic normality, but also introduces a bias, since the bandwidth σ_n equals zero only in the limit, which pops up in the limiting distribution.

3.2.2 Asymptotic Distribution for $\hat{t}_n^s$

Deriving this asymptotic behavior of $\hat{t}_n^s$ requires an additional assumption on the rate of bandwidth σ_n and the kernel function k. Since both are chosen by the policy maker, the assumption is not a restriction on unobservables but a condition on properly picking σ_n and k.

Assumption 5. (Bandwidth and kernel)

5.1 (Rate of σ_n) $\frac{\log n}{n\sigma_n^4} \to 0$ *as $n \to \infty$.*

5.2 (Kernel function) *Kernel function $k(\cdot) : \mathbb{R} \to \mathbb{R}$ satisfies Assumption 4 and the following:*

- *$k(\cdot)$ is continuous, bounded, and with limits $\lim_{x\to-\infty} k(x) = 0$ and $\lim_{x\to\infty} k(x) = 1$.*

- *$k(\cdot)$ is twice differentiable, with uniformly bounded derivatives k' and k''.*
- *$\int k'(x)^4 dx$, $\int k''(x)^2 dx$, and $\int |x^2 k''(x)| dx$ are finite.*
- *For some integer $h \geq 2$ and each integer $i \in [1, h]$, $\int |x^i k'(x)| dx = 0$ for $i < h$ and $\int |x^h k'(x)| dx = d \neq 0$, with d finite.*
- *For any integer $i \in [0, h]$, any $\eta > 0$, and any sequence $\sigma_n \to 0$, $\lim_{n\to\infty} \sigma_n^{i-h} \int_{|\sigma_n x|>\eta} |x^i k'(x)| dx = 0$, and $\lim_{n\to\infty} \sigma_n^{-1} \int_{|\sigma_n x|>\eta} |k''(x)| dx = 0$.*
- *$\int x k''(x) dx = 1$, $\lim_{n\to\infty} \int_{|\sigma_n x|>\eta} |x k''(x)| dx = 0$.*

An example of a function k satisfying Assumption 5.2 with $h = 2$ is the cumulative distribution function of the standard normal distribution.

I can now derive the asymptotic distribution of $\hat{t}_n^s$.

Theorem 4. *Consider the SWM policy $\hat{t}_n^s$ defined in equation* (3.2) *and the optimal policy t^* defined in equation* (2)*. Under Assumptions 1, 2 with $s = h + 1$ for some $h \geq 2$, and 5, as $n \to \infty$:*

1. *if $n\sigma_n^{2h+1} \to \infty$,*

$$\sigma_n^{-h}(\hat{t}_n^s - t^*) \to^p H^{-1}A;$$

2. *if $n\sigma_n^{2h+1} \to \lambda < \infty$,*

$$(n\sigma_n)^{1/2}(\hat{t}_n^s - t^*) \to^d \mathcal{N}(\lambda^{1/2}H^{-1}A, H^{-2}\alpha_2 K);$$

where A, α_1, and α_2 are:

$$A = -\frac{1}{h!}\alpha_1 \int_y (Y_1 - Y_0)\, \varphi_x^h(y, t^*) dy$$
$$\alpha_1 = \int_\zeta \zeta^h k'(\zeta)\, d\zeta$$
$$\alpha_2 = \int_\zeta k'(\zeta)^2\, d\zeta.$$

The asymptotic distribution of $(n\sigma_n)^{1/2}(\hat{t}_n^s - t^*)$ is normal, centered at the asymptotic bias $\lambda^{1/2}H^{-1}A$ introduced by the smoothing function, which exploits local information giving non-zero weights to treated units in the untreated region and vice versa. The bias and the variance of the distribution depend on the population distribution through K and H, as for the EWM policy, and also through A, a new term that determines the bias. In the definition of A, φ_x^h is the h derivative with respect to x of $\varphi(y_0, y_1, x)$, the joint density distribution of Y_0, Y_1, and X: the integral in the expression for A is the expectation of the h-derivative of $f_x(X)\tau(X)$ computed in $X = t^*$, whose existence is guaranteed by Assumption 2.3 with $s = h + 1$. α_1 and α_2 depends only on kernel function k, and are hence known.

The difference between Theorems 4 and 2 is the degree of smoothness required by Assumption 2.3. In Theorem 2, the objective function $W(t)$ is required to be twice differentiable, while Theorem 4 requires $h + 1$ continuous derivatives, where h directly impacts rate of convergence of the policy: to achieve a rate faster than the EWM policy, $\hat{t}_n^s$ requires $W(t)$ to be three times differentiable.

As for the EWM policy, results in Theorem 4 can be used to derive asymptotic valid confidence intervals for $\hat{t}_n^s$ (see appendix A), and, combined with equation 2, to characterize the asymptotic distribution of the regret $\mathcal{R}(\hat{t}_n^s)$, as derived in the next corollary.

Corollary 4.1. *Asymptotic distribution of regret $\mathcal{R}(\hat{t}_n^s)$ is:*

$$n\sigma_n \mathcal{R}(\hat{t}_n^s) \to^d \frac{1}{2}\frac{\alpha_2 K}{H}\chi^2\left(1, \frac{\lambda A^2}{\alpha_2 K}\right)$$

where $\chi^2\left(1, \frac{\lambda A^2}{\alpha_2 K}\right)$ is a non-centered chi-squared distribution with 1 degree of freedom and non-central parameter $\frac{\lambda A^2}{\alpha_2 K}$. The expected value of the asymptotic distribution is:

$$\frac{1}{2}\frac{\alpha_2 K}{H}\left(1 + \frac{\lambda A^2}{\alpha_2 K}\right) = \frac{\alpha_2}{2}\frac{K}{H} + \frac{1}{2}\frac{\lambda A^2}{H}.$$

Let $\sigma_n = (\lambda/n)^{1/(2h+1)}$ with $\lambda \in (0, \infty)$. The expectation of the asymptotic regret is minimized by setting $\lambda = \lambda^* = \frac{\alpha_2 K}{2hA^2}$: in this case, the expectation of the asymptotic distribution scaled by

$n^{\frac{2h}{2h+1}}$ *is* $A^{\frac{2}{2h+1}} K^{\frac{2h}{2h+1}} H^{-1} C^s$*, where* $C^s = \frac{2h+1}{2} \left(\frac{\alpha_2}{2h}\right)^{\frac{2h}{2h+1}}$ *is a constant not dependent on* P*.*

With the optimal bandwidth $\sigma_n = O_p(n^{-\frac{1}{2h+1}})$, the regret converges at $n^{\frac{2h}{2h+1}}$ rate. For $h \geq 2$, this implies that the regret converges faster with the SWM than with the EWM policy: the extra smoothness assumption has been exploited to achieve a better rate for the asymptotic regret. The corollary is valid for distributions satisfying Assumptions 1, 2 and 3, which imply a bounded A: the state space can be characterized setting $\overline{A} < \infty$ as the upper bound for $|A|$, implying a maximum for the expectation of the asymptotic regret equal to $\overline{A}^{\frac{2}{2h+1}} \overline{K}^{\frac{2h}{2h+1}} \underline{H}^{-1} C^s$.

3.3 Comparison of Regrets $\mathcal{R}(\hat{t}_n^e)$ and $\mathcal{R}(\hat{t}_n^s)$

Corollaries 2.1 and 4.1 showed that the regret with the SWM policy has a faster convergence rate than with the EWM policy. Assuming the accuracy of the asymptotic approximations, the comparison can extend beyond the rates to investigate how differently the asymptotic distributions depend on P, through H, K, and A. Since it is more intuitive to compare expectations than entire distributions, I will focus on this comparison, with the understanding that similar reasoning also applies to alternative statistics of the asymptotic distributions, such as medians or different quantiles.

The findings in the corollaries suggest that, for any sample size, one can select distributions P in a manner that makes each expectation of both the asymptotic regrets smaller. Hence, both the EWM and the SWM can be the better policy. If the asymptotic behavior is reflected in finite sample, this implies that in a given application where P is unknown it is impossible to determine which policy would guarantee a smaller expected regret.

In statistical decision theory, this ambiguity is recognized: focusing on the worst case scenario, the policy maker should choose the policy that guarantees the smaller regret when the expectation of the asymptotic regret is maximized. In this case, for any sample size, it means to compare $n^{-\frac{2}{3}} \overline{K}^{\frac{2}{3}} \underline{H}^{-\frac{1}{3}} C^e$ and $n^{-\frac{2h}{2h+1}} \overline{A}^{\frac{2}{2h+1}} \overline{K}^{\frac{2h}{2h+1}} \underline{H}^{-1} C^s$, and choose the policy accordingly. Which policy is uniformly better does hence depends on the application, and

specifically on values of parameters $\overline{K}$, $\underline{H}$, and $\overline{A}$, which the policy maker must set in advance.

It is interesting to investigate for which distributions P one policy is expected to do relatively better than the other, in a pointwise sense. It's clear that the SWM policy does relatively better with increasing sample size due to its faster convergence rate, for any P. Conversely, for any sample size, when $h = 2$, the ratio of the expectations of asymptotic regrets is $\frac{H^{\frac{2}{3}}}{A^{\frac{2}{5}}K^{\frac{2}{15}}}\frac{C^e}{C^s}$: the SWM policy does relatively better for larger H and smaller K and A. The negative impact of A is straightforward, since it only affects the bias for the SWM policy, without influencing the distribution of the EWM. The role of H and K is less intuitive, since they appear in both distributions but with different exponents. It's worth noting that the SWM policy holds a relative advantage when the derivative of the CATE is higher, and its variance is lower- scenarios where the benefits of selecting a threshold closer to the optimal are larger. Further comparisons of the regrets could explore the interplay between sample size and the distribution P, focusing on sequences P_n that change with sample size. Appendix B delves into this investigation.

It is important to acknowledge that the relevance of these discussions on asymptotic results relies on their ability to approximate behaviors in finite samples, as, after all, the policymaker has only access to a finite sample of experimental data. Guided by the theoretical results just derived, Monte Carlo simulations in the next section allow to analyze the finite sample regrets associated with the EWM and the SWM policies.

4 Monte Carlo Simulations

I examine the finite sample properties of the EWM and SWM policies using Monte Carlo simulations. The scope of this section is twofold: first, I will provide examples of data generating processes which lead different rankings for the two policies in terms of asymptotic regret. Then, I will verify how the asymptotic results approximate the finite sample distributions, and compare the finite sample regrets of the two policies.

As data generating process, consider the following distribution P of (Y_0, Y_1, D, X):

$$
\begin{aligned}
X &\sim \mathcal{N}(0,1) \\
\epsilon_1 &\sim \mathcal{N}(0,\gamma) \\
Y_1 &= X^3 + \beta_2 X^2 + \beta_1 X + \epsilon_1 \\
Y_0 &\sim \mathcal{N}(0,\gamma) \\
D &\sim \text{Bern}(p).
\end{aligned}
$$

Under P, the potential outcome Y_0 does not depend on the index X, and the treatment is randomly assigned with constant probability p. Parameter values are chosen such that $\mathbb{E}[Y_1|X = x]$ and hence $\mathbb{E}[Y_1 - Y_0|X = x]$ are increasing function of x, and the optimal threshold t^* is 0. It can be verified that such P implies the following:

$$
\begin{aligned}
K =& \phi(0)\left(\frac{\gamma^2}{p} + \frac{\gamma^2}{1-p}\right) \\
H =& \phi(0)\beta_1 \\
A =& -\phi(0)\beta_2 \\
W(t) =& \beta_2\left(1 - \Phi(t) + t\phi(t)\right) + \beta_1\phi(t) + \left(t^2\phi(t) + 2\phi(t)\right)
\end{aligned}
$$

where $\phi(t)$ and $\Phi(t)$ are the pdf and the cdf of the standard normal distribution, respectively.

I consider two models characterized by different parameter values:

Model	γ	β_1	β_2	p
1	1	1	-0.5	0.5
2	3	0.5	-1	0.5

Table 1 reports values for K, H, and A, and asymptotic expected regrets for both policies for sample of size $n \in \{500, 1000, 2000, 3000\}$. For the SWM policy, asymptotic regrets are computed using the infeasible optimal bandwidth σ_n^*. Compared to Model 1, Model 2 entails larger K and A, and smaller H. Results from the previous section hence imply the regret

with the EWM policy being relatively lower in Model 2. Consider the case with $n = 500$. In model 1, the asymptotic expected regret is higher with the EWM policy, whereas in model 2, it's higher with the SWM policy: this confirms that the ranking of asymptotic expected regrets depends on the unknown data distribution P. Because of the fastest rate, though, as n increases the SWM policy exhibits relatively better performance. Regardless of the specific distribution P, there exists a certain sample size beyond which the asymptotic expected regret with the SWM policy becomes smaller. In model 2, when $n = 1,000$, the inversion of ranking already occurs.

Table 1: Values of K, H, and A, and asymptotic expected regret using both EWM and SWM policies across different models, are presented. The regret for the SWM policy is computed using the optimal bandwidth σ_n^*. To facilitate the reading, asymptotic expected regrets have been scaled by a factor of 10,000.

Model	n	EWM	SWM	K	H	A
1	500	96.190	39.714	1.596	0.399	0.199
	1,000	60.596	22.809	1.596	0.399	0.199
	2,000	38.173	13.101	1.596	0.399	0.199
	3,000	29.131	9.471	1.596	0.399	0.199
2	500	400.166	439.442	9.575	0.199	0.399
	1,000	252.089	252.393	9.575	0.199	0.399
	2,000	158.806	144.962	9.575	0.199	0.399
	3,000	121.192	104.805	9.575	0.199	0.399

To investigate the finite sample distributions of regret, I draw samples of size n from P 5,000 times for each model. Each sample is used to estimate the thresholds $\hat{t}_n^e$ and $\hat{t}_n^s$. Estimating $\hat{t}_n^s$ requires specifying a bandwidth σ_n, for which I adopt the following method: I use the estimated policy $\hat{t}_n^e$ to compute $\hat{A}_n$ and $\hat{K}_n$, which are then used to compute the optimal $\hat{\lambda}_n^*$, and the optimal bandwidth $\hat{\sigma}_n^*$. The SWM policy $\hat{t}_n^s$ is hence estimated with this bandwidth $\hat{\sigma}_n^*$, which consistently estimates the optimal one if $\hat{A}_n$ and $\hat{K}_n$ consistently estimate A and K. I also consider $\hat{t}_n^s$ with the infeasible optimal σ_n^* computed from the data generating process.

Estimates for $\hat{t}_n^e$ and $\hat{t}_n^s$ are used to compute regrets $\mathcal{R}(\hat{t}_n^e)$ and $\mathcal{R}(\hat{t}_n^s)$. I thus obtain

the finite sample distributions of the regret, which can be compared with the asymptotic distributions derived in Corollaries 2.1 and 4.1. Table 2 presents the mean of these finite sample and asymptotic distributions, also depicted in Figures 1 and 2. Corresponding tables and figures for the median regret are provided in Appendix C.

The last column of each table reports the ratio between the finite sample mean regrets for the EWM and SWM policy, facilitating the comparison: a ratio larger than one indicates that the SWM policy outperforms the EWM policy. These ratios increase with the sample size, reflecting the faster asymptotic convergence rate of the SWM policy. Similar to the asymptotic results, in finite sample the SWM policy does relatively better as the sample size increases.

Table 2: Finite sample and asymptotic expected regrets with the EWM and SWM policies across different models are presented. Finite sample (empirical) values are computed through 5,000 Monte Carlo simulations. The last column displays the ratio between finite sample expected regrets with the EWM and SWM policies.

Model	n	EWM		SWM			Ratio
		empirical	asymptotic	empirical σ_n^*	empirical $\hat{\sigma}_n^*$	asymptotic	
1	500	89.635	96.190	31.492	77.747	39.714	1.153
	1,000	58.799	60.596	18.248	44.240	22.809	1.329
	2,000	37.297	38.173	10.887	29.176	13.101	1.278
	3,000	28.615	29.131	7.872	16.189	9.471	1.768
2	500	276.867	400.166	215.876	291.654	439.442	0.949
	1,000	201.182	252.089	151.181	209.069	252.393	0.962
	2,000	149.572	158.806	107.733	151.069	144.962	0.990
	3,000	124.168	121.192	90.251	125.565	104.805	0.989

Simulations enable comparison between finite sample regrets and their asymptotic counterparts. Across all models and sample sizes, the asymptotic approximation for the feasible SWM policy (with the estimated bandwidth $\hat{\sigma}_n^*$) is relatively less accurate. Simulations suggest that this is partly attributable to the need for estimating an additional tuning parameter, the bandwidth σ_n. When the SWM policy is estimated using the infeasible optimal bandwidth σ_n^*, in fact, the asymptotic approximation is more accurate and the regret is smaller.

In Model 1, as illustrated in Figure 1, both the finite sample and the asymptotic expected

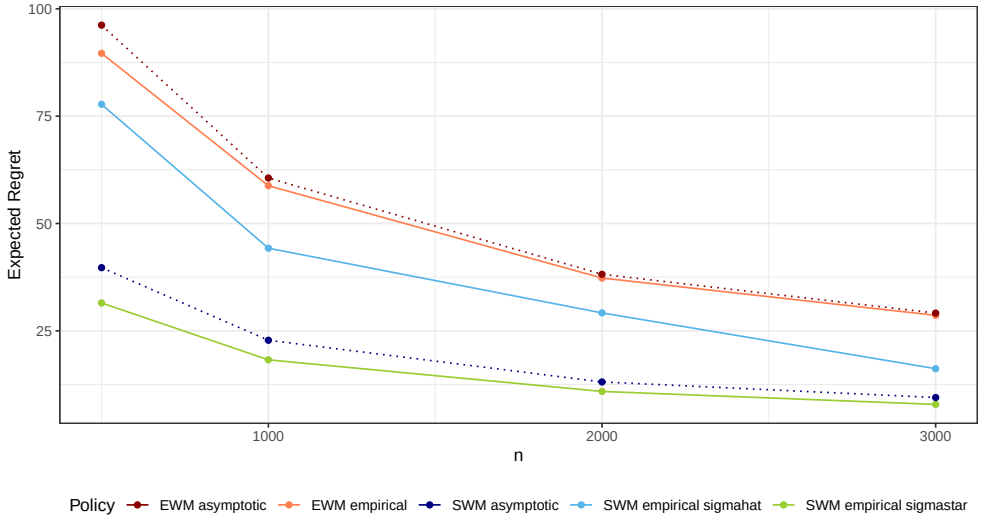

Figure 1: The figure illustrates asymptotic and finite sample expected regrets for the EWM and SWM policies in Model 1, corresponding to the values reported in Table 2.

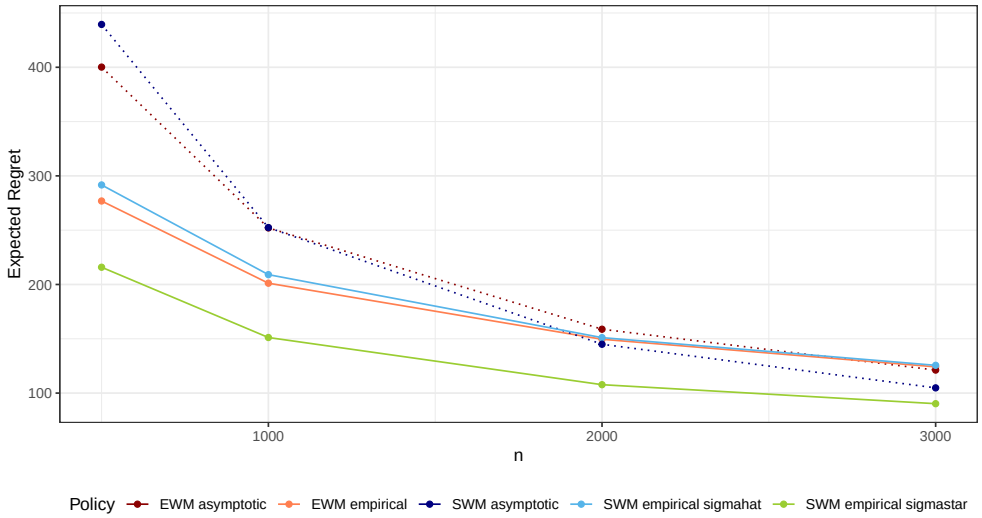

Figure 2: The figure illustrates asymptotic and finite sample expected regrets for the EWM and SWM policies in Model 2, corresponding to the values reported in Table 2.

regrets are lower for the SWM policy. Conversely, in Model 2 (illustrated in Figure 2), the finite sample expected regret is lower with the EWM policy. This confirms the impossibility of ranking the policies in a pointwise sense: different distributions P result in different rankings for the finite sample expected regrets.

It is important to note that the ranking of the EWM and the SWM policies indicated by the asymptotic results may differ from the actual finite sample comparison. Consider, for example, Model 2 with $n = 2,000$: despite the asymptotic analysis suggesting a smaller expected regret with the SWM policy, the EWM actually guarantees a smaller regret. In this scenario, even if P were known in advance, choosing according to the asymptotic approximation would not have been optimal. As n increases, the approximation improves, and the rankings based on asymptotic analysis and finite sample comparisons coincide.

Monte Carlo simulations have confirmed that the asymptotic results can approximate some finite sample behavior of the regrets, highlighting some caveats to consider when applying conclusions from asymptotic analysis to finite sample regrets with the EWM and SWM policies. However, they have not yet provided insight into the practical significance of the differences between the two policies, whether these differences are relevant or negligible in real-world scenarios. An empirical illustration is useful to answer these questions, illustrating the different implications that the policies may have.

5 Empirical Illustration

I consider the same empirical setting as Kitagawa and Tetenov (2018): experimental data from the National Job Training Partnership Act (JTPA) Study. Bloom et al. (1997) describes the experiment in detail. The study randomized whether applicants would be eligible to receive a mix of training, job-search assistance, and other services provided by the JTPA for a period of 18 months. Background information on the applicants was collected before treatment assignment, alongside administrative and survey data on the applicants' earnings

over the subsequent 30 months.

I consider the same sample of 9,223 observations as in Kitagawa and Tetenov (2018). The outcome variable Y represents the total individual earnings during the 30 months following program assignment, and the treatment variable D is a binary indicator denoting whether the individual was assigned to the program (intention-to-treat). The threshold policy is implemented by considering the individual's earnings in the year preceding the assignment as the index X. Treatment is exclusively assigned to workers with prior earnings below the threshold, based on the expectation that program services yield a more substantial positive effect for individuals who previously experienced lower earnings. Experimental data are employed to determine the threshold beyond which the treatment, on average, harms the recipients.

The bandwidth to estimate the SWM policy is chosen as in the Monte Carlo simulation, using the EWM policy $\hat{t}_n^e$ to compute $\hat{A}_n$, $\hat{K}_n$, the optimal $\hat{\lambda}_n^*$, and then the optimal bandwidth $\hat{\sigma}_n^*$. Table 3 reports the threshold estimates, including the confidence intervals constructed as discussed in Appendix A. The threshold with the SWM policy is 700 dollars lower than with the EWM (5,924 vs 6,614 dollars), a drop of 10.5%. The lower threshold implies that the treatment would target fewer workers: if the EWM policy were implemented, 82.9% of the workers in the sample would receive the program services, compared to the 78.9% with the SWM policy, resulting in a 4 percentage point decrease.

Table 3: Summary of the Empirical Welfare Maximizer (EWM) and Smoothed Welfare Maximizer (SWM) policies.

	EWM	SWM
Optimal threshold	6614	5924
Confidence interval	(4880.6, 8347.4)	(5832.7, 6411.3)
Bootstrapped confidence interval	(4748.5, 8060)	
Expected asymptotic $\mathcal{R}$	41.299	5.296
Median asymptotic $\mathcal{R}$	19.47	2.597
% of workers treated	82.912	78.911

Results in corollaries 2.1 and 4.1 allow to estimate the expected and the median regret

with the two policies: the asymptotic approximation suggests that the average regret drops from 42 to 5.5 dollars per worker, when comparing the EWM and SWM. In this context, the SWM would guarantee an average gain of 37.6 dollars per worker over the 30-month period under study, equating to 909 dollars on average for the workers who would change their treatment assignment under the two policies.

The numbers in the table should be considered with care, and clearly, the intention of this empirical illustration was not to advocate for a specific new job-training policy. Rather, the application aimed to assess if the EWM and SWM policies may have implications with relevant economic differences. Results suggest this is the case: together with theoretical and simulation findings, this implies that the choice between the EWM and the SWM policy should be thoughtfully considered, as it may determine relevant improvement in population welfare.

6 Conclusion

In this paper, I addressed the problem of using experimental data to estimate optimal threshold policies when the policymaker seeks to minimize the regret associated with implementing the policy in the population. I first examined the Empirical Welfare Maximizer threshold policy, deriving its asymptotic distribution, and showing how it links to the asymptotic distribution of its regret. I then introduced the Smoothed Welfare Maximizer policy, replacing the indicator function in the EWM policy with a smooth kernel function. Under the assumptions commonly made in the policy learning literature, the convergence rate for the worst-case regret of the SWM is faster than with the EWM policy. A comparative analysis of the asymptotic distributions of the two policies was conducted, to investigate how differently their regrets depend on the data distribution P. Monte Carlo simulations corroborated the asymptotic finding that the SWM policy may perform better than the commonly studied EWM policy also in finite sample. An empirical illustration displayed that the implications

of the two policies can remarkably differ in real-world application.

Three sets of problems remain open for future research, to extend the results of this paper in diverse directions. While this study compared the EWM policy with its smoothed counterpart SWM, the literature in statistical decision theory has also examined alternative policy functions. Consider, for example, the Augmented Inverse Propensity Weighted policy proposed by Athey and Wager (2021): what is the asymptotic distribution of the AIPW policy's regret in the context of threshold policies? How differently does it depend on P? Is it possible, analogously to what I did in this paper for the EWM policy, to modify the AIPW policy by smoothing the indicator function in its definition?

Second, it would be interesting to extend the smoothing approach from the EWM policy to other policy classes. Threshold policies are convenient as they depend on only one parameter. Still, besides more convoluted derivations, the same intuition of smoothing the indicator function also seems valid for the linear index or the multiple indices policy. The questions to explore include adapting the theory developed in this paper to these policy classes and whether this approach could be generalized even to all cases where the EWM policy is applied.

Lastly, the framework developed in this paper for using experimental data to estimate optimal policies could inform experimental design. While the existing literature mainly focuses on optimal design for estimating the average treatment effect, it could be valuable to consider scenarios when estimating the threshold policy is the goal: how should the experimental design be adapted? How the allocation of units to treatment and control groups would change? The results presented in this paper, elucidating the connection between the distribution P and the regret of the policy, provide a natural foundation for exploring experimental designs optimal for threshold policy estimation.

References

Abrevaya, J. and J. Huang (2005). On the bootstrap of the maximum score estimator. *Econometrica 73*(4), 1175–1204.

Aiken, E., S. Bellue, D. Karlan, C. Udry, and J. E. Blumenstock (2022). Machine learning and phone data can improve targeting of humanitarian aid. *Nature 603*(7903), 864–870.

Amemiya, T. (1985). *Advanced econometrics*. Harvard university press.

Athey, S. and S. Wager (2021). Policy learning with observational data. *Econometrica 89*(1), 133–161.

Banerjee, M. and I. W. McKeague (2007). Confidence sets for split points in decision trees. *The Annals of Statistics 35*(2), 543–574.

Banerjee, M. and J. A. Wellner (2001). Likelihood ratio tests for monotone functions. *Annals of Statistics*, 1699–1731.

Bloom, H. S., L. L. Orr, S. H. Bell, G. Cave, F. Doolittle, W. Lin, and J. M. Bos (1997). The benefits and costs of jtpa title ii-a programs: Key findings from the national job training partnership act study. *Journal of human resources*, 549–576.

Card, D., C. Dobkin, and N. Maestas (2008). The impact of nearly universal insurance coverage on health care utilization: evidence from medicare. *American Economic Review 98*(5), 2242–2258.

Cattaneo, M. D., M. Jansson, and K. Nagasawa (2020). Bootstrap-based inference for cube root asymptotics. *Econometrica 88*(5), 2203–2219.

Chernoff, H. (1964). Estimation of the mode. *Annals of the Institute of Statistical Mathematics 16*(1), 31–41.

Crost, B., J. Felter, and P. Johnston (2014). Aid under fire: Development projects and civil conflict. *American Economic Review 104*(6), 1833–1856.

Groeneboom, P. and J. A. Wellner (2001). Computing chernoff's distribution. *Journal of Computational and Graphical Statistics 10*(2), 388–400.

Haushofer, J., P. Niehaus, C. Paramo, E. Miguel, and M. W. Walker (2022). Targeting impact versus deprivation. Technical report, National Bureau of Economic Research.

Hirano, K. and J. R. Porter (2009). Asymptotics for statistical treatment rules. *Econometrica 77*(5), 1683–1701.

Horowitz, J. L. (1992). A smoothed maximum score estimator for the binary response model. *Econometrica: journal of the Econometric Society*, 505–531.

Hussam, R., N. Rigol, and B. N. Roth (2022). Targeting high ability entrepreneurs using community information: Mechanism design in the field. *American Economic Review 112*(3), 861–98.

Kamath, P. S. and W. R. Kim (2007). The model for end-stage liver disease (meld). *Hepatology 45*(3), 797–805.

Kim, J. and D. Pollard (1990). Cube root asymptotics. *The Annals of Statistics*, 191–219.

Kitagawa, T., S. Lee, and C. Qiu (2022). Treatment choice with nonlinear regret. *arXiv preprint arXiv:2205.08586*.

Kitagawa, T. and A. Tetenov (2018). Who should be treated? empirical welfare maximization methods for treatment choice. *Econometrica 86*(2), 591–616.

Leboeuf, J.-S., F. LeBlanc, and M. Marchand (2020). Decision trees as partitioning machines to characterize their generalization properties. *Advances in Neural Information Processing Systems 33*, 18135–18145.

Léger, C. and B. MacGibbon (2006). On the bootstrap in cube root asymptotics. *Canadian Journal of Statistics 34*(1), 29–44.

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of econometrics 3*(3), 205–228.

Manski, C. F. (2004). Statistical treatment rules for heterogeneous populations. *Econometrica 72*(4), 1221–1246.

Manski, C. F. (2021). Econometrics for decision making: Building foundations sketched by haavelmo and wald. *Econometrica 89*(6), 2827–2853.

Manski, C. F. (2023). Probabilistic prediction for binary treatment choice: with focus on personalized medicine. *Journal of Econometrics 234*(2), 647–663.

Manski, C. F. and A. Tetenov (2023). Statistical decision theory respecting stochastic dominance. *The Japanese Economic Review 74*(4), 447–469.

Mbakop, E. and M. Tabord-Meehan (2021). Model selection for treatment choice: Penalized welfare maximization. *Econometrica 89*(2), 825–848.

Mohammadi, L., S. Van De Geer, and J. Shawe-Taylor (2005). Asymptotics in empirical risk minimization. *Journal of Machine Learning Research 6*(12).

Rai, Y. (2018). Statistical inference for treatment assignment policies. *Unpublished Manuscript*.

Shigeoka, H. (2014). The effect of patient cost sharing on utilization, health, and risk protection. *American Economic Review 104*(7), 2152–2184.

Stoye, J. (2012). Minimax regret treatment choice with covariates or with limited validity of experiments. *Journal of Econometrics 166*(1), 138–156.

Sun, H., E. Munro, G. Kalashnov, S. Du, and S. Wager (2021). Treatment allocation under uncertain costs. *arXiv preprint arXiv:2103.11066*.

Sun, L. (2021). Empirical welfare maximization with constraints. *arXiv preprint arXiv:2103.15298*.

Taylor, J. (2003). Corporation income tax brackets and rates, 1909-2002. *Statistics of Income. SOI Bulletin 23*(2), 284–291.

Viviano, D. and J. Bradic (2023). Fair policy targeting. *Journal of the American Statistical Association*, 1–14.

A Confidence Intervals for Threshold Policies

Results derived in Section 3 can be used to construct confidence intervals that asymptotically cover the optimal threshold policy with a given probability, and to conduct hypotheses tests. It is important to remark that, in a decision problem setting, hypotheses testing does not have a clearly motivated justification, and indeed, statistical decision theory is the alternative approach to deal with decisions under uncertainty, as pointed out in Manski (2021). Rather than advocating for confidence intervals and hypothesis tests for threshold policies, this appendix aims to provide a procedure agnostic on why one may be interested in it.

For the EWM policy, Rai (2018) proposes some confidence intervals uniformly valid for several policy classes. They rely on test inversion of a certain bootstrap procedure, which compares the welfare generated by all the policies in the class. My procedure is much simpler for the EWM threshold policies, and I directly construct confidence intervals from the asymptotic distributions derived in Theorem 2. An analogous approach, built over results in Theorem 4, is then used to construct confidence intervals for the SWM policy.

A.1 Empirical Welfare Maximizer Policy

Consider the asymptotic distribution for the EWM threshold policy derived in Theorem 2:

$$n^{1/3}\left(\hat{t}_n^e - t^*\right) \to^d (2\sqrt{K}/H)^{2/3} \arg\max_r \left(B(r) - r^2\right).$$

If H and K were known, confidence intervals for the optimal policy t^* with asymptotic coverage $1-\alpha$ could be constructed as $(\hat{t}_n^e - w_n^e, \hat{t}_n^e + w_n^e)$, where

$$w_n^e = n^{-1/3}(2\sqrt{K}/H)^{2/3} c_{\alpha/2}$$

and $c_{\alpha/2}$ is the critical value, the upper $\alpha/2$ quantile of the distribution of $\max_r B(r) - r^2$.

In practice, H and K are unknown and should be estimated. They are defined as:

$$\begin{aligned} K =& f_x(t^*)\left(\frac{1}{p(\mathbf{X})}\mathbb{E}[Y_1^2|X=t^*] + \frac{1}{1-p(\mathbf{X})}\mathbb{E}[Y_0^2|X=t^*]\right) \\ H =& f_x(t^*)\left(\frac{\partial \mathbb{E}\left[Y_1 - Y_0|X=t^*\right]}{\partial X}\right). \end{aligned}$$

and can be estimated by a plug-in method: consider kernel density estimator $\hat{f}_x(x)$ for $f_x(x)$, and local linear estimators $\hat{\kappa}_j(x)$ and $\hat{\nu}_j'(x)$ for $\kappa_j(x) = \mathbb{E}\left[Y_j^2|X=x, D=j\right]$ and $\nu_j'(x) = \frac{\partial \nu_j(x)}{\partial x} = \frac{\partial \mathbb{E}[Y_j|X=x,D=j]}{\partial x}$. Define estimators $\hat{K}_n$ and $\hat{H}_n$ by:

$$\hat{K}_n = \hat{f}_x(\hat{t}_n^e)\left(\frac{1}{p(\mathbf{X})}\hat{\kappa}_1(\hat{t}_n^e) + \frac{1}{1-p(\mathbf{X})}\hat{\kappa}_0(\hat{t}_n^e)\right) \tag{5}$$

and

$$\hat{H}_n = \hat{f}_x(\hat{t}_n^e)(\hat{\nu}_1'(\hat{t}_n^e) - \hat{\nu}_0'(\hat{t}_n^e)). \tag{6}$$

Under the additional assumption that the second derivatives of f_x, ν_1 and ν_0 are continuous and bounded in a neighborhood of t^*, and with the proper choice of bandwidth sequences,

$\hat{K}_n$ and $\hat{H}_n$ are consistent estimators for K and H.

Feasible confidence intervals with asymptotic coverage $1-\alpha$ can hence be constructed as $(\hat{t}_n^e - \hat{w}_n^e, \hat{t}_n^e + \hat{w}_n^e)$, where

$$\hat{w}_n^e = n^{-1/3}(2\sqrt{\hat{K}_n}/\hat{H}_n)^{2/3} c_{\alpha/2}.$$

A.1.1 Bootstrap

To avoid relying on tabulated values for $c_{\alpha/2}$ and on estimation of K, an alternative approach to inference for the EWM policy is the bootstrap. Nonparametric bootstrap is not valid for $\hat{t}_n^e$ and, more generally, for "cube root asymptotics" estimators (Abrevaya and Huang, 2005; Léger and MacGibbon, 2006). Nonetheless, Cattaneo et al. (2020) provide a consistent bootstrap procedure for estimators of this type. Consistency is achieved by altering the shape of the criterion function defining the estimator whose distribution must be approximated. The standard nonparametric bootstrap is inconsistent for $Q_0(r) = -\frac{1}{2}Hr$ as defined in the proof of Theorem 2, and hence the procedure in Cattaneo et al. (2020) directly estimates this non-random part.

Let $\{Z_i^b\}$ be a random sample from the empirical distribution P_n, and define the estimator $\hat{t}_n^b$ as:

$$\begin{aligned} \hat{t}_n^b = \arg\max_t \frac{1}{n}\sum_{i=1}^n \left[\left(\frac{D_i^b Y_i^b}{p(\mathbf{X}_i^b)} - \frac{(1-D_i^b)Y_i^b}{(1-p(\mathbf{X}_i^b))}\right)\mathbf{1}\{X_i^b > t\}\right] - \\ \frac{1}{n}\sum_{i=1}^n \left[\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right)\mathbf{1}\{X_i > t\}\right] - \frac{1}{2}(t-\hat{t}_n)^2\hat{H}_n. \end{aligned} \tag{7}$$

The bootstrap procedure proposed by Cattaneo et al. (2020) is the following:

1. Compute $\hat{t}_n^e$ as described in equation (3.1).

2. Using $\hat{t}_n^e$, compute $\hat{H}_n$ as described in equation (A.1).

3. Using $\hat{t}_n^e$, $\hat{H}_n$, and the bootstrap sample $\{Z_i^b\}$, compute $\hat{t}_n^b$ as described in equation

(A.1.1).

4. Iterate step 3 to obtain the distribution of $n^{1/3}\left(\hat{t}_n^b - \hat{t}_n^e\right)$, and use it as an estimate for the distribution of $n^{1/3}\left(\hat{t}_n^e - t^*\right)$.

To be valid, the procedure needs an additional assumption.

Assumption 6. (Bounded 4th moment) *Potential outcomes distribution are such that* $\frac{1}{n^{2/3}}\mathbb{E}[Y_1^4|X = t^*] = o(1)$ *and* $\frac{1}{n^{2/3}}\mathbb{E}[Y_0^4|X = t^*] = o(1)$.

Assumption 6 guarantees that the envelope G_R is such that $PG_R^4 = o(R^{-1})$. Theorem 5 proves that the distribution of $n^{1/3}\left(\hat{t}_n^b - \hat{t}_n^e\right)$ consistently estimates the distribution of $n^{1/3}\left(\hat{t}_n^e - t^*\right)$, and validate the bootstrap procedure.

Theorem 5. *Consider estimators* $\hat{t}_n^e$ *defined in equation* (3.1) *and* $\hat{t}_n^b$ *defined in equation* (A.1.1) *and the estimand* t^* *defined in equation* (2). *Under Assumptions 1, 2 with* $s = 2$, *3, and 6, as* $\hat{H}_n \to^p H$ *and* $n \to \infty$,

$$n^{1/3}\left(\hat{t}_n^b - \hat{t}_n\right) \to^d (2\sqrt{K}/H)^{2/3} \arg\max_r \left(B(r) - r^2\right)$$

where the limiting distribution is the same as in Theorem 2.

Distribution of $n^{1/3}\left(\hat{t}_n^b - \hat{t}_n^e\right)$ can hence be used to construct asymptotic valid confidence intervals and run hypothesis tests for $\hat{t}_n^e$.

A.2 Smoothed Welfare Maximizer Policy

Consider the asymptotic distribution for the SWM threshold policy derived in Theorem 4, for $n\sigma_n^{2h+1} \to \lambda < \infty$:

$$(n\sigma_n)^{1/2}(\hat{t}_n^s - t^*) \to^d \mathcal{N}(\lambda^{1/2}H^{-1}A, H^{-2}\alpha_2 K).$$

λ, σ_n, and α_2 are known. If also K, H, and A were known, confidence intervals for the optimal policy t^* with asymptotic coverage $1-\alpha$ could be constructed as $(\hat{t}_n^s - b_n - w_n^s, \hat{t}_n^s - b_n + w_n^s)$, where

$$b_n = (n\sigma_n)^{-1/2}\lambda^{1/2}\frac{A}{H}$$
$$w_n^s = (n\sigma_n)^{-1/2}(\sqrt{\alpha_2 K}/H)c_{\alpha/2}$$

and $c_{\alpha/2}$ the upper $\alpha/2$ quantile of the standard normal distribution.

In practice, K, H, and A are unknown and should be estimated. As usual with inference involving bandwidths and kernels, two approaches are available: estimate and remove the asymptotic bias, or undersmooth.

For the first approach, consider estimators in equation (A.1) for $\hat{K}_n$ and in equation (A.1) for $\hat{H}_n$, substituting $\hat{t}_n^e$ with $\hat{t}_n^s$. For A, recall that

$$\begin{aligned} A &= -\frac{1}{h!}\alpha_1 \int_y (Y_1 - Y_0)\, \varphi_x^h(y, t^*)dy \\ &= -\frac{1}{h!}\alpha_1 \left[2f_x'(t^*)[\lambda_1'(t^*) - \lambda_0'(t^*)] + f_x(t^*)[\lambda_1''(t^*) - \lambda_0''(t^*)]\right] \end{aligned}$$

where $f_x(x)$ is the probability density function of X and $\nu_j(x) = \mathbb{E}[Y_j|X = x, D = j]$. Consider kernel density estimator $\hat{f}_x(x)$ and $\hat{f}_x'(x)$ for $f_x(x)$ and $f_x'(x)$, and local linear estimators $\hat{\nu}_j'(x)$ and $\hat{\nu}_j''(x)$ for $\nu_j'(x) = \frac{\partial \nu_j(x)}{\partial x}$ and $\nu_j''(x) = \frac{\partial^2 \nu_j(x)}{\partial^2 x}$. Define estimator $\hat{A}_n$ by:

$$\hat{A}_n = -\frac{1}{h!}\alpha_1 \left[2\hat{f}_x'(\hat{t}_n^s)[\hat{\lambda}_1'(\hat{t}_n^s) - \hat{\lambda}_0'(\hat{t}_n^s)] + \hat{f}_x(\hat{t}_n^s)[\hat{\lambda}_1''(\hat{t}_n^s) - \hat{\lambda}_0''(\hat{t}_n^s)]\right]$$

which consistently estimate A the additional assumption that the third derivatives of f_x, ν_1 and ν_0 are continuous and bounded in a neighborhood of t^*, and with the proper choice of bandwidth sequences.

Confidence intervals with asymptotic coverage $1-\alpha$ can hence be constructed as $(\hat{t}_n^s -$

$\hat{b}_n - \hat{w}_n^s, \hat{t}_n^s - \hat{b}_n + \hat{w}_n^s)$, where

$$\hat{b}_n = (n\sigma_n)^{-1/2}\lambda^{1/2}\frac{\hat{A}_n}{\hat{H}_n}$$
$$\hat{w}_n^s = (n\sigma_n)^{-1/2}(\sqrt{\alpha_2\hat{K}_n}/\hat{H}_n)c_{\alpha/2}$$

The second approach relies on undersmoothing, and chooses a suboptimally small σ_n to eliminate the asymptotic bias, with no need to estimate A. Instead of a bandwidth sequence $\sigma_n = O_p(n^{-\frac{1}{2h+1}})$, it considers a sequence $\sigma_n = o_p(n^{-\frac{1}{2h+1}})$ such that $n\sigma_n^{2h+1} \to \lambda = 0$, and ensures $b_n \to 0$. Confidence intervals with asymptotic coverage $1-\alpha$ can hence be constructed as $(\hat{t}_n^s - \hat{w}_n^s, \hat{t}_n^s + \hat{w}_n^s)$, with $\hat{w}_n^s$ defined as above.

B Local Asymptotics

The interplay between the population distribution P and the sample size can be studied considering a local asymptotic framework, considering a *sequence* of population distributions $\{P_n\}$ that varies with n. I focus on the following two sequences.

Definition 1. (Sequence $\{P_n^1\}$) *The sequence of distributions* $\{P_n^1\}$ *is such that* $\frac{\sqrt{K_n}}{H_n} = n^\gamma$, *with* $\gamma \in [0, \frac{h}{2h+1})$, *and* $\frac{A_n}{H_n} = 1$.

Definition 2. (Sequence $\{P_n^2\}$) *The sequence of distributions* $\{P_n^2\}$ *is such that* $\frac{\sqrt{K_n}}{H_n} = 1$, *and* $\frac{A_n}{H_n} = n^\gamma$, *with* $\gamma \in [0, \frac{h}{2h+1})$.

Sequence $\{P_n^1\}$ mimics a scenario where $\sqrt{K}$ is large compared to H. This occurs when the variance of the conditional treatment effect, $\mathrm{Var}(Y_1 - Y_0|X = t^*)$, is large compared to the derivative of the conditional ATE, $\frac{\partial \mathbb{E}_P[Y_1 - Y_0|X=t^*]}{\partial X}$, or compared to the density of the index, $f_x(t^*)$. In these situations, the population welfare remains relatively stable for thresholds in the neighborhood of the optimal one. The limit case with $\gamma = \frac{1}{2}$ coincides with the local asymptotic framework proposed by Hirano and Porter (2009) and studied by Athey and Wager (2021).

Sequence $\{P_n^2\}$, instead, mimics a situation where A is large compared to H. When $h = 2$, this happens when the second derivative of the conditional ATE, $\frac{\partial^2 \mathbb{E}_P[Y_1 - Y_0 | X = t^*]}{\partial^2 X}$, is large compared to the first derivative, $\frac{\partial \mathbb{E}_P[Y_1 - Y_0 | X = t^*]}{\partial X}$. Since the asymptotic bias of $\hat{t}_n^s$ is $\lambda \frac{A_n}{H_n}$, sequence $\{P_n^2\}$ represents situation where this bias is large relatively to the sample size.

Let r_n^e and r_n^s denote the rates of convergence of $\mathcal{R}(\hat{t}_n^e)$ and $\mathcal{R}(\hat{t}_n^s)$, i.e. let r_n^e and r_n^s be sequences such that $\mathcal{R}(\hat{t}_n^e) = O_p(r_n^e)$ and $\mathcal{R}(\hat{t}_n^s) = O_p(r_n^s)$. The following theorem establishes a relationship between r_n^e and r_n^s under $\{P_n^1\}$ and $\{P_n^2\}$.

Theorem 6. *Let Assumptions 1, 2 with $s = h + 1$ for some $h \geq 2$, 3, and 5 hold. Consider sequences $\{P_n^1\}$ and $\{P_n^2\}$ of data generating processes. The limit of the ratio $\frac{r_n^e}{r_n^s}$ depends on γ as follows:*

- $\frac{r_n^e}{r_n^s} \to \infty$ *if* $\gamma \in (\bar{\gamma}, \frac{h}{2h+1})$
- $\frac{r_n^e}{r_n^s} = O_p(1)$ *if* $\gamma = \bar{\gamma}$
- $\frac{r_n^e}{r_n^s} \to 0$ *if* $\gamma \in [0, \bar{\gamma})$.

where $\bar{\gamma} = \frac{h-1}{2h+1}$ *under* $\{P_n^1\}$, *and* $\bar{\gamma} = \frac{1}{3}\frac{h-1}{2h+1}$ *under* $\{P_n^2\}$.

Theorem 6 shows how, under sequences $\{P_n^1\}$ and $\{P_n^2\}$, the EWM policy guarantees a regret convergence rate faster than the SWM policy whenever the parameter γ exceeds a certain value $\bar{\gamma}$. For sequence $\{P_n^1\}$, the result depends on the fact that the term $\frac{\sqrt{K_n}}{H_n}$ enters the asymptotic distributions of $\hat{t}_n^e$ and $\hat{t}_n^s$ with different powers ($\frac{2}{3}$ and 1, respectively): if $\frac{\sqrt{K_n}}{H_n}$ is large enough relatively to the sample size, the exponent lower than one makes r_n^e faster (and hence in the limit $\hat{t}_n^e$ preferable). For sequence $\{P_n^2\}$, the result is due to the asymptotic bias of the SWM policy, proportional to $\frac{A_n}{H_n}$. Since the asymptotic distribution of the regret of the EWM policy remains constant in $\{P_n^2\}$, when the bias of $\hat{t}_n^s$ is large enough compared to sample size, $\hat{t}_n^e$ becomes preferable. The value of $\bar{\gamma}$ is increasing in h, the order of the kernel k: a smoother objective function $\mathbb{E}[(Y_1 - Y_0)\mathbf{1}\{X > t\}]$ amplifies the benefit of using the SWM policy, expanding the region of values of γ where the SWM policy has a faster rate compared to the EWM.

Considering the sequence $\{P_n^1\}$ with $\gamma = \frac{1}{2}$, Athey and Wager (2021) show that their AIPW policy achieves the uniform fastest asymptotic convergence rate. Results in Theorem 6 are different: without focusing on a single specific sequence, their goal is to shed light on how the asymptotic behavior of the EWM and the SWM policies depends on P and sample size.

C Monte Carlo Simulation

I report the analogous of Table 2 and Figures 1 and 2 for the median regret. Comments made for the mean also extend to the median regret.

Table 4: Finite sample and asymptotic median regret with the EWM and SWM policies across different models. Finite sample (empirical) values are computed through 5,000 Monte Carlo simulations. The last column reports the ratio between finite sample median regrets with the EWM and SWM policies.

Model	n	EWM		SWM			Ratio
		empirical	asymptotic	empirical σ_n^*	empirical $\hat{\sigma}_n^*$	asymptotic	
1	500	39.809	45.347	15.113	24.169	18.459	1.647
	1,000	27.547	28.567	8.692	14.145	10.602	1.948
	2,000	18.816	17.996	5.073	8.507	6.089	2.212
	3,000	14.043	13.733	3.471	5.491	4.402	2.558
2	500	147.029	188.651	116.784	159.876	204.255	0.920
	1,000	110.372	118.843	82.473	118.695	117.314	0.930
	2,000	87.325	74.866	55.731	79.436	67.379	1.099
	3,000	72.122	57.134	46.007	68.656	48.714	1.050

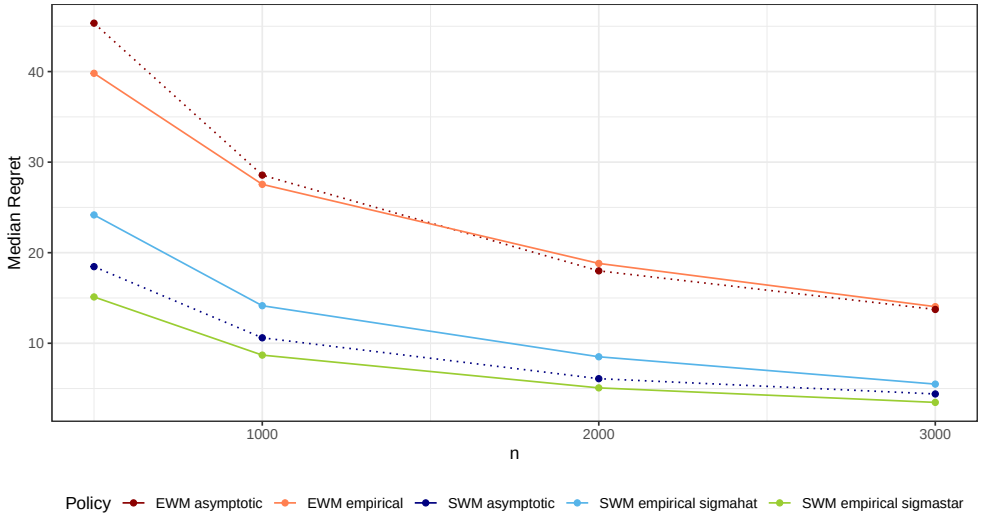

Figure 3: Figure reports asymptotic and finite sample median regrets for the EWM and SWM policies, illustrating the values reported in Table 4.

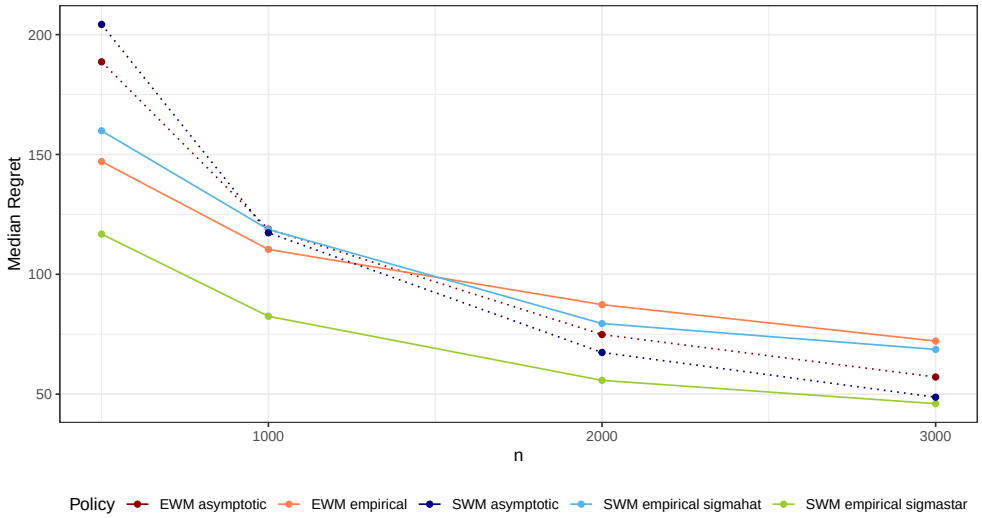

Figure 4: Figure reports asymptotic and finite sample median regrets for the EWM and SWM policies, illustrating the values reported in Table 4.

D Proofs

Theorem 1

Theorem 1. *Consider the EWM policy $\hat{t}_n^e$ defined in equation* (3.1) *and the optimal policy t^* defined in equation* (2)*. Under Assumptions 1 and 2 with $s = 0$,*

$$\hat{t}_n^e \to^{a.s.} t^*$$

i.e. $\hat{t}_n^e$ is a consistent estimator for t^*.

Proof. Estimator $\hat{t}_n^e$ in (3.1) can be written as

$$\begin{aligned}\hat{t}_n^e &= \arg\max_t \frac{1}{n}\sum_{i=1}^n \left[\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right)\left(\mathbf{1}\{X_i > t\} - \mathbf{1}\{X_i > t^*\}\right)\right] \\ &= \arg\max_t \frac{1}{n}\sum_{i=1}^n m\left(Z_i, t\right)\end{aligned}$$

where

$$m\left(Z,t\right) = \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\left(\mathbf{1}\{X > t\} - \mathbf{1}\{X > t^*\}\right)$$

and $\{Z_i\}$ is a sample of n observation from distribution P.

Define $P_n m\left(\cdot, t\right) = \sum_{i=1}^n m\left(Z_i, t\right)$ and $Pm\left(Z,t\right) = \mathbb{E}_P[m\left(Z,t\right)]$. With this notation, equation (2) and (3.1) can be written as

$$t^* = \arg\max_t Pm\left(Z,t\right)$$

$$\hat{t}_n^e = \arg\max_t P_n m\left(\cdot, t\right).$$

Threshold policies I am considering can be seen as a tree partition of depth 1. Tree partitions of finite depth are a VC class (Leboeuf et al., 2020), and hence $m\left(\cdot, t\right)$ is a

manageable class of functions. Consider the envelope function $F = 2\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right|$. Note that $\mathbb{E}\left[\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right|^2\right] = \frac{1}{p(\mathbf{X})}\mathbb{E}[Y_1^2] + \frac{1}{1-p(\mathbf{X})}\mathbb{E}[Y_0^2]$: Assumption 2.2 guarantees the existence of $\mathbb{E}\left[\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right|^2\right]$.

It follows from corollary 3.2 in Kim and Pollard (1990) that

$$\sup_t |P_n m(\cdot, t) - Pm(Z, t)| \to^{a.s.} 0.$$

Under Assumptions 2.1 and 2.3, $Pm(Z, t)$ is continuous in t, and t^* is the unique maximizer. Hence,

$$\begin{aligned}
&\sup_t |P_n m(\cdot, t) - Pm(Z, t)| + Pm(Z, t^*) \geq \\
&\left|P_n m(\cdot, \hat{t}_n^e) - Pm(Z, \hat{t}_n^e)\right| + Pm(Z, t^*) \geq \\
&\left|P_n m(\cdot, \hat{t}_n^e) - Pm(Z, \hat{t}_n^e)\right| + Pm(Z, \hat{t}_n^e) \geq \\
&P_n m(\cdot, \hat{t}_n^e) \geq P_n m(\cdot, t^*) \to Pm(Z, t^*)
\end{aligned}$$

where the second inequality is due to the fact that t^* is the maximizer of $Pm(Z, t)$, the third comes from the triangular inequality, the fourth from $\hat{t}_n^e$ being the maximizer of $P_n m(\cdot, t)$, and the last limit comes from LLN. This prove that $P_n m(\cdot, \hat{t}_n^e) \to^{a.s.} Pm(Z, t^*)$, and hence $Pm(Z, \hat{t}_n^e) \to^{a.s.} Pm(Z, t^*)$. Since t^* is the unique maximizer of $Pm(Z, t)$ and $Pm(Z, t)$ is continuous, $\hat{t}_n^e \to^{a.s.} t^*$. It means that $\hat{t}_n^e$ is a consistent estimator for t^*. □

Theorem 2

Theorem 2. *Consider the EWM policy $\hat{t}_n^e$ defined in equation* (3.1) *and the optimal policy t^* defined in equation* (2)*. Under Assumptions 1, 2 with $s = 2$ and 3, as $n \to \infty$,*

$$n^{1/3}\left(\hat{t}_n^e - t^*\right) \to^d (2\sqrt{K}/H)^{2/3} \arg\max_r \left(B(r) - r^2\right)$$

where $B(r)$ is the two-sided Brownian motion process, and K and H are

$$
\begin{aligned}
K =& f_x(t^*)\left(\frac{1}{p(\boldsymbol{X})}\mathbb{E}[Y_1^2|X=t^*]+\frac{1}{1-p(\boldsymbol{X})}\mathbb{E}[Y_0^2|X=t^*]\right)\\
H =& f_x(t^*)\left(\frac{\partial\mathbb{E}_P\left[Y_1-Y_0|X=t^*\right]}{\partial X}\right).
\end{aligned}
$$

Proof. The proof shows that conditions for the main theorem in Kim and Pollard (1990) hold; hence, their result is valid for $\hat{t}_n^e$. For completeness, I report the theorem.

Theorem Kim and Pollard. *Consider estimators defined by maximization of processes*

$$P_n g(\cdot,\theta)=\frac{1}{n}\sum_{i=1}^{n} g\left(\xi_i,\theta\right)$$

where $\{\xi_i\}_i$ is a sequence of i.i.d. observations from a distribution P and $\{g(\cdot,\theta):\theta\in\Theta\}$ is a class of functions indexed by a subset Θ in $\mathbb{R}^k$. The envelope $G_R(\cdot)$ is defined as the supremum of $g(\cdot,\theta)$ over the class

$$\mathcal{G}_R=\left\{|g(\cdot,\theta)|:|\theta-\theta_0|\leq R\right\},\quad R>0.$$

Let $\{\theta_n\}$ be a sequence of estimators for which

1. *$P_n g\left(\cdot,\theta_n\right)\geq\sup_{\theta\in\Theta}P_n g(\cdot,\theta)-o_P\left(n^{-2/3}\right)$.*
2. *θ_n converges in probability to the unique θ_0 that maximizes $Pg(\cdot,\theta)$.*
3. *θ_0 is an interior point of Θ.*

Let the functions be standardized so that $g\left(\cdot,\theta_0\right)\equiv 0$. If the classes $\mathcal{G}_R$ for R near 0 are uniformly manageable for the envelopes G_R and satisfy:

4. *$Pg(\cdot,\theta)$ is twice differentiable with second derivative matrix $-H$ at θ_0.*
5. *$K(s,r)=\lim_{\alpha\to\infty}\alpha Pg\left(\cdot,\theta_0+r/\alpha\right)g\left(\cdot,\theta_0+s/\alpha\right)$ exists for each s,r in $\mathbb{R}^k$ and $\lim_{\alpha\to\infty}\alpha Pg\left(\cdot,\theta_0+r/$ 0 for each $\varepsilon>0$ and r in $\mathbb{R}^k$.*

6. *$PG_R^2 = O(R)$ as $R \to 0$ and for each $\varepsilon > 0$ there is a constant C such that $PG_R^2 \mathbf{1}\{G_R > C\} \leq \varepsilon R$ for R near 0.*

7. *$P\left|g\left(\cdot, \theta_1\right) - g\left(\cdot, \theta_2\right)\right| = O\left(\left|\theta_1 - \theta_2\right|\right)$ near θ_0.*

Then, the process $n^{2/3} P_n g\left(\cdot, \theta_0 + r n^{-1/3}\right)$ converges in distribution to a Gaussian process $Q(r)$ with continuous sample paths, expected value $-\frac{1}{2} r' H r$ and covariance kernel K. If H is positive definite and if Q has nondegenerate increments, then $n^{1/3}\left(\theta_n - \theta_0\right)$ converges in distribution to the (almost surely unique) random vector that maximizes Q.

I apply Theorem D by taking $\xi_i = Z_i, \theta = t, \theta_n = \hat{t}_n^e, \theta_0 = t^*$, $g(\cdot, \theta) = m(\cdot, t)$, where $m(\cdot, t)$ is standardized:

$$m\left(Z_i, t\right) = \left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1 - D_i) Y_i}{(1 - p(\mathbf{X}_i))}\right)\left(\mathbf{1}\{X_i > t\} - \mathbf{1}\{X_i > t^*\}\right).$$

First, I will verify that conditions 1-7 apply to my setting:

1. $P_n m(\cdot, t)$ takes only finite $(n+1)$ values; hence, condition 1 is satisfied with the equality.

2. In Theorem 1, I proved that $\hat{t}_n^e$ is a consistent estimator for t^*.

3. t^* is an interior point of $\mathcal{T}$ by Assumption 2.1.

I need to prove that the classes $\mathcal{G}_R$ for R near 0 are uniformly manageable for the envelopes G_R. The envelope function $G_R(\cdot)$ is defined as

$$\begin{aligned} G_R(z) &= \sup\left\{m(z, t) : |t - t^*| \leq R\right\} \\ &= \sup_{|t - t^*| \leq R}\left[\left(\frac{dy}{p(\mathbf{x})} - \frac{(1 - d) y}{(1 - p(\mathbf{x}))}\right)\left(\mathbf{1}\{x > t\} - \mathbf{1}\{x > t^*\}\right)\right] \\ &= \left|\frac{dy}{p(\mathbf{x})} - \frac{(1 - d) y}{(1 - p(\mathbf{x}))}\right| \mathbf{1}\{|x - t^*| < R\} \end{aligned}$$

and I have:

$$
\begin{aligned}
PG_R^2 &= \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \mathbf{1}\{|X-t^*| < R\}\right] \\
&= \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \middle| X \in (t^*-R, t^*+R)\right] \\
&= 2R\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \middle| X = t^*\right] + o(1) = O(R)
\end{aligned} \tag{8}
$$

where the second to last equality comes from Assumption 2.2. The envelope function is uniformly square-integrable for R near 0, and therefore, the classes $\mathcal{G}_R$ are uniformly manageable.

4. Define $h(t) = Pm(Z,t)$ and consider derivatives:

$$
\begin{aligned}
h(t) =& \mathbb{E}_P\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)(\mathbf{1}\{X > t\} - \mathbf{1}\{X > t^*\})\right] = \\
& \mathbb{E}_P\left[(Y_1 - Y_0)(\mathbf{1}\{X > t\} - \mathbf{1}\{X > t^*\})\right] \\
h'(t) =& -f_x(t)\mathbb{E}_P\left[Y_1 - Y_0 \middle| X = t\right] \\
h''(t) =& -f_x'(t)\mathbb{E}_P\left[Y_1 - Y_0 \middle| X = t\right] - f_x(t)\left(\frac{\partial \mathbb{E}_P\left[Y_1 - Y_0 | X = t\right]}{\partial X}\right).
\end{aligned}
$$

Assumption 2.3 with $s = 2$ guarantees the existence of h' and h''. Since $\mathbb{E}_P\left[Y_1 - Y_0 \middle| X = t^*\right] = 0$, H is given by

$$
H = -h''(t^*) = f_x(t^*)\left(\frac{\partial \mathbb{E}_P\left[Y_1 - Y_0 | X = t^*\right]}{\partial X}\right).
$$

5. This condition is divided into two parts. First, I prove the existence of $K(s,r) =$

$\lim_{\alpha\to\infty} \alpha Pm\left(\cdot, t^* + r/\alpha\right) m\left(\cdot, t^* + s/\alpha\right)$ for each s, r in $\mathbb{R}$. Covariance K is:

$$Pm\left(\cdot, t^* + r/\alpha\right) m\left(\cdot, t^* + s/\alpha\right) = \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \right.$$
$$\left. (\mathbf{1}\{X > t^* + r/\alpha\} - \mathbf{1}\{X > t^*\})(\mathbf{1}\{X > t^* + s/\alpha\} - \mathbf{1}\{X > t^*\})\right].$$

If $rs < 0$, covariance and $K(s, r)$ are 0. If $rs > 0$, and suppose $r > 0$:

$$Pm\left(\cdot, t^* + r/\alpha\right) m\left(\cdot, t^* + s/\alpha\right) =$$
$$\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \Big| X \in (t^*, t^* + \min\{r, s\}/\alpha)\right]$$

and hence

$$K(s, r) = \lim_{\alpha\to\infty} \alpha\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \Big| X \in (t^*, t^* + \min\{r, s\}/\alpha)\right]$$
$$= \min\{r, s\} f_x(t^*)\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \Big| X = t^*\right].$$

where the equality is due to continuity of f_x (Assumption 2.3 with $s = 2$). Boundedness of the quantity follows from Assumptions 2.2 and 2.3.

Now, I will prove that $\lim_{\alpha\to\infty} \alpha Pm\left(\cdot, t^* + r/\alpha\right)^2 \mathbf{1}\left\{\left|m\left(\cdot, t^* + r/\alpha\right)\right| > \varepsilon\alpha\right\} = 0$ for each

$\varepsilon > 0$ and r in $\mathbb{R}$. I have:

$$
\begin{aligned}
&\alpha P m\left(\cdot, t^* + r/\alpha\right)^2 \mathbf{1}\left\{\left|m\left(\cdot, t^* + r/\alpha\right)\right| > \varepsilon\alpha\right\} = \\
&\alpha \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2\right. \\
&\left.(\mathbf{1}\{X > t^* + r/\alpha\} - \mathbf{1}\{X > t^*\})^2 \mathbf{1}\left\{\left|m\left(\cdot, t^* + r/\alpha\right)\right| > \varepsilon\alpha\right\}\right] = \\
&\alpha \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2\right. \\
&\left.(\mathbf{1}\{X > t^* + r/\alpha\} - \mathbf{1}\{X > t^*\})^2 \mathbf{1}\left\{\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right| > \varepsilon\alpha\right\}\right] \leq \\
&\alpha \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \mathbf{1}\left\{\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right| > \varepsilon\alpha\right\}\right].
\end{aligned}
$$

Some algebra gives

$$
\begin{aligned}
&\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \mathbf{1}\left\{\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right| > \epsilon\right\}\right] = \\
&\mathbb{E}\left[\left(\frac{DY_1^2}{p(\mathbf{X})^2} + \frac{(1-D)Y_0^2}{(1-p(\mathbf{X}))^2}\right) \mathbf{1}\left\{\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right| > \epsilon\right\}\right] = \\
&\mathbb{E}\left[\frac{Y_1^2}{p(\mathbf{X})} \mathbf{1}\left\{\left|\frac{Y_1}{p(\mathbf{X})}\right| > \epsilon\right\}\right] + \mathbb{E}\left[\frac{Y_0^2}{(1-p(\mathbf{X}))} \mathbf{1}\left\{\left|\frac{Y_0}{(1-p(\mathbf{X}))}\right| > \epsilon\right\}\right] = \\
&\frac{1}{p(\mathbf{X})} \mathbb{E}\left[Y_1^2 \mathbf{1}\left\{|Y_1| > \epsilon_1\right\}\right] + \frac{1}{(1-p(\mathbf{X}))} \mathbb{E}\left[Y_0^2 \mathbf{1}\left\{|Y_0| > \epsilon_0\right\}\right]
\end{aligned}
$$

and hence the condition is satisfied if

$$
\begin{aligned}
\lim_{\alpha \to \infty} \alpha \mathbb{E}\left[Y_1^2 \mathbf{1}\left\{|Y_1| > \epsilon\alpha\right\}\right] &= 0 \\
\lim_{\alpha \to \infty} \alpha \mathbb{E}\left[Y_0^2 \mathbf{1}\left\{|Y_0| > \epsilon\alpha\right\}\right] &= 0.
\end{aligned}
$$

Consider the limit for Y_1:

$$\lim_{\alpha\to\infty} \alpha\mathbb{E}\left[Y_1^2\mathbf{1}\left\{|Y_1| > \epsilon\alpha\right\}\right] = \lim_{\alpha\to\infty} \frac{\int_{\epsilon\alpha} Y_1^2\varphi_1(y_1)dy_1}{\alpha^{-1}} = \lim_{\alpha\to\infty} \frac{\epsilon^3\alpha^2\varphi_1(\epsilon\alpha)}{\alpha^{-2}} =$$
$$\lim_{\alpha\to\infty} \epsilon^3\alpha^4\varphi_1(\epsilon\alpha) = \lim_{\alpha\to\infty} \epsilon^3\alpha^4|\epsilon\alpha|^{-(4+\delta)}o(1) = 0$$

where the second to last equality follows from Assumption 3.

6. I showed that $PG_R^2 = O(R)$ as $R \to 0$ in equation (D). I need to prove that for each $\varepsilon > 0$ there is a constant C such that $PG_R^2\mathbf{1}\{G_R > C\} \leq \varepsilon R$ for R near 0. For any $\varepsilon > 0$ and $C > 0$:

$$PG_R^2\mathbf{1}\{G_R > C\} = \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \mathbf{1}\{|X - t^*| < R\}\mathbf{1}\{G_R > C\}\right] \leq$$
$$\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \mathbf{1}\{|X - t^*| < R\}\mathbf{1}\left\{\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right| > C\right\}\right] \leq$$

$$\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \mathbf{1}\left\{\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right| > C\right\}\right] \to 0$$

where the last limit is taken for $C \to \infty$, and follows from Assumption 3.

7. I need to show that $P\left|m\left(\cdot, t_1\right) - m\left(\cdot, t_2\right)\right| = O\left(|t_1 - t_2|\right)$ near t^*. Consider $t_2 > t_1$:

$$P\left|m\left(\cdot, t_1\right) - m\left(\cdot, t_2\right)\right| = \mathbb{E}\left[\left|\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\left(\mathbf{1}\{X > t_1\} - \mathbf{1}\{X > t_2\}\right)\right|\right] \leq$$
$$M_x\mathbb{E}\left[\left|\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\right| \Big| X \in (t_1, t_2)\right] \leq M_x M_y|t_2 - t_1|$$

where $M_x = \max_{x\in(t_1,t_2)} f_x(x)$ and $M_y = \max_{x\in(t_1,t_2)} \mathbb{E}\left[\left|\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\right| \Big| X\right]$. $M_x < \infty$ and $M_y < \infty$ because of Assumption 2.3.

Assumptions 1-7 of Theorem D by Kim and Pollard (1990) are hence satisfied. It follows

that, for $n \to \infty$,

$$n^{1/3}\left(\hat{t}_n^e - t^*\right) \to^d \arg\max_r Q(r)$$

where $Q(r) = Q_1(r) + Q_0(r)$, and Q_1 is a non degenerate zero-mean Gaussian process with covariance K, while $Q_0(r)$ is non-random and $Q_0(r) = -\frac{1}{2}r^2 H$.

Limiting distribution $\arg\max_r Q(r)$ is of Chernoff (1964) type. It can be shown (Banerjee and Wellner, 2001) that

$$\arg\max_r Q(r) =^d (2\sqrt{K}/H)^{2/3} \arg\max_r B(r) - r^2$$

where $B(r)$ is the two-sided Brownian motion process, K is:

$$\begin{aligned} K =& f_x(t^*)\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 \Big| X = t^*\right] \\ =& f_x(t^*)\left(\frac{1}{p(\mathbf{X})}\mathbb{E}[Y_1^2|X = t^*] + \frac{1}{1-p(\mathbf{X})}\mathbb{E}[Y_0^2|X = t^*]\right) \end{aligned}$$

and H is:

$$H = f_x(t^*)\left(\frac{\partial \mathbb{E}_P\left[Y_1 - Y_0 | X = t^*\right]}{\partial X}\right).$$

This completes the proof of the theorem. □

Corollary 2.1

Corollary 2.1. *The asymptotic distribution of regret $\mathcal{R}(\hat{t}_n^e)$ is:*

$$n^{\frac{2}{3}}\mathcal{R}(\hat{t}_n^e) \to^d \left(\frac{2K^2}{H}\right)^{\frac{1}{3}}\left(\arg\max_r B(r) - r^2\right)^2.$$

The expected value of the asymptotic distribution is $K^{\frac{2}{3}}H^{-\frac{1}{3}}C^e$*, where*

$$C^e = \sqrt[3]{2}\mathbb{E}\left[\left(\arg\max_r B(r) - r^2\right)^2\right]$$

is a constant not dependent on P*.*

Proof. Result in equation (2) for $\hat{t}_n^e$ implies

$$n^{\frac{2}{3}}\mathcal{R}(\hat{t}_n^e) = \frac{1}{2}W''(\tilde{t})\left(n^{\frac{1}{3}}\left(\hat{t}_n^e - t^*\right)\right)^2,$$

where $|\tilde{t} - t^*| \leq |\hat{t}_n - t^*|$. By continuous mapping theorem

$$W''(\tilde{t}) \to^p W''(t^*) = H$$

and hence by Slutsky's theorem

$$n^{\frac{2}{3}}\left(W(\hat{t}_n) - W(t^*)\right) \to^d \left(\frac{2K^2}{H}\right)^{\frac{1}{3}}\left(\arg\max_r B(r) - r^2\right)^2.$$

□

Theorem 3

Theorem 3. *Consider the SWM policy* $\hat{t}_n^s$ *defined in equation* (3.2) *and the optimal policy* t^* *defined in equation* (2)*. Under Assumptions 1, 2 with* $s = 0$ *and 4,*

$$\hat{t}_n^s \to^{a.s.} t^*$$

i.e. $\hat{t}_n^s$ *is a consistent estimator for* t^**.*

Proof. To prove the result, I show that conditions for Theorem 4.1.1 in Amemiya (1985) hold, and hence $\hat{t}_n^s$ is consistent for t^*.

First, define function $m^s(Z,t)$:

$$m^s(Z,t) = \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\left(k\left(\frac{X-t}{\sigma_n}\right) - k\left(\frac{X-t^*}{\sigma_n}\right)\right)$$

and recall definitions of $m(Z,t)$, P_n, and P introduced in the proof of Theorem 2:

$$\begin{aligned} m(Z,t) &= \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)(\mathbf{1}\{X > t\} - \mathbf{1}\{X > t^*\}) \\ P_n m(\cdot,t) &= \sum_{i=1}^{n} m(Z_i,t) \\ Pm(Z,t) &= \mathbb{E}_P[m(Z,t)]. \end{aligned}$$

In this notation, $\hat{t}_n^s = \arg\max_t P_n m^s(.,t)$ and $t^* = \arg\max_t Pm(Z,t)$. I can now show that conditions A, B, and C for Theorem 4.1.1 in Amemiya (1985) hold:

A) Parameter space $\mathcal{T}$ is compact by Assumption 2.1.

B) Function $P_n m^s(Z_i,t)$ is continuous in t for all Z and is a measurable function of Z for all $t \in \mathcal{T}$, as $k(\cdot)$ is continuous by Assumption 4.

C1) I need to prove that $P_n m^s(Z_i,t)$ converges a.s. to $Pm(Z,t)$ uniformly in $t \in \mathcal{T}$ as $n \to \infty$, i.e. $\sup_t |P_n m^s(\cdot,t) - Pm(Z,t)| \to^{a.s.} 0$. Note that:

$$\begin{aligned} &\sup_t |P_n m^s(\cdot,t) - Pm(Z,t)| \leq \\ &\sup_t |P_n m^s(\cdot,t) - Pm^s(Z,t)| + \sup_t |Pm^s(\cdot,t) - Pm(Z,t)|. \end{aligned}$$

I need to show that the two addends on the right-hand side converge to zero.

To show uniform convergence of $P_n m^s(\cdot,t)$ to $Pm^s(Z,t)$, I consider sufficient conditions provided by Theorem 4.2.1 in Amemiya (1985). $m^s(Z,t)$ is continuous in $t \in \mathcal{T}$ with $\mathcal{T}$ compact, and measurable in Z. I only need to show that $\mathbb{E}[\sup_{t\in\mathcal{T}} |m^s(Z,t)|] < \infty$. By Assumption 4, $k(\cdot)$ is a bounded function, i.e. it exists an M such that $|k(x)| < M$ for all

x. Hence $\mathbb{E}[\sup_{t\in\mathcal{T}} |m^s(Z,t)|] \leq M\mathbb{E}\left[\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right|\right]$, and $\mathbb{E}\left[\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right|\right] < \infty$ by Assumption 2.2.

To show uniform convergence of $Pm^s(\cdot,t)$ to $Pm(Z,t)$, note that $t \in \mathcal{T}$, where $\mathcal{T}$ is bounded, and hence the result holds for $\sigma_n \to 0$.

C2) By Assumption 2.1, t^* is the unique global maximum of $Pm(Z,t)$.

Assumptions A, B, C of Theorem 4.1.1 in Amemiya (1985) are satisfied, and hence $\hat{t}_n^s \to^{a.s.} t^*$. □

Lemmas

Proof of Theorem 4 requires some intermediate lemmas, stated and proved below. Arguments follows the ideas in Horowitz (1992), but are adapted to my context. I report the entire proof for completeness, even when it overlaps with the original in Horowitz (1992).

To make the notation simpler, define:

$$\hat{S}_n(t,\sigma_n) = \frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{D_iY_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right)k\left(\frac{X_i-t}{\sigma_n}\right)\right]$$

and note that $\hat{t}_n^s = \arg\max_t \hat{S}_n(t,\sigma_n)$. Then define:

$$\hat{S}_n^1(t,\sigma_n) = \frac{\partial\hat{S}_n(t,\sigma_n)}{\partial t} = -\frac{1}{\sigma_n}\frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{D_iY_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right)k'\left(\frac{X_i-t}{\sigma_n}\right)\right]$$
$$\hat{S}_n^2(t,\sigma_n) = \frac{\partial\hat{S}_n(t,\sigma_n)}{\partial^2 t} = \frac{1}{\sigma_n^2}\frac{1}{n}\sum_{i=1}^{n}\left[\left(\frac{D_iY_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right)k''\left(\frac{X_i-t}{\sigma_n}\right)\right].$$

Indicate with $\varphi_{y,x}(y,x)$ the joint distribution of Y_1, Y_0, and X, and with $\varphi_{y|x}(y|x)$ the conditional distribution, where $\varphi_{y,x}(y,x) = \varphi_{y|x}(y|x)f_x(x)$.

Lemma 1

Lemma 1. *Let Assumptions 1, 2 with $s = h + 1$ for some $h \geq 2$, and 5 hold. Then*

$$\lim_{n\to\infty} \mathbb{E}\left[\sigma_n^{-h}\hat{S}_n^1(t^*, \sigma_n)\right] = A$$
$$\lim_{n\to\infty} \mathrm{Var}\left[(n\sigma_n)^{1/2}\hat{S}_n^1(t^*, \sigma_n)\right] = \alpha_2 K.$$

Proof. First, I will prove that $\lim_{n\to\infty} \mathbb{E}\left[\sigma_n^{-h}\hat{S}_n^1(t^*, \sigma_n)\right] = A$:

$$\begin{aligned}
\mathbb{E}\left[\sigma_n^{-h}\hat{S}_n^1(t^*, \sigma_n)\right] &= -\frac{\sigma_n^{-h}}{\sigma_n}\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'\left(\frac{X_i - t^*}{\sigma_n}\right)\right] = \\
&= -\frac{\sigma_n^{-h}}{\sigma_n}\mathbb{E}\left[(Y_1 - Y_0)\,k'\left(\frac{X_i - t^*}{\sigma_n}\right)\right] = \\
&= -\sigma_n^{-h}\int_x\int_y (Y_1 - Y_0)\,\frac{1}{\sigma_n}k'\left(\frac{X_i - t^*}{\sigma_n}\right)\varphi_{y,x}(y,x)dydx = \\
&= -\sigma_n^{-h}\int_\zeta\int_y (Y_1 - Y_0)\,k'(\zeta)\,\varphi_{y,x}(y, t^* + \zeta\sigma_n)dyd\zeta
\end{aligned}$$

where in the last line I made the substitution $\zeta = \frac{X_i - t^*}{\sigma_n}$. Consider the Taylor expansion of φ around $\varphi(y, t^*)$:

$$\begin{aligned}
\varphi(y, t^* + \zeta\sigma_n) &= \varphi(y, t^*) + \zeta\sigma_n\varphi_2^1(y, t^*) + \frac{1}{2}(\zeta\sigma_n)^2\varphi_2^2(y, t^*) + \dots \\
&= \varphi(y, t^*) + \left(\sum_{i=1}^{h-1}\frac{1}{i!}\varphi_2^i(y, x = t^*)\zeta^i\sigma_n^i\right) + \frac{1}{h!}\varphi_2^h(y, \tilde{t})\zeta^h\sigma_n^h
\end{aligned}$$

with $|\tilde{t} - t^*| \leq |t^* + \zeta\sigma_n - t^*|$. Existence of $\varphi_2^m(y, x)$, the m-derivatives of $\varphi(y, x)$ with respect to its second argument, is guaranteed by Assumption 2.3 with $s = h + 1$.

Write $\mathbb{E}\left[\sigma_n^{-h}\hat{S}_n^1(t^*,\sigma_n)\right]$ as $I_{n1}+I_{n2}+I_{n3}$, where:

$$
\begin{aligned}
I_{n1} =& -\sigma_n^{-h}\int_\zeta\int_y (Y_1-Y_0)\,k'(\zeta)\,\varphi(y,t^*)dyd\zeta \\
=& -\sigma_n^{-h}\int_\zeta k'(\zeta)\,d\zeta\underbrace{\int_y (Y_1-Y_0)\,dy}_{=\mathbb{E}[Y_1-Y_0|X=t^*]=0} = 0 \\
I_{n2} =& -\sigma_n^{-h}\int_\zeta\int_y (Y_1-Y_0)\,k'(\zeta)\left(\sum_{i=1}^{h-1}\frac{1}{i!}\varphi_2^i(y,x=t^*)\zeta^i\sigma_n^i\right)dyd\zeta \\
=& \sigma_n^{-h}\int_y (Y_1-Y_0)\left(\sum_{i=1}^{h-1}\frac{1}{i!}\varphi_2^i(y,x=t^*)\sigma_n^i\underbrace{\int_\zeta k'(\zeta)\,\zeta^i d\zeta}_{=0}\right)dy = 0.
\end{aligned}
$$

Result on I_{n1} follows from definition of t^*, while Assumption 5.2 guarantees result on I_{n2}. Finally, consider I_{n3}:

$$
\begin{aligned}
I_{n3} =& -\sigma_n^{-h}\int_\zeta\int_y (Y_1-Y_0)\,k'(\zeta)\,\frac{1}{h!}\varphi_2^h(y,\tilde{t})\zeta^h\sigma_n^h dyd\zeta \\
=& -\frac{1}{h!}\int_\zeta k'(\zeta)\,\zeta^h d\zeta\int_y (Y_1-Y_0)\,\varphi_2^h(y,\tilde{t})dy
\end{aligned}
$$

and conclude that:

$$
\lim_{n\to\infty}\mathbb{E}\left[\sigma_n^{-h}\hat{S}_n^1(t^*,\sigma_n)\right] = -\frac{1}{h!}\int_\zeta \zeta^h k'(\zeta)\,d\zeta\int_y (Y_1-Y_0)\,\varphi_2^h(y,t^*)dy = A.
$$

Now I will prove that $\lim_{n\to\infty} \text{Var}\left[(n\sigma_n)^{1/2}\hat{S}_n^1(t^*,\sigma_n)\right] = \alpha_2 K$. Note that:

$$\text{Var}\left[(n\sigma_n)^{1/2}\hat{S}_n^1(t^*,\sigma_n)\right] = \text{Var}\left[(n\sigma_n)^{-1/2}\sum_{i=1}^{n}\left[\left(\frac{D_iY_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right)k'\left(\frac{X_i-t^*}{\sigma_n}\right)\right]\right] =$$

$$\begin{aligned}
&=\sigma_n \text{Var}\left[\frac{1}{\sigma_n}\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'\left(\frac{X-t^*}{\sigma_n}\right)\right] = \\
&=\sigma_n\mathbb{E}\left[\frac{1}{\sigma_n^2}\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 k'\left(\frac{X-t^*}{\sigma_n}\right)^2\right] - \\
&\quad \sigma_n\mathbb{E}\left[\frac{1}{\sigma_n}\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'\left(\frac{X-t^*}{\sigma_n}\right)\right]^2.
\end{aligned}$$

Second term in the last expression goes to 0 as $n \to \infty$. For the first term, observe that:

$$\begin{aligned}
&\sigma_n\mathbb{E}\left[\frac{1}{\sigma_n}\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 k'\left(\frac{X-t^*}{\sigma_n}\right)^2\right] = \\
&\int_x\int_y\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 k'\left(\frac{X-t^*}{\sigma_n}\right)^2 \frac{1}{\sigma_n}\varphi_{y,x}(y,x)dydx = \\
&\int_\zeta\int_y\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 k'(\zeta)^2\,\varphi_{y,x}(y,t^*+\zeta\sigma_n)dyd\zeta
\end{aligned}$$

where in the last line I made the substitution $\zeta = \frac{X_i-t^*}{\sigma_n}$. Conclude that

$$\begin{aligned}
\text{Var}\left[(n\sigma_n)^{1/2}\hat{S}_n^1(t^*,\sigma_n)\right] &= \int_\zeta k'(\zeta)^2\,d\zeta f_x(t^*)\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^2 |X=t^*\right] \\
&=\alpha_2 K.
\end{aligned}$$

Note that $\alpha_2 K$ is bounded by Assumptions 2.2, 2.3, and 5.2. □

Lemma 2

Lemma 2. *Let Assumptions 1, 2 with $s = h+1$ for some $h \geq 2$, and 5 hold. If $n\sigma_n^{2h+1} \to \infty$, $\sigma_n^{-h}\hat{S}_n^1(t^*,\sigma_n)$ converges in probability to A. If $n\sigma_n^{2h+1}$ has a finite limit λ, $(n\sigma_n)^{1/2}\hat{S}_n^1(t^*,\sigma_n)$*

converges in distribution to $\mathcal{N}(\lambda^{1/2}A, \alpha_2 K)$.

Proof. Note that

$$\mathrm{Var}\left[(\sigma_n)^{-h}\hat{S}_n^1(t^*,\sigma_n)\right] = \underbrace{(n\sigma_n^{2h+1})^{-1}}_{\to 0}\underbrace{\mathrm{Var}\left[(n\sigma_n)^{1/2}\hat{S}_n^1(t^*,\sigma_n)\right]}_{\to \alpha_2 K}.$$

So the first result follows from lemma 1 and Chebyshev's inequality.

For the second result, first note that under the stated assumptions and from lemma 1,

$$\mathbb{E}\left[(n\sigma_n)^{1/2}\hat{S}_n^1(t^*,\sigma_n)\right] = \underbrace{(n\sigma_n^{2h+1})^{1/2}}_{\to \lambda^{1/2}}\underbrace{\mathbb{E}\left[\sigma_n^{-h}\hat{S}_n^1(t^*,\sigma_n)\right]}_{A}$$

and so the result follows if I show that

$$U_n = (n\sigma_n)^{1/2}\left(\hat{S}_n^1(t^*,\sigma_n) - \mathbb{E}\left[\hat{S}_n^1(t^*,\sigma_n)\right]\right) \to^d \mathcal{N}(0,\alpha_2 K).$$

Note that

$$\begin{aligned} U_n =& (n\sigma_n)^{1/2}\frac{1}{n}\sum_{i=1}^n\left[\underbrace{\left(\frac{D_iY_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right)k'\left(\frac{X_i-t}{\sigma_n}\right)\frac{1}{\sigma_n}}_{=B} - \mathbb{E}[B]\right] = \\ =& \sum_{i=1}^n\left(\frac{\sigma_n}{n}\right)^{1/2}(B-\mathbb{E}[B]) \end{aligned}$$

and hence U_n has characteristic function $\psi(\tau)^n$, where

$$\psi(\tau) = \mathbb{E}\left[\exp\left(i\tau\left(\frac{\sigma_n}{n}\right)^{1/2}(B-\mathbb{E}[B])\right)\right]$$

and

$$\psi^{'}(\tau) = \mathbb{E}\left[i\left(\frac{\sigma_n}{n}\right)^{1/2}(B - \mathbb{E}[B])\exp\left(i\tau\left(\frac{\sigma_n}{n}\right)^{1/2}(B - \mathbb{E}[B])\right)\right]$$
$$\psi^{''}(\tau) = \mathbb{E}\left[-\frac{\sigma_n}{n}(B - \mathbb{E}[B])\exp\left(i\tau\left(\frac{\sigma_n}{n}\right)^{1/2}(B - \mathbb{E}[B])\right)\right].$$

Note that $\psi^{'}(0) = 0$ and $\psi^{''}(0) = -\frac{\sigma_n}{n}\operatorname{Var}(B) = -\frac{1}{n}(\alpha_2 K + o(1))$, since lemma 1 proved that $\lim_{n\to\infty}\sigma_n \operatorname{Var}(B) = \alpha_2 K$.

A Taylor series expansion of $\psi(\tau)$ about $\tau = 0$ yields:

$$\psi(\tau) = \underbrace{\psi(0)}_{=1} + \underbrace{\psi^{'}(0)}_{=0}\tau + \frac{1}{2}\underbrace{\psi^{''}(0)}_{=-\frac{\alpha_2 K}{n}}\tau^2 + o\left(\frac{\tau^2}{n}\right) = 1 - \frac{1}{2n}\alpha_2 K\tau^2 + o\left(\frac{\tau^2}{n}\right)$$

and hence the characteristic function of U_n has limit:

$$\lim_{n\to\infty}\left[1 - \frac{1}{2n}\alpha_2 K\tau^2 + o\left(\frac{\tau^2}{n}\right)\right]^n = \exp\left(-\alpha_2 K\frac{\tau^2}{2}\right).$$

Since $\exp\left(-\alpha_2 K\frac{\tau^2}{2}\right)$ is the characteristic function of $\mathcal{N}(0, \alpha_2 K)$, the second result of the lemma holds. □

Lemma 3

Lemma 3. *Let Assumptions 1, 2 with $s = h+1$ for some $h \geq 2$, and 5 hold. Let $\eta > 0$ be such that $\varphi_{y,x}(y, x)$ has second derivative uniformly bounded for almost every X if $|X - t^*| < \eta$. For $\theta \in \mathbb{R}$, define $\hat{S}_n^{\theta}(\theta)$ by*

$$\hat{S}_n^{\theta}(\theta) = -(n\sigma_n^2)^{-1}\sum_{i=1}^{n}\left[\left(\frac{D_i Y_i}{p(\boldsymbol{X}_i)} - \frac{(1 - D_i)Y_i}{(1 - p(\boldsymbol{X}_i))}\right)k'\left(\frac{X_i - t^*}{\sigma_n} + \theta\right)\right].$$

Define the sets $\Theta_n (n = 1, 2, \dots)$ *by* $\Theta_n = \{\theta : \theta \in \mathbb{R}, \sigma_n |\theta| \leq \frac{\eta}{2}\}$. *Then*

$$\text{plim}_{n\to\infty} \sup_{\theta\in\Theta_n} |\hat{S}_n^\theta(\theta) - \mathbb{E}[\hat{S}_n^\theta(\theta)]| = 0.$$

In addition, there are finite numbers α_1 *and* α_2 *such that for all* $\theta \in \Theta_n$

$$|\mathbb{E}[\hat{S}_n^\theta(\theta)] - H\theta| \leq o(1) + \alpha_1 \sigma_n |\theta| + \alpha_2 \sigma_n \theta^2$$

uniformly over $\theta \in \Theta_n$.

Proof. To prove the first result, first define

$$\begin{aligned} -g_i(\theta) = &\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right) k'\left(\frac{X_i - t^*}{\sigma_n} + \theta\right) - \\ &\mathbb{E}\left[\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right) k'\left(\frac{X_i - t^*}{\sigma_n} + \theta\right)\right]. \end{aligned}$$

It is necessary to prove that for any $\varepsilon > 0$

$$\lim_{n\to\infty} \Pr\left[\sup_{\theta\in\Theta_n} \left(n\sigma_n^2\right)^{-1} \left|\sum_{n=1}^{N} g_i(\theta)\right| > \varepsilon\right] = 0.$$

Given any $\delta > 0$, divide each set Θ_n into nonoverlapping subsets $\Theta_{nj} (j = 1, 2, \dots)$ such that the distance between any two points in the same subset does not exceed $\delta\sigma_n^2$ and the number Γ_n of subsets does not exceed $C\sigma_n^{-3(q-1)}$ for some $C > 0$. Let $\{\theta_{Ni}\}$ be a set of vectors such

that $\theta_{nj} \in \Theta_{nj}$. Then

$$\begin{aligned}
\Pr\left[\sup_{\theta\in\Theta_n}\left(n\sigma_n^2\right)^{-1}\left|\sum_{n=1}^{n} g_i(\theta)\right| > \varepsilon\right] &= \\
&= \Pr\left[\bigcup_{j=1}^{\Gamma_n}\sup_{\theta\in\Theta_{nj}}\left(n\sigma_n^2\right)^{-1}\left|\sum_{i=1}^{n} g_i(\theta)\right| > \varepsilon\right] \\
&\leqslant \sum_{j=1}^{\Gamma_n}\Pr\left[\sup_{\theta\in\Theta_{nj}}\left(n\sigma_n^2\right)^{-1}\left|\sum_{i=1}^{n} g_i(\theta)\right| > \varepsilon\right] \\
&\leqslant \underbrace{\sum_{j=1}^{\Gamma_n}\Pr\left[\left(n\sigma_n^2\right)^{-1}\left|\sum_{i=1}^{n} g_i\left(\theta_{nj}\right)\right| > \varepsilon/2\right]}_{=B_1} \\
&\quad + \underbrace{\sum_{j=1}^{\Gamma_n}\Pr\left[\left(n\sigma_n^2\right)^{-1}\sum_{i=1}^{n}\sup_{\theta\in\Theta_{nj}}\left|g_i(\theta) - g_i\left(\theta_{nj}\right)\right| > \varepsilon/2\right]}_{=B_2},
\end{aligned}$$

where the last two lines follow from the triangle inequality. By Hoeffding's inequality, there are finite numbers $c_1 > 0$ and $c_2 > 0$ such that

$$\Pr\left[\left(n\sigma_n^2\right)^{-1}\left|\sum_{i=1}^{n} g_i\left(\theta_{nj}\right)\right| > \varepsilon/2\right] \leqslant c_1 \exp\left(-c_2 n\sigma_n^4\right).$$

Therefore, B_1 is bounded by $Cc_1\sigma_n^{-3(q-1)}\exp\left(-c_2 n\sigma_n^4\right)$, which converges to 0 as $n \to \infty$ by Assumption 5.1. In addition, by Assumption 4 there is a finite $c_3 > 0$ such that if $\theta \in \Theta_{nj}$,

$$\begin{aligned}
&\left|-\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right)\left[k'\left(\frac{X_i - t^*}{\sigma_n} + \theta\right) - k'\left(\frac{X_i - t^*}{\sigma_n} + \theta_{nj}\right)\right]\right| \\
&\qquad \leqslant c_3\left|\theta - \theta_{nj}\right| \leqslant c_3\delta\sigma_n^2.
\end{aligned}$$

So

$$\left(n\sigma_n^2\right)^{-1}\sum_{i=1}^{n}\sup_{\theta\in\Theta_{nj}}\left|g_i(\theta) - g_i\left(\theta_{nj}\right)\right| \leqslant 2c_3\delta.$$

Choose $\delta < \varepsilon/4c_3$. Then B_2 is 0. This establishes $\operatorname{plim}_{n\to\infty}\sup_{\theta\in\Theta_n}\left|\hat{S}_n^\theta(\theta) - \mathbb{E}[\hat{S}_n^\theta(\theta)]\right| = 0$.

To prove the second result, start noting that

$$\mathbb{E}[\hat{S}_n^{\theta}(\theta)] = -\sigma_n^{-2}\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right) k'\left(\frac{X-t^*}{\sigma_n} + \theta\right)\right] = I_{n1} + I_{n2}$$

where

$$\begin{aligned} I_{n1} &= -\sigma_n^{-2}\int_{|X-t^*|\leq\eta}\int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right) k'\left(\frac{X-t^*}{\sigma_n} + \theta\right)\varphi(y,x)dydx \\ I_{n2} &= -\sigma_n^{-2}\int_{|X-t^*|>\eta}\int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right) k'\left(\frac{X-t^*}{\sigma_n} + \theta\right)\varphi(y,x)dydx. \end{aligned}$$

First, consider I_{n2} and observe that

$$I_{n2} = -\sigma_n^{-2}\int_{|X-t^*|>\eta} k'\left(\frac{X-t^*}{\sigma_n} + \theta\right) \underbrace{\int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\varphi(y|x)dy}_{=\mathbb{E}[Y_1-Y_0|X]} f_x(x)dx$$

and since $\mathbb{E}[Y_1 - Y_0|X]$ is bounded by Assumption 2.2,

$$|I_{n2}| \leq \left|C\sigma_n^{-2}\int_{|X-t^*|>\eta} k'\left(\frac{X-t^*}{\sigma_n} + \theta\right) f_x(x)dx\right|.$$

Define $\zeta = \frac{X-t^*}{\sigma_n} + \theta$. Since $\sigma_n|\theta| \leq \frac{\eta}{2}$, when $|X - t^*| > \eta$

$$\begin{aligned} |\zeta| &= \left|\frac{X-t^*}{\sigma_n} + \theta\right| = \pm\left(\frac{X-t^*}{\sigma_n} + \theta\right) \geq \pm\left(\frac{X-t^*}{\sigma_n}\right) - |\theta| = \left|\frac{X-t^*}{\sigma_n}\right| - |\theta| \\ &\geq \left|\frac{X-t^*}{\sigma_n}\right| - \frac{\eta}{2\sigma_n} > \frac{\eta}{\sigma_n} - \frac{\eta}{2\sigma_n} = \frac{\eta}{2\sigma_n}. \end{aligned}$$

and so the event $|X - t^*| > \eta$ implies $|\zeta| > \frac{\eta}{2\sigma_n}$. Then

$$
\begin{aligned}
|I_{n2}| &\leq \left| C\sigma_n^{-2} \int_{|X-t^*|>\eta} k'\left(\frac{X-t^*}{\sigma_n} + \theta\right) f_x(x)dx \right| \\
&= \left| C\sigma_n^{-1} \int_{|X-t^*|>\eta} k'(\zeta)\, f_x(t^* - \theta\sigma_n)d\zeta \right| \\
&\leq \left| C \underbrace{f_x(t^* - \theta\sigma_n)}_{\to f_x(t^*)} \underbrace{\sigma_n^{-1} \int_{|\zeta|>\eta/\sigma_n} k'(\zeta)\, d\zeta}_{\to 0} \right|.
\end{aligned}
$$

The fact that $f_x(t^* - \theta\sigma_n) \to f_x(t^*)$ bounded by Assumption 2.3 with $s = h + 1$ and $\sigma_n^{-1} \int_{|\zeta|>\eta/\sigma_n} k'(\zeta)\, d\zeta \to 0$ by Assumption 5.2 implies

$$
\text{plim}_{n\to\infty} \sup_{\theta\in\Theta_n} |I_{n2}| = 0.
$$

Recall that I_{n1} is defined as

$$
I_{n1} = -\sigma_n^{-2} \int_{|X-t^*|\leq\eta} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) k'\left(\frac{X-t^*}{\sigma_n} + \theta\right) \varphi(y,x)dydx.
$$

Consider a Taylor expansion of $\varphi(y,x)$ about $x = t^*$:

$$
\varphi(y,x) = \varphi(y,t^*) + \varphi'(y,t^*)(x-t^*) + \frac{1}{2}\varphi''(y,\tilde{t})(x-t^*)^2
$$

with $|\tilde{t}-t^*| \leq |x-t^*|$. Write I_{n1} as $J_{n1}+J_{n2}+J_{n3}$ where

$$
\begin{aligned}
J_{n1} =& -\sigma_n^{-2}\int_{|X-t^*|\leq\eta}\int_y \left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'\left(\frac{X-t^*}{\sigma_n}+\theta\right)\varphi(y,t^*)dydx \\
=& -\sigma_n^{-2}\underbrace{\mathbb{E}[Y_1-Y_0|X=t^*]}_{=0}\int_{|X-t^*|\leq\eta} k'\left(\frac{X-t^*}{\sigma_n}+\theta\right)f_x(t^*)dx=0 \\
J_{n2} =& -\sigma_n^{-2}\int_{|X-t^*|\leq\eta}\int_y \left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'\left(\frac{X-t^*}{\sigma_n}+\theta\right)\varphi'(y,t^*)(x-t^*)dydx \\
J_{n3} =& -\sigma_n^{-2}\int_{|X-t^*|\leq\eta}\int_y \left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'\left(\frac{X-t^*}{\sigma_n}+\theta\right)\frac{1}{2}\varphi''(y,\tilde{t})(x-t^*)^2dydx.
\end{aligned}
$$

Consider J_{n2} and the substitution $\zeta=\frac{X-t^*}{\sigma_n}+\theta$:

$$
\begin{aligned}
J_{n2} =& -\int_{|\zeta-\theta|\leq\eta/\sigma_n}\int_y \left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'(\zeta)(\zeta-\theta)\varphi'(y,t^*)dyd\zeta \\
=& -\int_{|\zeta-\theta|\leq\eta/\sigma_n}\int_y \left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'(\zeta)\zeta\varphi'(y,t^*)dyd\zeta+ \\
& \int_{|\zeta-\theta|\leq\eta/\sigma_n}\int_y \left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)k'(\zeta)\theta\varphi'(y,t^*)dyd\zeta \\
=& -\int_{|\zeta-\theta|\leq\eta/\sigma_n}\zeta k'(\zeta)d\zeta\underbrace{\int_y \left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\varphi'(y,t^*)dy}_{=H}+ \\
& \theta\int_{|\zeta-\theta|\leq\eta/\sigma_n}k'(\zeta)d\zeta\underbrace{\int_y \left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)\varphi'(y,t^*)dy}_{=H}
\end{aligned}
$$

where $H=f_x(t^*)\left(\frac{\partial\mathbb{E}_P[Y_1-Y_0|X=t^*]}{\partial X}\right)$ is bounded by Assumption 2.3. Since $\int\zeta k'(\zeta)d\zeta=0$ and $\sigma_n|\theta|\leq\frac{\eta}{2}$:

$$
\left|\int_{|\zeta-\theta|\leq\eta/\sigma_n}\zeta k'(\zeta)d\zeta\right| = \left|\int_{|\zeta-\theta|>\eta/\sigma_n}\zeta k'(\zeta)d\zeta\right| \leq \left|\int_{|\zeta|>\eta/2\sigma_n}\zeta k'(\zeta)d\zeta\right|.
$$

By Assumption 5.2, $\left|\int_{|\zeta|>\eta/2\sigma_n}\zeta k'(\zeta)d\zeta\right|$ converges to 0 uniformly over $\theta\in\Theta_n$. It means that $\int_{|\zeta-\theta|\leq\eta/\sigma_n}\zeta k'(\zeta)d\zeta$ converges uniformly to 0.

Consider $\theta H \int_{|\zeta-\theta|\leq\eta/\sigma_n} k'(\zeta)\,d\zeta$, and note that, since $\int k'(\zeta)\,d\zeta = 1$,

$$\left|\theta H - \theta H \int_{|\zeta-\theta|\leq\eta/\sigma_n} k'(\zeta)\,d\zeta\right| = \left|\theta H \int_{|\zeta-\theta|>\eta/\sigma_n} k'(\zeta)\,d\zeta\right| \leq$$
$$|\sigma_n \theta H|\sigma_n^{-1} \int_{|\zeta-\theta|>\eta/\sigma_n} k'(\zeta)\,d\zeta \leq \frac{\eta}{2}\sigma_n^{-1} \int_{|\zeta-\theta|>\eta/\sigma_n} k'(\zeta)\,d\zeta.$$

The last term is bounded uniformly over n and $\theta \in \Theta_n$ and converges to 0 by Assumption 5.2. It means that

$$\lim_{n\to\infty} \left|\sup_{\theta\in\Theta_n} J_{n2} - \theta H\right| = 0.$$

Finally, consider J_{n3}:

$$\begin{aligned}
|J_{n3}| &= \left|\frac{1}{2}\sigma_n^{-2} \int_{|X-t^*|\leq\eta} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right) k'\left(\frac{X-t^*}{\sigma_n} + \theta\right) \varphi''(y,\tilde{t})(x-t^*)^2 dydx\right| \\
&= \left|\frac{1}{2}\sigma_n \int_{|\zeta-\theta|\leq\eta/\sigma_n} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right) k'(\zeta)\,\varphi''(y,\tilde{t})(\zeta-\theta)^2 dydx\right| \\
&\leq \underbrace{\left|\frac{1}{2}\sigma_n \int_{|\zeta-\theta|\leq\eta/\sigma_n} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right) k'(\zeta)\,\zeta^2\varphi''(y,\tilde{t}) dydx\right|}_{=o(1)} + \\
&\quad \sigma_n|\theta| \underbrace{\left|\int_{|\zeta-\theta|\leq\eta/\sigma_n} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right) k'(\zeta)\,\zeta\varphi''(y,\tilde{t}) dydx\right|}_{=\alpha_1} + \\
&\quad \sigma_n|\theta|^2 \underbrace{\left|\int_{|\zeta-\theta|\leq\eta/\sigma_n} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right) k'(\zeta)\,\varphi''(y,\tilde{t}) dydx\right|}_{=\alpha_2}.
\end{aligned}$$

Combine results on I_{n2} and J_{n1}, J_{n2}, and J_{n3} to get

$$|\mathbb{E}[\hat{S}_n^\theta(\theta)] - H\theta| = |J_{n1} + J_{n2} + J_{n3} + I_{n2} - H\theta| \leq o(1) + \alpha_1\sigma_n|\theta| + \alpha_2\sigma_n\theta^2$$

uniformly over $\theta \in \Theta_n$, which proves the second part of the lemma. □

Lemma 4

Lemma 4. *Let Assumptions 1, 2 with $s = h + 1$ for some $h \geq 2$, and 5 hold. Define $\hat{\theta}_n = \frac{t^* - \hat{t}_n^s}{\sigma_n}$. Then* $\text{plim}_{n\to\infty} \hat{\theta}_n = 0$.

Proof. Consider $\hat{S}_n^\theta(\theta_n)$:

$$\begin{aligned}\hat{S}_n^\theta(\theta_n) &= -\left(n\sigma_n^2\right)^{-1} \sum_{i=1}^n \left[\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right) k'\left(\frac{X_i - t^*}{\sigma_n} + \theta_n\right)\right] \\ &= -\left(n\sigma_n^2\right)^{-1} \sum_{i=1}^n \left[\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))}\right) k'\left(\frac{X_i - \hat{t}_n^s}{\sigma_n}\right)\right].\end{aligned}$$

By Theorem 3, $\hat{t}_n^s \to^{a.s.} t^*$. With probability approaching 1, then, $\hat{t}_n^s$ is an interior point of $\mathcal{T}$. It means that, with probability approaching 1, $\hat{S}_n^\theta(\theta_n) = \hat{S}_n^1(\hat{t}_n^s, \sigma_n) = 0$. Hence lemma 3 gives

$$|H\theta_n| \leq o(1) + \alpha_1 \sigma_n |\theta_n| + \alpha_2 \sigma_n \theta_n^2$$

with $H \neq 0$ by Assumptions 2.1 and 2.3.

I will hence prove $\text{plim}_{n\to\infty} \hat{\theta}_n = 0$ by contradiction. First, assume that $\hat{\theta}_n$ has finite limit different from 0. The left-hand side of the previous inequality would be positive, while the right-hand side converges to 0. This contradicts the inequality. Then assume the limit is unbounded. By Theorem 3, $\text{plim}_{n\to\infty} \sigma_n \hat{\theta}_n = 0$. This gives the contradiction

$$\underbrace{\frac{|H\hat{\theta}_n|}{|\hat{\theta}_n|}}_{=|H|>0} \leq \underbrace{\frac{o(1)}{|\hat{\theta}_n|}}_{\to 0} + \underbrace{\frac{\alpha_1 \sigma_n |\hat{\theta}_n|}{|\hat{\theta}_n|}}_{=\alpha_1\sigma_n \to 0} + \underbrace{\frac{\alpha_2 \sigma_n \hat{\theta}_n^2}{|\hat{\theta}_n|}}_{=\alpha_2\sigma_n|\hat{\theta}_n| \to 0}$$

and proves that $\text{plim}_{n\to\infty} \hat{\theta}_n = 0$. □

Lemma 5

Lemma 5. *Let Assumptions 1, 2 with $s = h + 1$ for some $h \geq 2$, and 5 hold. Consider a sequence t_n such that $\frac{t_n - t^*}{\sigma_n} \to 0$. Then*

$$\text{plim}_{n\to\infty} \hat{S}_n^2(t_n, \sigma_n) = -H.$$

Proof. To prove the lemma it is sufficient to show that $\mathbb{E}[\hat{S}_n^2(t_n, \sigma_n)] \to H$ and $\text{Var}(\hat{S}_n^2(t_n, \sigma_n)) \to 0$. Recall that

$$\hat{S}_n^2(t, \sigma_n) = \frac{1}{\sigma_n^2} \frac{1}{n} \sum_{i=1}^{n} \left[\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1 - D_i) Y_i}{(1 - p(\mathbf{X}_i))} \right) k'' \left(\frac{X_i - t}{\sigma_n} \right) \right]$$

and hence

$$\begin{aligned} \mathbb{E}[\hat{S}_n^2(t_n, \sigma_n)] =& \frac{1}{\sigma_n^2} \mathbb{E} \left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1 - D)Y}{1 - p(\mathbf{X})} \right) k'' \left(\frac{X - t_n}{\sigma_n} \right) \right] \\ =& \frac{1}{\sigma_n^2} \int_x \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1 - D)Y}{1 - p(\mathbf{X})} \right) k'' \left(\frac{X - t_n}{\sigma_n} \right) \varphi(y, x) dy dx. \end{aligned}$$

Consider a Taylor expansion of $\varphi(y, x)$ about $x = t^*$:

$$\varphi(y, x) = \varphi(y, t^*) + \varphi'(y, t^*)(x - t^*) + \frac{1}{2} \varphi''(y, \tilde{t})(x - t^*)^2$$

with $|\tilde{t} - t^*| \leq |x - t^*|$. Let $\eta > 0$ be such that $\varphi_{y,x}(y, x)$ has second derivative uniformly bounded for almost every X if $|X - t^*| < \eta$, and write $\mathbb{E}[\hat{S}_n^2(t_n, \sigma_n)]$ as $I_{n1} + I_{n2} + I_{n3} + I_{n4}$,

where:

$$
\begin{aligned}
I_{n1} =& \sigma_n^{-2} \int_{|X-t^*|\leq\eta} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) k''\left(\frac{X-t_n}{\sigma_n} \right) \varphi(y,t^*) dy dx \\
=& \sigma_n^{-2} \underbrace{\mathbb{E}[Y_1 - Y_0 | X = t^*]}_{=0} \int_{|X-t^*|\leq\eta} k''\left(\frac{X-t_n}{\sigma_n} \right) f_x(t^*) dx = 0 \\
I_{n2} =& \sigma_n^{-2} \int_{|X-t^*|\leq\eta} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) k''\left(\frac{X-t_n}{\sigma_n} \right) \varphi'(y,t^*)(x-t^*) dy dx \\
I_{n3} =& \sigma_n^{-2} \int_{|X-t^*|\leq\eta} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) k''\left(\frac{X-t_n}{\sigma_n} \right) \frac{1}{2}\varphi''(y,\tilde{t})(x-t^*)^2 dy dx \\
I_{n4} =& \frac{1}{\sigma_n^2} \int_{|X-t^*|>\eta} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{1-p(\mathbf{X})} \right) k''\left(\frac{X-t_n}{\sigma_n} \right) \varphi(y,x) dy dx.
\end{aligned}
$$

Consider the substitution $\zeta = \frac{X-t^*}{\sigma_n} + \theta_n = \frac{X-t_n}{\sigma_n}$:

$$
\begin{aligned}
I_{n2} =& \int_{|\zeta-\theta_n|\leq\eta/\sigma_n} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) k''(\zeta)\, \varphi'(y,t^*)(\zeta-\theta_n) dy d\zeta \\
=& \int_{|\zeta-\theta_n|\leq\eta/\sigma_n} k''(\zeta)(\zeta-\theta_n) d\zeta \underbrace{\int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) \varphi'(y,t^*) dy}_{=H} \\
=& H \int_{|\zeta-\theta_n|\leq\eta/\sigma_n} \zeta k''(\zeta)\, d\zeta - \theta_n H \int_{|\zeta-\theta_n|\leq\eta/\sigma_n} k''(\zeta)\, d\zeta \\
=& H \left(\int \zeta k''(\zeta)\, d\zeta - \int_{|\zeta-\theta_n|>\eta/\sigma_n} \zeta k''(\zeta)\, d\zeta \right) - \theta_n H \int_{|\zeta-\theta_n|\leq\eta/\sigma_n} k''(\zeta)\, d\zeta
\end{aligned}
$$

Under Assumption 5.2, $\int \zeta k''(\zeta)\, d\zeta = -1$, $\int_{|\zeta-\theta_n|>\eta/\sigma_n} \zeta k''(\zeta)\, d\zeta \to^p 0$ and $\int_{|\zeta-\theta_n|\leq\eta/\sigma_n} k''(\zeta)\, d\zeta$ is bounded. Since $\theta_n \to^p 0$, $I_{n2} \to^p -H$.

Consider I_{n3} and the substitution $\zeta = \frac{X-t^*}{\sigma_n} + \theta_n = \frac{X-t_n}{\sigma_n}$:

$$
I_{n3} = \frac{1}{2}\sigma_n \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) \varphi''(y,\tilde{t}) dy \int_{|\zeta-\theta_n|\leq\eta/\sigma_n} k''(\zeta)(\zeta-\theta_n) d\zeta.
$$

Integrals are bounded by Assumptions 2.3 with $s = h+1$ and 5.2, and hence $I_{n3} \to 0$.

Finally, consider I_{n4} and the substitution $\zeta = \frac{X-t^*}{\sigma_n} + \theta_n = \frac{X-t_n}{\sigma_n}$:

$$\begin{aligned} I_{n4} =& \frac{1}{\sigma_n} \int_{|\zeta-\theta_n|>\eta/\sigma_n} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{1-p(\mathbf{X})} \right) k''(\zeta)\, \varphi(y, t^* + \sigma_n(\zeta - \theta_n)) dy d\zeta \\ \to^p & \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{1-p(\mathbf{X})} \right) \varphi(y, t^*) dy \frac{1}{\sigma_n} \int_{|\zeta-\theta_n|>\eta/\sigma_n} k''(\zeta)\, d\zeta. \end{aligned}$$

$\int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{1-p(\mathbf{X})} \right) \varphi(y, t^*) dy$ is bounded by Assumptions 2.2 and $\frac{1}{\sigma_n} \int_{|\zeta-\theta_n|>\eta/\sigma_n} k''(\zeta)\, d\zeta \to^p 0$ by Assumptions 5.2, and hence $I_{n4} \to^p 0$.

Combine results on I_{n1}, I_{n2}, I_{n3}, and I_{n4} to conclude $\mathbb{E}[\hat{S}_n^2(t_n, \sigma_n)] \to H$. Consider now the variance:

$$\begin{aligned} \mathrm{Var}(\hat{S}_n^2(t_n, \sigma_n)) =& \mathrm{Var}\left(\frac{1}{\sigma_n^2} \frac{1}{n} \sum_{i=1}^n \left[\left(\frac{D_i Y_i}{p(\mathbf{X}_i)} - \frac{(1-D_i)Y_i}{(1-p(\mathbf{X}_i))} \right) k''\left(\frac{X_i - t_n}{\sigma_n} \right) \right] \right) \\ =& \frac{1}{n\sigma_n^4} \mathrm{Var}\left(\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) k''\left(\frac{X - t_n}{\sigma_n} \right) \right) \\ =& \frac{1}{n\sigma_n^4} \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right)^2 k''\left(\frac{X - t_n}{\sigma_n} \right)^2 \right] - \\ & \frac{1}{n} \left(\underbrace{\frac{1}{\sigma_n} \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right) k''\left(\frac{X - t_n}{\sigma_n} \right) \right]}_{\to^p 0} \right)^2 \end{aligned}$$

and the substitution $\zeta = \frac{X-t^*}{\sigma_n} + \theta_n = \frac{X-t_n}{\sigma_n}$:

$$\begin{aligned} & \frac{1}{n\sigma_n^4} \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right)^2 k''\left(\frac{X - t_n}{\sigma_n} \right)^2 \right] = \\ & \frac{1}{n\sigma_n^3} \int_\zeta \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right)^2 k''(\zeta)^2\, \varphi(y, t^* + \sigma_n(\zeta - \theta_n)) dy d\zeta \to^p \\ & \frac{1}{n\sigma_n^3} \int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))} \right)^2 \varphi(y, t^*) dy \int_\zeta k''(\zeta)^2\, d\zeta. \end{aligned}$$

$\int_y \left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{1-p(\mathbf{X})} \right) \varphi(y, t^*) dy$ and $\int_\zeta k''(\zeta)^2\, d\zeta$ are bounded by Assumptions 2.2 and 5.2. Since by Assumption 5.1 $n\sigma_n^3 \to \infty$, conclude that $\mathrm{Var}(\hat{S}_n^2(t_n, \sigma_n)) \to 0$. □

Theorem 4

Theorem 4. *Consider the SWM policy $\hat{t}_n^s$ defined in equation* (3.2) *and the optimal policy t^* defined in equation* (2)*. Under Assumptions 1, 2 with $s = h + 1$ for some $h \geq 2$, and 5, as $n \to \infty$:*

1. *if $n\sigma_n^{2h+1} \to \infty$,*

$$\sigma_n^{-h}(\hat{t}_n^s - t^*) \to^p H^{-1}A;$$

2. *if $n\sigma_n^{2h+1} \to \lambda < \infty$,*

$$(n\sigma_n)^{1/2}(\hat{t}_n^s - t^*) \to^d \mathcal{N}(\lambda^{1/2}H^{-1}A, H^{-2}\alpha_2 K);$$

where A, α_1, and α_2 are:

$$\begin{aligned} A &= -\frac{1}{h!}\alpha_1 \int_y (Y_1 - Y_0)\, \varphi_x^h(y, t^*) dy \\ \alpha_1 &= \int_\zeta \zeta^h k'(\zeta)\, d\zeta \\ \alpha_2 &= \int_\zeta k'(\zeta)^2\, d\zeta. \end{aligned}$$

Proof. Consider a Taylor expansion of $\hat{S}_n^1(t, \sigma_n)$ about $t = t^*$:

$$\hat{S}_n^1(\hat{t}_n^s, \sigma_n) = \hat{S}_n^1(t^*, \sigma_n) + \hat{S}_n^2(\tilde{t}, \sigma_n)(\hat{t}_n^s - t^*)$$

with $|\tilde{t} - t^*| \leq |\hat{t}_n^s - t^*|$. By Theorem 3, $\hat{t}_n^s \to^{a.s.} t^*$, and hence with probability approaching 1 $\hat{t}_n^s$ is an interior point of $\mathcal{T}$. It means that, with probability approaching 1, $\hat{S}_n^1(\hat{t}_n^s, \sigma_n) = 0$.

To prove the first result of the theorem, note that with probability approaching one as $n \to \infty$

$$\sigma_n^{-h}\hat{S}_n^1(t^*, \sigma_n) + \sigma_n^{-h}\hat{S}_n^2(\tilde{t}, \sigma_n)(\hat{t}_n^s - t^*) = 0.$$

By lemmas 4 and 5, $\text{plim}_{n\to\infty} \hat{S}_n^2(t_n, \sigma_n) = -H$, and $H \neq 0$ by Assumptions 2.1 and 2.3. Hence

$$\sigma_n^{-h}(\hat{t}_n^s - t^*) = H^{-1}\sigma_n^{-h}\hat{S}_n^1(t^*, \sigma_n) + o_p(1)$$

and since $\sigma_n^{-h}\hat{S}_n^1(t^*, \sigma_n) \to^p A$ by lemma 2,

$$\sigma_n^{-h}(\hat{t}_n^s - t^*) \to^p H^{-1}A.$$

Analogously, to prove the second result note that

$$(n\sigma_n)^{1/2}\hat{S}_n^1(t^*, \sigma_n) + (n\sigma_n)^{1/2}\hat{S}_n^2(\tilde{t}, \sigma_n)(\hat{t}_n^s - t^*) = 0.$$

with probability approaching 1, and hence

$$(n\sigma_n)^{1/2}(\hat{t}_n^s - t^*) = H^{-1}(n\sigma_n)^{1/2}\hat{S}_n^1(t^*, \sigma_n).$$

Since $(n\sigma_n)^{1/2}\hat{S}_n^1(t^*, \sigma_n) \to^d \mathcal{N}(\lambda^{1/2}A, \alpha_2 K)$ by lemma 2,

$$(n\sigma_n)^{1/2}(\hat{t}_n^s - t^*) \to^d \mathcal{N}(\lambda^{1/2}H^{-1}A, H^{-2}\alpha_2 K).$$

For the third result, first compute the asymptotic bias and the asymptotic variance of $\hat{t}_n^s - t^*$:

$$\begin{aligned}
\mathbb{E}[\hat{t}_n^s - t^*] &= -\lambda^{1/2}\frac{A}{H}(n\sigma_n)^{-1/2} = -\lambda^{1/2}\frac{A}{H}\lambda^{-\frac{1}{2(2h+1)}}n^{-\frac{h}{2h+1}} \\
\text{Var}(\hat{t}_n^s - t^*) &= \frac{\alpha_2 K}{H^2}\lambda^{-\frac{1}{2h+1}}n^{-\frac{2h}{2h+1}}
\end{aligned}$$

and then the MSE:

$$MSE = \frac{1}{H^2} n^{-\frac{2h}{2h+1}} \left[\alpha_2 K \lambda^{-\frac{1}{2h+1}} + A^2 \lambda^{\frac{2h}{2h+1}} \right]$$

which is minimize setting

$$\lambda = \lambda^* = \frac{\alpha_2 K}{2hA^2}.$$

□

Corollary 4.1

Corollary 4.1. *Asymptotic distribution of regret $\mathcal{R}(\hat{t}_n^s)$ is:*

$$n\sigma_n \mathcal{R}(\hat{t}_n^s) \to^d \frac{1}{2} \frac{\alpha_2 K}{H} \chi^2 \left(1, \frac{\lambda A^2}{\alpha_2 K}\right)$$

where $\chi^2 \left(1, \frac{\lambda A^2}{\alpha_2 K}\right)$ is a non-centered chi-squared distribution with 1 degree of freedom and non-central parameter $\frac{\lambda A^2}{\alpha_2 K}$. The expected value of the asymptotic distribution is:

$$\frac{1}{2} \frac{\alpha_2 K}{H} \left(1 + \frac{\lambda A^2}{\alpha_2 K}\right) = \frac{\alpha_2}{2} \frac{K}{H} + \frac{1}{2} \frac{\lambda A^2}{H}.$$

Let $\sigma_n = (\lambda/n)^{1/(2h+1)}$ with $\lambda \in (0, \infty)$. The expectation of the asymptotic regret is minimized by setting $\lambda = \lambda^* = \frac{\alpha_2 K}{2hA^2}$: in this case the expectation of the asymptotic distribution scaled by $n^{\frac{2h}{2h+1}}$ is $A^{\frac{2}{2h+1}} K^{\frac{2h}{2h+1}} H^{-1} C^s$, where $C^s = \frac{2h+1}{2} \left(\frac{\alpha_2}{2h}\right)^{\frac{2h}{2h+1}}$ is a constant not dependent on P.

Proof. Result in equation (2) for $\hat{t}_n^s$ implies

$$n\sigma_n \mathcal{R}(\hat{t}_n^s) \to^d \frac{1}{2} W''(\tilde{t}) \left((n\sigma_n)^{1/2} \left(\hat{t}_n^s - t^*\right)\right)^2.$$

where $|\tilde{t} - t^*| \leq |\hat{t}_n^s - t^*|$. By continuous mapping theorem

$$W''(\tilde{t}) \to^p W''(t^*) = H$$

and hence by Slutsky's theorem

$$n\sigma_n \mathcal{R}(\hat{t}_n^s) \to^d \frac{1}{2} H \left(\mathcal{N}(\lambda^{1/2} H^{-1} A, H^{-2}\alpha_2 K)\right)^2 =^d \frac{1}{2}\frac{\alpha_2 K}{H}\left(\mathcal{N}(\lambda^{1/2} A/\sqrt{D}, 1)\right)^2.$$

By definition, $\chi^2\left(1, \frac{\lambda A^2}{\alpha_2 K}\right) =^d \left(\mathcal{N}(\lambda^{1/2} A/\sqrt{D}, 1)\right)^2$, and $\mathbb{E}\left[\chi^2\left(1, \frac{\lambda A^2}{\alpha_2 K}\right)\right] = \left(1 + \frac{\lambda A^2}{\alpha_2 K}\right)$.

When $\sigma_n = (\lambda/n)^{1/(2h+1)}$, the expectation of asymptotic regret is minimized by

$$\lambda^* = \arg\min_\lambda (n\sigma_n)^{-1}\left(\frac{\alpha_2}{2}\frac{K}{H} + \frac{1}{2}\frac{\lambda A^2}{H}\right) = \arg\min_\lambda \alpha_2 K \lambda^{-\frac{1}{2h+1}} + A^2 \lambda^{\frac{2h}{2h+1}}$$

which is solved by $\lambda^* = \frac{\alpha_2 K}{2hA^2}$.

By substituting σ_n by $(\lambda/n)^{1/(2h+1)}$, and λ by $\frac{\alpha_2 K}{2hA^2}$, the expectation of the asymptotic regret multiplied by $(n\sigma_n)^{-1}$ is

$$(n\sigma_n)^{-1}\left(\frac{\alpha_2}{2}\frac{K}{H} + \frac{1}{2}\frac{\lambda A^2}{H}\right) = n^{-\frac{2h}{2h+1}} A^{\frac{2}{2h+1}} K^{\frac{2h}{2h+1}} H^{-1} \frac{2h+1}{2}\left(\frac{\alpha_2}{2h}\right)^{\frac{2h}{2h+1}}$$

and the expectation of the asymptotic distribution scaled by $n^{\frac{2h}{2h+1}}$ is $A^{\frac{2}{2h+1}} K^{\frac{2h}{2h+1}} H^{-1} C^s$, where $C^s = \frac{2h+1}{2}\left(\frac{\alpha_2}{2h}\right)^{\frac{2h}{2h+1}}$ is a constant not dependent on P. □

Theorem 6

Theorem 5. *Let Assumptions 1, 2 with $s = h+1$ for some $h \geq 2$, 3, and 5 hold. Consider sequences $\{P_n^1\}$ and $\{P_n^2\}$ of data generating processes. The rank of r_n^e and r_n^s depends on γ as follows:*

- $r_n^e > r_n^s$ *if* $\gamma \in (\bar{\gamma}, \frac{h}{2h+1})$
- $r_n^e = r_n^s$ *if* $\gamma = \bar{\gamma}$

- $r_n^e < r_n^s$ *if* $\gamma \in [0, \bar{\gamma})$.

where $\bar{\gamma} = \frac{h-1}{2h+1}$ *under* $\{P_n^1\}$, *and* $\bar{\gamma} = \frac{1}{3}\frac{h-1}{2h+1}$ *under* $\{P_n^2\}$.

Proof. Under Assumptions 1, 2 with $s = h+1$ for some $h \geq 2$, 3, and 5, Theorems 2 and 4, for a fixed population distribution P and with $\sigma_n = (\lambda/n)^{1/(2h+1)}$, imply

$$n^{1/3}\left(\hat{t}_n^e - t^*\right) \to^d \left(\frac{\sqrt{K}}{H}\right)^{2/3} D^e \tag{9}$$
$$n^{\frac{h}{2h+1}}(\hat{t}_n^s - t^*) \to^d \frac{\sqrt{K}}{H} D^s + \frac{A}{H}$$

where D^e and D^s are known distributions which do not depend on population parameters.

Results in equation (D) are valid for any distribution P_n^1 belonging to the sequence $\{P_n^1\}$, and hence along the sequence imply:

$$n^{\frac{1}{3}-\frac{2}{3}\gamma}\left(\hat{t}_n^e - t^*\right) \to^d D^e$$
$$n^{\frac{h}{2h+1}-\gamma}(\hat{t}_n^s - t^*) \to^d D^s + O_p(n^{-\gamma}).$$

The comparison of the rate of convergence of $\hat{t}_n^e$ and $\hat{t}_n^s$ when data are distributed according to $\{P_n^1\}$ depends hence on γ, in the following way:

- $\gamma \in (\frac{h-1}{2h+1}, \frac{h}{2h+1})$: $\frac{1}{3} - \frac{2}{3}\gamma > \frac{h}{2h+1} - \gamma$, and hence the rate of convergence of $\hat{t}_n^e$ is faster.
- $\gamma = \frac{h-1}{2h+1}$: $\frac{1}{3} - \frac{2}{3}\gamma = \frac{h}{2h+1} - \gamma$, and hence the rate of convergence of $\hat{t}_n^e$ and $\hat{t}_n^s$ is the same rate.
- $\gamma \in [0, \frac{h-1}{2h+1})$: $\frac{1}{3} - \frac{2}{3}\gamma < \frac{h}{2h+1} - \gamma$, a and hence the rate of convergence of $\hat{t}_n^s$ is faster.

Result in equation (2) for $\hat{t}_n^s$ implies

$$r_n^e \mathcal{R}(\hat{t}_n^e) \to^d \frac{1}{2} W''(\tilde{t}) \left((r_n^e)^{1/2}\left(\hat{t}_n^e - t^*\right)\right)^2.$$

where $|\tilde{t} - t^*| \leq |\hat{t}_n^e - t^*|$ and $(r_n^e)^{1/2}$ is the rate of convergence of $\hat{t}_n^e$. Since $\hat{t}_n^e$ is consistent for t^*, by continuous mapping theorem $W''(\tilde{t}) \to^p W''(t^*) = H$, and hence

$$r_n^e \mathcal{R}(\hat{t}_n^e) \to^d \frac{1}{2} H \left((r_n^e)^{1/2} \left(\hat{t}_n^e - t^* \right) \right)^2 .$$

An analogous result can be proved for $\mathcal{R}(\hat{t}_n^s)$. The order relationship between r_n^e and r_n^s is then the same as the order relationship between $(r_n^e)^{1/2}$ and $(r_n^s)^{1/2}$. Specifically, it depends on γ in the same way, and then the result of the theorem for $\{P_n^1\}$ follows.

An analogous argument proves the result for a sequence $\{P_n^2\}$. Along this sequence, results in equation (D) imply:

$$n^{\frac{1}{3}} \left(\hat{t}_n^e - t^* \right) \to^d D^e$$
$$n^{\frac{h}{2h+1} - \gamma} (\hat{t}_n^s - t^*) \to^d 1 + O_p(n^{-\gamma}).$$

The comparison of the rate of convergence of $\hat{t}_n^e$ and $\hat{t}_n^s$ when data are distributed according to $\{P_n^2\}$ depends hence on γ, in the following way:

- $\gamma \in (\frac{1}{3}\frac{h-1}{2h+1}, \frac{h}{2h+1})$: the rate of convergence of $\hat{t}_n^e$ is faster.
- $\gamma = \frac{1}{3}\frac{h-1}{2h+1}$: the rate of $\hat{t}_n^e$ and $\hat{t}_n^s$ is the same.
- $\gamma \in [0, \frac{1}{3}\frac{h-1}{2h+1})$: the rate of convergence of $\hat{t}_n^s$ is faster.

As for $\{P_n^1\}$, under $\{P_n^2\}$ the order relationship between r_n^e and r_n^s is the same as the order relationship between rates of convergence of $\hat{t}_n^e$ and $\hat{t}_n^s$. It depends on γ in the same way, and hence the result of the theorem for $\{P_n^2\}$ follows.

□

Theorem 5

Theorem 6. *Consider estimators $\hat{t}_n^e$ defined in equation* (3.1) *and $\hat{t}_n^b$ defined in equation* (A.1.1) *and the estimand t^* defined in equation* (2)*. Under Assumptions 1, 2 with $s = 2$, 3, and 6, as $\hat{H}_n \to^p H$ and $n \to \infty$,*

$$n^{1/3}\left(\hat{t}_n^b - \hat{t}_n\right) \to^d (2\sqrt{K}/H)^{2/3} \arg\max_r \left(B(r) - r^2\right)$$

where the limiting distribution is the same as in Theorem 2.

Proof. The result follows from the main theorem in Cattaneo et al. (2020). I will show that the assumptions for their results hold. My case is the benchmark case with $m_n = m_0$ (in my notation, m), $M_n = M_0$ (in my notation, Pm) and $q_n = 1$. Hence, in my case, class $\mathcal{M}_n$ coincides with m. I will verify the five conditions CRA:

1. Consider envelope $F = 2\left|\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right|$: Assumption 2.2 guarantees it is square integrable.

   Since $M_n = M_0$, $\sup_t |M_n(t) - M_0(t)| = 0$. Under Assumption 2.1, t^* is the unique maximizer of m, and hence $\sup_{t \neq t^*} m(t) < m(t^*)$.

2. t^* is an interior point of $\mathcal{T}$ by Assumption 2.1. Assumption 2.3 with $s = 2$ guarantees that Pm is twice continuously differentiable in a neighborhood of t^*.

3. I proved that this condition is satisfied in the proof of Theorem 2, under Assumption 2.2. In my notation, $\delta = R$, $\mathcal{D}_n^{\delta'} = \mathcal{G}_R$, $\bar{d}_n^{\delta'} = G_R$.

4. Note that:

$$\begin{aligned}\mathbb{E}[G_R^4] &= \mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^4 \mathbf{1}\{|X - t^*| < R\}\right] \\ &= 2R\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})} - \frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^4 \middle| X = t^*\right] + o(1).\end{aligned}$$

and that

$$\mathbb{E}\left[\left(\frac{DY}{p(\mathbf{X})}-\frac{(1-D)Y}{(1-p(\mathbf{X}))}\right)^4\right]=\frac{1}{p(\mathbf{X})^3}\mathbb{E}\left[Y_1^4\right]+\frac{1}{(1-p(\mathbf{X}))^3}\mathbb{E}\left[Y_0^4\right].$$

Let $R=O(n^{-1/3})$. It follows from Assumption 6 that

$$\begin{aligned}n^{-1/3}\mathbb{E}[G_R^4]=&2n^{-1/3}R\left(\frac{1}{p(\mathbf{X})^3}\mathbb{E}\left[Y_1^4\right]+\frac{1}{(1-p(\mathbf{X}))^3}\mathbb{E}\left[Y_0^4\right]\right)+o(1)\\=&o(1).\end{aligned}$$

The second part of assumption 4 is the same as the first part of assumption 5 in Theorem D by Kim and Pollard (1990). The only difference is that it must be valid for $t=t^*$ and t in a neighborhood of t^*. Since Assumptions 2.2 and 2.3 with $s=2$ are valid also in a neighborhood of t^*, the argument provided before holds also here.

5. The first part of this assumption is the same as assumption 6 in Kim and Pollard (1990), and the second part is the same as assumption 7 in Kim and Pollard (1990). Under Assumptions 2.3 with $s=2$, and 3, arguments provided before hold.

CRA assumptions 1-5 by Cattaneo et al. (2020) are satisfied; hence, their results in Theorem 1 hold. It implies

$$n^{1/3}\left(\hat{t}_n^b-\hat{t}_n\right)\to^d Q(r)$$

where $Q(r)=Q_1(r)+Q_0(r)$, and Q_1 is a non degenerate zero-mean Gaussian process, while $Q_0(r)=-\frac{1}{2}r^2H$. Process $Q(r)$ is the same as in Theorem 2. □