arXiv:2205.08586v2 [econ.EM] 20 Jul 2022

Treatment Choice with Nonlinear Regret*

Toru Kitagawa† Sokbae Lee‡ Chen Qiu§

July 2022

Abstract

The literature on treatment choice focuses on the mean of welfare regret. Ignoring other features of the regret distribution, however, can lead to an undesirable rule that suffers from a high chance of welfare loss due to sampling uncertainty. We propose to minimize the mean of a *nonlinear transformation* of welfare regret. This paradigm shift alters optimal rules drastically. We show that for a wide class of nonlinear criteria, admissible rules are fractional. Focusing on mean square regret, we derive the closed-form probabilities of randomization for finite-sample Bayes and minimax optimal rules when data are normal with known variance. The minimax optimal rule is a simple logit based on the sample mean and agrees with the posterior probability for positive treatment effect under the least favorable prior. The Bayes optimal rule with an uninformative prior is different but produces quantitatively comparable mean square regret. We extend these results to limit experiments and discuss our findings through sample size calculations.

Keywords: Statistical decision theory, treatment assignment rules, mean square regret, limit experiments

*We would like to thank Tommaso Denti, Charles Manski, Francesca Molinari, José Luis Montiel Olea, Jörg Stoye, Aleksey Tetenov, and the participants at Bristol Econometric Study Group Conference, Cowles Foundation Conference, Cornell Brown Bag Seminar, IAAE 2022, NASMES 2022 for helpful comments. The authors gratefully acknowledge financial support from ERC grants (numbers 715940 for Kitagawa and 646917 for Lee) and the ESRC Centre for Microdata Methods and Practice (CeMMAP) (grant number RES-589-28-0001).

†Department of Economics, Brown University and Department of Economics, University College London. Email: toru_kitagawa@brown.edu

‡Department of Economics, Columbia University. Email: sl3841@columbia.edu

§Department of Economics, Cornell University. Email: cq62@cornell.edu

1 Introduction

Evidence-based policy making using randomized control trial data is becoming increasingly common in various fields of economics. How should we use data to inform an optimal policy decision in terms of social welfare? Building on the framework of statistical decision theory as laid out in Wald (1950), the literature on statistical treatment choice initiated by Manski (2004) analyzes how to use data to inform a welfare optimal policy. Following Savage (1951) and Manski (2004), researchers often focus on the average of welfare regret across the sampled data (called *expected regret*) and obtain an optimal decision rule by minimizing a worst-case expected regret.

When it comes to the ranking of different statistical decision rules, once we eliminate those that are stochastically dominated, it becomes less obvious how we should compare decision rules that do not stochastically dominate each other. Focusing on the expected regret, as suggested by Manski (2004), provides a natural starting point. In general, regardless of whether we consider a Bayes or minimax criterion, optimal decision rules defined in terms of their expected regret are nonrandomized, i.e., given a sample, optimal decision rules either treat everyone, or no-one in the population. There is, however, no compelling argument why we should limit our attention to the mean of regret, as has been acknowledged by Manski and Tetenov (2014) and Manski (2021a). In fact, focusing solely on the mean of regret and ignoring other features of the distribution of regret (e.g., second- or higher-order moments and tail probabilities) can lead to rules that incur a large welfare loss due to random sampling errors, especially when the sample size is small. As an artificial example, suppose that the outcome of interest is $+1$ or -1 (success or failure) and imagine that we observe 100 successes and 99 failures (the status quo is zero for everyone). The empirical success (ES) rule, which is asymptotically optimal in terms of the mean of regret, suggests that everyone in the entire population should be treated. If there is a swing of one outcome from $+1$ to -1 though, then the same ES rule now dictates that no-one should be treated. Such high sensitivity of treatment decisions with respect to sampling uncertainty implies that, given a sample, there is always a non-negligible probability that ES rule incurs a large welfare loss.

To address these concerns, this paper proposes a novel approach to statistical treatment choice by optimizing a *nonlinear transformation* of welfare regret. In what follows, we let $g(\cdot)$ be a nonlinear transformation of the regret loss. We assess the performance of each treatment rule via the expected value of the transformed regret loss that it delivers. In the spirit of Wald (1950), this average nonlinear regret over realizations of the sampling process becomes the risk function. We refer to this risk as a *nonlinear regret risk*. Due to the nonlinearity of $g(\cdot)$, information relating to other moments of the regret distribution is encoded in the risk function. For example, when $g(r) = r^2$, the associated risk function is the sum of the squared expected regret and the variance of regret, penalizing decision rules that lead to a high variance of regret. We refer to this nonlinear regret risk as *mean square regret*.

This shift of criterion towards a nonlinear transformation of regret changes optimal rules drastically. We show that, for a wide class of nonlinear regret risks, including mean square regret, any deterministic decision rule is dominated by some randomized decision rule (i.e., fractional assignment rule). That is, deterministic decision rules are inadmissible once we take other moments or features of the regret distribution into account. This offers a novel decision-theoretic justification for implementing a randomized (fractional) treatment assignment rule, which differs from justifications given in the existing literature so far, which all use the standard expected regret as the criterion. See, for example, Manski (2009) for a detailed review of randomized rules with standard regret under ambiguity and other non-standard settings, including nonlinear welfare, interacting treatments, learning and other non-cooperative aspects. More specifically, when the linear welfare is partially identified, Manski (2000, 2005, 2007a,b) show that minimax regret optimal rules are randomized even with the true knowledge of the identified set. One approach is to plug in an estimate of the identified set, treating it *as if* it is the true object (Manski, 2013; Cassidy and Manski, 2019; Manski, 2021b). The other approach is to directly consider finite sample minimax regret optimal rules, which can be also randomized (Stoye, 2012; Yata, 2021; Manski, 2022). As shown by Manski and Tetenov (2007) and Manski (2009), randomized rules can also be justified by a nonlinear welfare in a point-identified setting.

Our results justify randomized rules in a standard setting without ambiguity or nonlinear welfare. We provide general results on Bayes and minimax optimal rules based on nonlinear regret risks. For mean square regret, we derive both Bayes and minimax optimal decision rules, not only in Gaussian finite samples with known variance but also asymptotically. These optimal rules are randomized, with the probability of assignment to the treatment dependent on the t-statistic for the average treatment effect estimated from experimental data. The probability of randomized assignment has a simple and insightful expression that is easy to compute in practice. For example, an asymptotically minimax optimal rule for the previous artificial example would allocate treatment to the population with a probability of only 54%, dropping to 46% if one outcome switches.

We show that the form that our treatment rules take is closely related to the posterior distribution for the average treatment effect. In particular, the minimax mean square regret rule coincides with the posterior probability-matching assignment under the least favorable prior. The posterior probability-matching assignment, known as the Thompson sampling algorithm (Thompson, 1933), possesses a desirable exploration-exploitation property in bandit problems. Our results show that the posterior probability-matching assignment can be justified in terms of minimax mean square regret, even in the static treatment choice problem where the exploration motive does not exist.

Given a nonlinear regret risk and a prior for the underlying potential outcome distributions, we obtain the Bayes optimal rules. Consistent with our admissibility results, Bayes optimal rules are, in general, also randomized rules. For mean square regret, we show that the Bayes optimal rule is a *tilted* posterior-probability matching rule, where the probability of random assignment corresponds to the posterior probability tilted by a weighting term determined by $g(\cdot)$. In a special case where the prior for the average treatment effect is supported only on two symmetric points, the tilting term is nullified and the Bayes optimal rule boils down to the Thompson-sampling type posterior-probability matching rule. For the minimax optimal rule in a Gaussian experiment with known variance, we can show that a least favorable prior is supported on two symmetric points. Hence, the minimax optimal rule follows the posterior-probability matching assignment rule, and is a logistic transformation of the sample mean. This minimax mean square regret rule is easy to compute and tuning parameter free.

Imagine the outcome of interest now follows a normal distribution $N(1, 1)$ with unit mean and unit variance, whereas the status quo is zero for everyone. In this scenario, the infeasible optimal rule is to treat everyone and the regret of any decision rule is supported on $[0, 1]$. Suppose the planner observes one observation from the

$N(1,1)$ distribution and needs to make a treatment choice. The ES rule is optimal in terms expected regret, but could be far from ideal in terms of other features of the regret distribution. In fact, if the planner adopted ES rule, then there would be a mass of 16% probability that she ended up with the largest possible regret of one. In contrast with the mean regret criterion commonly used in the literature, our mean square regret criterion penalizes rules with large variance of the regret distribution. If, instead, the planner implemented our proposed minimax rule, she could avert such high chance of welfare loss: the probability of incurring a regret larger than 0.95 is only 1.4%. Also see Figure 1.1 for a comparison of the distributions of the regret for ES rule and our proposed mean square regret minimax optimal rule.

In practice, the planner often has a preference for deterministic rules, and calculates a sufficient sample size for their randomized experiment based on these deterministic rules. We also show that implementing these deterministic rules can lead to a large efficiency loss in terms mean square regret. For example, to guarantee the same mean square regret with our proposed minimax optimal rule, ES rule and hypothesis testing rule require 40% and 1100% more observations, respectively.

	ES rule	Our proposed minimax rule
Mean of regret	0.1587	0.2077
Standard deviation of regret	0.3653	0.2650
Mean square regret	0.1587	0.1133

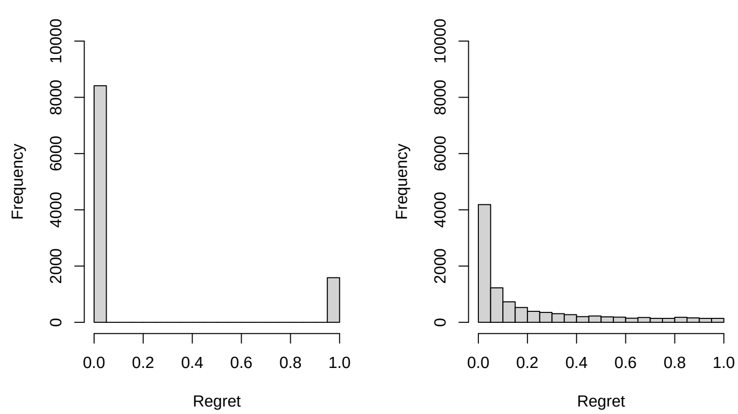

Figure 1.1: Summary statistics and empirical distributions of regret for the ES rule (left) and our proposed minimax optimal rule (right) in one $N(1,1)$ experiment, 10000 simulations.

Following Hirano and Porter (2009, 2020), we extend our finite sample results to a large sample setting by engaging with the limit experiments framework introduced by Le Cam (1986). Even when potential outcome distributions are non-Gaussian but belong to a regular parametric class, we can obtain a Gaussian limit experiment with a known variance. Therefore, we can apply our results from a finite sample Gaussian experiment to a limit experiment and find feasible and asymptotically optimal rules with some efficient estimator of the parameters. Interestingly, in the limit experiment, the Bayes optimal rule under the mean square regret remains different from the minimax optimal rule, although the resulting mean square regret is quantitatively similar between the two rules. This is in contrast with the *linear* regret risk, for which it is known that the Bayes optimal and minimax optimal rules in the limit experiment are the same empirical success rule.

In a series of papers, Manski and Tetenov explored optimal treatment rules in frameworks that go beyond the classical paradigm of the statistical decision theory laid out by Wald (1950). Manski (1988, 2011) argues to maximize a functional of the welfare distribution that at least weakly respects stochastic dominance. Manski and Tetenov (2014) consider the performance of a statistical treatment rule measured in terms of quantiles of the welfare. Our approach is distinct from the approaches taken by the aforementioned papers. In particular, we select treatment rules based on the distributions of their regret. Motivated by the risk aversion of policy makers, Manski and Tetenov (2007) consider a concave and monotone transformation of welfare measured in terms of a binary outcome, and define regret in terms of the transformed welfare. Manski and Tetenov (2007) further show that the fractional monotone rules are essentially complete. However, the set of fractional monotone rules does not exclude deterministic rules. Our approach is also different from the approach of risk averse welfare criteria taken by Manski and Tetenov (2007). To compare rules based on features of the regret distribution other than the mean, we look at a nonlinear transformation of regret. Manski and Tetenov (2007), in contrast, look at a concave transformation of the outcome. In Appendix A, we discuss how our analysis differs from that of Manski and Tetenov (2007) in greater detail.

The literature on the treatment choice problem has become an area of active research since the pioneering works of Manski (2000, 2002, 2004) and Dehejia (2005) introduced a decision theoretic framework to the problem. When the welfare is point-identified, minimax regret treatment choice rules for finite samples are derived in Schlag (2006) and Stoye (2009). Tetenov (2012) considers asymmetric regret criteria. Hirano and Porter (2009, 2020) introduce an asymptotic framework to analyze treatment rules with limit experiments. When the welfare is partially identified, Manski (2000, 2005, 2007a,b, 2009) analyzes the treatment choice given the knowledge of the identified set. Treatment allocation analysis without the knowledge of the identified set but with finite sample data include Stoye (2012), Christensen et al. (2020), Ishihara and Kitagawa (2021), Yata (2021) and Manski (2022). Chamberlain (2011) investigates a Bayesian approach to treatment choice, and Christensen et al. (2020) and Giacomini et al. (2021) discuss a robust Bayesian approach.

There is a growing literature on learning in the context of individualized treatment rules that map an individual's observable characteristics to a treatment. See Manski (2004), Bhattacharya and Dupas (2012), Kitagawa and Tetenov (2018, 2021), Mbakop and Tabord-Meehan (2021), and Athey and Wager (2021), among others. Our analysis does not incorporate individuals' observable covariates. Since the nonlinear regret risk aggregates the conditional nonlinear regret risk additively, it is straightforward to incorporate observable discrete covariates into our analysis, i.e., an optimal individualized fractional assignment rule that applies an optimal fractional assignment rule to each subpopulation of individuals sharing the same covariate value.

The rest of the paper is organised as follows. Section 2 introduces our setup. Section 3 studies the admissibility of decision rules with nonlinear regret risk. Section 4 presents finite sample results on Bayes and minimax optimal decision rules. Section 5 extends our results to the limit experiment framework and derives asymptotically optimal decision rules. Section 6 evaluates the efficiency loss of using common deterministic rules via the lens of sample size calculations. Section 7 concludes. Lengthy proofs and lemmas are reserved for the Appendix.

2 Setup

Consider the assignment of a binary treatment $D \in \{1, 0\}$ to a population of individuals whose treatment effects can be heterogeneous. Let $Y(1)$ be the potential outcome when $D = 1$ (with treatment) and $Y(0)$ be the potential outcome when $D = 0$ (no treatment). Denote by $P \in \mathcal{P}$ the joint distribution of $(Y(1), Y(0))$.

Define $\mu_1 := \mathbb{E}[Y(1)]$ and $\mu_0 := \mathbb{E}[Y(0)]$ as the means of the potential outcomes $Y(1)$ and $Y(0)$ under the distribution P. We assume that the welfare of the planner is determined by the mean outcome in the population. Defining the population average treatment effect as $\tau := \mu_1 - \mu_0$, the infeasible optimal treatment policy is as follows: allocate $D = 1$ to each individual in the population if $\tau \geq 0$ and allocate everyone $D = 0$ otherwise.

Since τ is unknown, the planner collects an experimental sample of the observed outcomes of n units randomly drawn from the population P. We assume the experimental design is known to the planner. The experiment generates a random vector $Z_n := \{Y_i, D_i\}_{i=1}^n \in \mathbf{Z}_n$, where Y_i is the observed outcome of unit i, D_i is the treatment status of unit i, and $\mathbf{Z}_n$ is the sampling space. Let P^n be the sampling distribution of Z_n, which depends on P as well as the known experimental design.

After observing data Z_n, the planner chooses a statistical treatment rule $\hat{\delta}$ that maps $Z_n \in \mathbf{Z}_n$ to a real number between 0 and 1, i.e.,

$$\hat{\delta} : \mathbf{Z}_n \to [0, 1],$$

where $\hat{\delta}(z_n)$ can be interpreted as the probability of treating one individual according to a randomization device after observing $Z_n = z_n$.

Remark 2.1. We say $\hat{\delta}$ is a *non-randomized* or *deterministic* rule if $\hat{\delta}(z_n) \in \{0, 1\}$ for almost all $z_n \in \mathbf{Z}_n$. We say $\hat{\delta}$ is *randomized* or *fractional* if $0 < \hat{\delta}(z_n) < 1$ for almost all $z_n \in \mathbf{Z}_n$.

Applying the statistical treatment rule $\hat{\delta}$ to the population yields a welfare of

$$W(\hat{\delta}) := W(\hat{\delta}, P) := \mu_1 \hat{\delta} + \mu_0 (1 - \hat{\delta})$$

to the planner. The infeasible optimal treatment policy that maximizes welfare is $\delta^* := \mathbf{1}\{\tau \geq 0\}$. Following Savage (1951) and Manski (2004), we define the regret of $\hat{\delta}$ as its welfare compared to the welfare of δ^*, i.e.,

$$Reg(\hat{\delta}) := Reg(\hat{\delta}, P) := \tau[1\{\tau \geq 0\} - \hat{\delta}].$$

Since $Reg(\hat{\delta})$ is a random object that depends on realizations of the random vector Z_n, Manski (2004) follows Wald (1950) in measuring the performance of $\hat{\delta}$ using its risk, i.e., the expected regret across realizations of the sampling process:

$$R(\hat{\delta}, P) := \mathbb{E}_{P^n}[Reg(\hat{\delta})] := \int_{z_n \in \mathbf{Z}_n} Reg(\hat{\delta}(z_n)) dP^n(z_n),$$

where $\mathbb{E}_{P^n}$ denotes the expectation with respect to P^n.

The risk criterion $R(\hat{\delta}, P)$ ranks treatment rules by the means of their regret. We, instead, consider a planner whose assessment of the performance of statistical treatment rules depends not only on the mean of regret but also on some other features of the regret distribution. To take other features of the regret distribution into consideration, we look at the nonlinear transformation of regret:

$$g(Reg(\hat{\delta})),$$

where $g : \mathbb{R}^+ \to \mathbb{R}$ is some nonlinear function. The planner's preference over statistical decision rules $\hat{\delta}$ is measured by the expected value of $g(Reg(\hat{\delta}))$ with respect to realizations of Z_n:

$$R_g(\hat{\delta}, P) := \mathbb{E}_{P^n}[g(Reg(\hat{\delta}))]. \tag{2.1}$$

We refer to the criterion $R_g(\hat{\delta}, P)$ as the *nonlinear regret risk*. Due to the nonlinearity of $g(\cdot)$, $R_g(\hat{\delta}, P)$ depends not only on the mean but also on other features of the regret distribution, including its higher-order moments. For instance, if we specify the quadratic function $g(r) = r^2$, the squared regret is

$$(Reg(\hat{\delta}))^2 = \tau^2[1\{\tau \geq 0\} - \hat{\delta}]^2.$$

This squared regret constitutes the new loss function, and we can evaluate the performance of $\hat{\delta}$ via *mean square regret*:

$$R_{sq}(\hat{\delta}, P) := \tau^2 \mathbb{E}_{P^n}[1\{\tau \geq 0\} - \hat{\delta}]^2.$$

Remark 2.2. Similar to in classical estimation theory, we can decompose

$$R_{sq}(\hat{\delta}, P) = \left[R(\hat{\delta}, P)\right]^2 + V(\hat{\delta}, P),$$

where $R(\hat{\delta}, P)$ is the mean regret risk, and

$$V(\hat{\delta}, P) := \mathbb{E}_{P^n}\left[\tau(1\{\tau \geq 0\} - \hat{\delta}) - \tau\mathbb{E}_{P^n}[1\{\tau \geq 0\} - \hat{\delta}]\right]^2$$

is the variance of the regret $Reg(\hat{\delta})$. Therefore, in addition to the standard mean regret criterion $R(\hat{\delta}, P)$, mean square regret also takes into account the variance of regret. Ranking treatment rules by the mean square regret criterion thus has the benefit of penalizing rules with high regret variance.

3 Inadmissibility of deterministic rules

Viewing the nonlinear regret risk $R_g(\hat{\delta}, P)$ defined in (2.1) as the risk criterion within Wald's framework of statistical decision theory, we introduce the following definition of admissibility of a statistical treatment rule:

Definition 3.1 (Admissibility and inadmissibility under nonlinear regret risk)**.**

(i) A statistical treatment choice rule $\hat{\delta} : \mathbf{Z}_n \to [0, 1]$ is admissible under the nonlinear regret risk $R_g(\hat{\delta}, P) = \mathbb{E}_{P^n}[g(Reg(\hat{\delta}))]$ if no $\hat{\delta}' \neq \hat{\delta}$ dominates $\hat{\delta}$, i.e., there is no $\hat{\delta}'$ such that $R_g(\hat{\delta}', P) \leq R_g(\hat{\delta}, P)$ holds for all P with the inequality strict for some P.

(ii) A statistical treatment choice rule $\hat{\delta} : \mathbf{Z}_n \to [0, 1]$ is inadmissible under the nonlinear regret $R_g(\hat{\delta}, P)$ if there exists a decision rule $\hat{\delta}' \neq \hat{\delta}$ that dominates $\hat{\delta}$.

As in standard statistical decision theory, admissibility of $\hat{\delta}$ defined through nonlinear regret is a minimal requirement that a desirable statistical treatment rule should satisfy.

Assumption G (Nonlinear Transformation)**.** The nonlinear transformation $g : \mathbb{R}^+ \to \mathbb{R}$ is differentiable and $g(\cdot)$ is strictly increasing on $\mathbb{R}^+ \setminus \{0\}$ with $g'(0) = 0$.

Assumption G puts mild restrictions on the shape of the nonlinear transformation. Together with Assumption G, the next theorem shows that, in terms of nonlinear regret risk, deterministic assignment rules are inadmissible.

Theorem 3.1. *Consider a deterministic rule $\hat{\delta}_D$ that is nondegenerate at some $P \in \mathcal{P}$ with $\tau \neq 0$, i.e., $P^n(\hat{\delta}_D = 1) \in (0,1)$ for some P with $\tau \neq 0$. If Assumption G holds, then there exists a randomized rule $\hat{\delta}_R$ that dominates $\hat{\delta}_D$.*

Proof. Given a deterministic rule $\hat{\delta}_D$, we establish the existence of a randomized rule $\hat{\delta}_R$ that yields $R_g(\hat{\delta}_D, P) \geq R_g(\hat{\delta}_R, P)$ for all $P \in \mathcal{P}$ with the inequality strict for some $P \in \mathcal{P}$.

Given $\hat{\delta}_D$, consider the following randomized rule:

$$\hat{\delta}_R = (1-\lambda)\hat{\delta}_D + \lambda(1-\hat{\delta}_D).$$

When $\hat{\delta}_R$ is implemented, its regret is

$$\begin{aligned} Reg(\hat{\delta}_R) &= \tau\left[1\{\tau \geq 0\} - (1-\lambda)\hat{\delta}_D - \lambda(1-\hat{\delta}_D)\right] \\ &= (1-\lambda)Reg(\hat{\delta}_D) + \lambda Reg(1-\hat{\delta}_D). \end{aligned}$$

Hence, the nonlinear regret risk is

$$R_g(\hat{\delta}_R, P) = \mathbb{E}_{P^n}\left[g\left((1-\lambda)Reg(\hat{\delta}_D) + \lambda Reg(1-\hat{\delta}_D)\right)\right].$$

We now take the directional derivative (from above) of $R_g(\hat{\delta}_R, P)$ with respect to λ at $\lambda = 0$,

$$\begin{aligned} \left.\frac{\partial R_g(\hat{\delta}_R, P)}{\partial \lambda}\right|_{\lambda \searrow 0} &= \mathbb{E}_{P^n}\left[g'(Reg(\hat{\delta}_D))\left(Reg(1-\hat{\delta}_D) - Reg(\hat{\delta}_D)\right)\right] \\ &= \mathbb{E}_{P^n}\left[g'(Reg(\hat{\delta}_D))\tau(2\hat{\delta}_D - 1)\right] \\ &= \tau\left[g'(Reg(1))P^n(\hat{\delta}_D = 1) - g'(Reg(0))P^n(\hat{\delta}_D = 0)\right]. \end{aligned}$$

If $\tau > 0$, $Reg(1) = 0$ and $Reg(0) = \tau$, so

$$\begin{aligned} \left.\frac{\partial R_g(\hat{\delta}_R, P)}{\partial \lambda}\right|_{\lambda \searrow 0} &= \tau\left[g'(0)P^n(\hat{\delta}_D = 1) - g'(\tau)P^n(\hat{\delta}_D = 0)\right] \\ &= -\tau g'(\tau)P^n(\hat{\delta}_D = 0) \leq 0 \end{aligned}$$

where the second equality follows from the assumption that $g'(0) = 0$, and the inequality in the last line follows from $g'(\tau) > 0$ due to strict monotonicity of $g(\cdot)$. Since $P^n(\hat{\delta}_D = 0) > 0$ for some $P \in \mathcal{P}$ with $\tau \neq 0$, the inequality in the last line holds with a strict inequality at those P.

In the case where $\tau < 0$, we have $Reg(1) = -\tau$ and $Reg(0) = 0$. We hence have

$$\left.\frac{\partial R_g(\hat{\delta}_R, P)}{\partial \lambda}\right|_{\lambda \searrow 0} = \tau g'(-\tau) P^n(\hat{\delta}_D = 1) \leq 0,$$

where the inequality is strict at some $P \in \mathcal{P}$ with $\tau \neq 0$ due to the nondegeneracy of $\hat{\delta}_D$.

Having shown that $\left.\frac{\partial R_g(\hat{\delta}_R, P)}{\partial \lambda}\right|_{\lambda \searrow 0} \leq 0$ for any P and is strictly negative at some P, we conclude that there exists $\lambda > 0$ in a neighborhood of zero such that the resulting randomized treatment choice rule dominates $\hat{\delta}_D$. This completes the proof. $\square$

This result shows that if we consider the space of decision rules to comprise nondegenerate rules (i.e., $P^n(\hat{\delta} = 1) \in (0, 1)$), deterministic assignment rules $\hat{\delta}_D \in \{0, 1\}$ are inadmissible. Equivalently, the class of randomized decision rules is essentially complete and any admissible decision rule among the nondegenerate decision rules has to be a randomized rule.

This theorem contrasts sharply with the known admissibility of deterministic rules in more standard formulations of the treatment choice problem, where the (negative) expected welfare corresponds to the risk criterion in Wald's framework of statistical decision theory. For hypothesis testing problems with monotone likelihood ratio distributions, Karlin and Rubin (1956) show that the class of deterministic threshold rules is essentially complete (i.e., for an arbitrary decision rule $\hat{\delta}$ including randomized ones, there exists a deterministic threshold rule that performs as well as $\hat{\delta}$). As exploited in Hirano and Porter (2009) and Tetenov (2012), the essential completeness of deterministic threshold rules carries over to the treatment choice problem, implying that optimal rules among the deterministic threshold treatment assignment rules are admissible.

4 Finite sample optimality

Let $\mathcal{D}$ be the set of statistical decision rules under consideration. We measure the performance of a rule $\hat{\delta} \in \mathcal{D}$ by its nonlinear regret risk $R_g(\hat{\delta}, P)$, which depends on the true data generating process P. In this section we look at two optimality criteria and derive general results on optimal rules for these criteria. We illustrate the usefulness of our results using specific parametric models.

4.1 Bayes optimality

Definition 4.1 (Bayes nonlinear risk and the Bayes optimal rule)**.** Let π be a prior distribution on $P \in \mathcal{P}$. The Bayes nonlinear (regret) risk of $\hat{\delta}$ with respect to the prior π is

$$r_g(\hat{\delta}, \pi) := \int_{P \in \mathcal{P}} R_g(\hat{\delta}, P) d\pi(P).$$

A Bayes optimal rule $\hat{\delta}_\pi$ with respect to the prior π is such that

$$r_g(\hat{\delta}_\pi, \pi) = \inf_{\hat{\delta} \in \mathcal{D}} r_g(\hat{\delta}, \pi).$$

Moreover, we say that a prior distribution π is *least favorable* if $r_g(\pi, \hat{\delta}_\pi) \geq r_g(\pi', \hat{\delta}_{\pi'})$ for all prior distributions π'.

We now characterize the Bayes optimal rule for the Bayes nonlinear risk. It turns out that under mild restrictions on the nonlinear transformation g, the associated Bayes optimal rule is also randomized. To proceed, let $\pi(P|z_n)$ be the posterior distribution of P given a prior π and $Z_n = z_n$.

Theorem 4.1. *Suppose Assumption G holds, and the following conditions are true:*

(i) $g(Reg(\hat{\delta})) \geq 0$ *for all* $\hat{\delta} \in \mathcal{D}$ *and* $P \in \mathcal{P}$.

(ii) There exists some treatment rule $\tilde{\delta} \in \mathcal{D}$ *such that* $R_g(\tilde{\delta}, P)$ *is finite.*

(iii) For almost all $z_n \in \mathbf{Z}_n$, *the posterior distribution* $\pi(P|z_n)$ *puts nonzero probability mass on both* $\{P \in \mathcal{P} : \tau(P) > 0\}$ *and* $\{P \in \mathcal{P} : \tau(P) < 0\}$.

Then for almost all $z_n \in \mathbf{Z}_n$, the Bayes optimal rule $\hat{\delta}_\pi$ exists, is randomized, and satisfies

$$\int \left[\tau(P) g'\Big(\tau(P)(\mathbf{1}\{\tau(P) \geq 0\} - \hat{\delta}_\pi)\Big)\right] d\pi(P|z_n) = 0. \tag{4.1}$$

Proof. Under conditions (i) and (ii), it is straightforward to show (see, for example, Theorem 1.1 in Lehmann and Casella (1998)) that the Bayes optimal rule $\hat{\delta}_\pi$ is such that

$$\hat{\delta}_\pi \in \min_{\hat{\delta} \in [0,1]} \int g(Reg(\hat{\delta})) d\pi(P|z_n), \text{ for almost all } z_n \in \mathbf{Z}_n, \tag{4.2}$$

provided the solution of (4.2) exists for almost all $z_n \in \mathbf{Z}_n$.

Then the existence of $\hat{\delta}_\pi$ follows from continuity of the objective function (4.2) in $\hat{\delta} \in [0,1]$, which itself follows from the fact that g is continuously differentiable. (4.1) follows from the first order condition for (4.2). To see $0 < \hat{\delta}_\pi < 1$ for almost all $z_n \in \mathbf{Z}_n$, note $g'(\tau) > 0$ for all $\tau > 0$ because g is strictly increasing on $\mathbb{R}^+ \setminus \{0\}$ by Assumption G. Thus,

$$\begin{aligned}
&\left[\frac{\partial}{\partial \hat{\delta}} \int g(Reg(\hat{\delta})) d\pi(P|z_n)\right]_{\hat{\delta} \searrow 0} \\
&= -\int \left[\tau(P) g'(\tau(P)\mathbf{1}\{\tau(P) \geq 0\})\right] d\pi(P|z_n) \\
&= -\left[\int_{P \in \mathcal{P}: \tau(P) > 0} \left[\tau(P) g'(\tau(P))\right] d\pi(P|z_n) + g'(0) \int_{P \in \mathcal{P}: \tau(P) < 0} \tau(P) d\pi(P|z_n)\right] \\
&= -\left[\int_{P \in \mathcal{P}: \tau(P) > 0} \left[\tau(P) g'(\tau(P))\right] d\pi(P|z_n)\right] < 0,
\end{aligned}$$

where the last inequality follows from Assumption G and condition (iii). Similarly,

$$\begin{aligned}
&\left[\frac{\partial}{\partial \hat{\delta}} \int g(Reg(\hat{\delta})) d\pi(P|z_n)\right]_{\hat{\delta} \nearrow 1} \\
&= -\int \left[\tau(P) g'(\tau(P)(\mathbf{1}\{\tau(P) \geq 0\} - 1))\right] d\pi(P|z_n) \\
&= -\left[g'(0) \int_{P \in \mathcal{P}: \tau(P) > 0} \tau(P) d\pi(P|z_n) + \int_{P \in \mathcal{P}: \tau(P) < 0} \left[g'(-\tau(P))\tau(P)\right] d\pi(P|z_n)\right] \\
&= -\int_{P \in \mathcal{P}: \tau(P) < 0} \left[g'(-\tau(P))\tau(P)\right] d\pi(P|z_n) > 0.
\end{aligned}$$

The above calculations imply that we can always reduce $\int g(Reg(\hat{\delta}))d\pi(P|z_n)$ by moving $\hat{\delta}$ away from both 0 and 1 and toward an interior point. Therefore, $\hat{\delta}_\pi$ must be such that $0 < \hat{\delta}_\pi < 1$, for almost all $z_n \in \mathbf{Z}_n$. $\square$

Remark 4.1. In general, the Bayes optimal rule depends on the nonlinear transformation g and the model specification for P. The calculation of the posterior expectation in (4.1), which requires integration with respect to the posterior distribution of P, can be complicated. To gain further insight, consider the simple case where $g(r) = r^2$ and $P = P_\tau$ is parameterized by the one dimensional parameter $\tau \in \mathbb{R}$, where $\tau = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$ (for example, the outcome is normal with known variance). It follows that the prior distribution is indexed by τ and written as $\pi(\tau)$, and the Bayes optimal rule with respect to the Bayes mean square regret

$$r_{sq}(\hat{\delta}, \pi) := \int R_{sq}(\hat{\delta}, P_\tau)d\pi(\tau)$$

is characterized as

$$\int \left[\tau^2(\mathbf{1}\{\tau \geq 0\} - \hat{\delta}_\pi)\right] d\pi(\tau|z_n) = 0, \tag{4.3}$$

where $\pi(\tau|z_n)$ is the posterior distribution of τ given the prior $\pi(\tau)$ and data $Z_n = z_n$, with $Z_n \sim P_\tau^n$.

Further to this, if the prior $\pi(\tau)$ is supported on two symmetric points $\tau \in \{a, -a\}$ for some $a > 0$, it follows that

$$\hat{\delta}_\pi(z_n) = \frac{\int a^2 1\{\tau \geq 0\}d\pi(\tau|z_n)}{\int a^2 d\pi(\tau|z_n)} = \underbrace{\int 1\{\tau \geq 0\}d\pi(\tau|z_n)}_{\text{posterior probability that treatment effect is non-negative}},$$

which is the exact form of the posterior probability matching rule, as used by Thompson (1933). If the prior is not supported on two symmetric points, it holds that

$$\hat{\delta}_\pi(z_n) = \frac{\int \tau^2 1\{\tau \geq 0\}d\pi(\tau|z_n)}{\int \tau^2 d\pi(\tau|z_n)} = \underbrace{\int 1\{\tau \geq 0\}d\pi(\tau|z_n)}_{\text{posterior probability matching}} \underbrace{\frac{\int \tau^2 d\pi(\tau|z_n, \tau \geq 0)}{\int \tau^2 d\pi(\tau|z_n)}}_{\text{weight}}, \tag{4.4}$$

where $\pi(\tau|z_n, \tau \geq 0)$ denotes the posterior distribution of τ conditional on $\tau \geq 0$.

Thus, for the mean square regret, the Bayes optimal rule is a *tilted* version of the posterior probability matching rule.

Remark 4.2. In contrast, for the *linear* regret risk $R(\hat{\delta}, P)$, the Bayes optimal rule is

$$\hat{\delta}(z_n) = \begin{cases} \hat{\delta}(z_n) = 1, & \int \tau(P) d\pi(P|z_n) > 0, \\ \hat{\delta}(z_n) \in [0,1], & \int \tau(P) d\pi(P|z_n) = 0, \\ \hat{\delta}(z_n) = 0, & \int \tau(P) d\pi(P|z_n) < 0, \end{cases}$$

which is deterministic unless $\int \tau d\pi(\tau|z_n) = 0$.

We now provide a simple example for which we derive the finite sample Bayes optimal rule with respect to a flat prior. This example also sheds some light on the form of the Bayes optimal rule in large samples, which is discussed in Section 5.

Example 4.1 (Testing an innovation with normal outcome and mean square regret)**.** Let $g(r) = r^2$. Suppose the distribution of $Y(0)$ is known to the planner and without loss of generality, $\mathbb{E}[Y(0)] = 0$. Therefore, the planner only needs to learn $\mathbb{E}[Y(1)]$ and in the experimental design she allocates all units to the treatment. Let $\bar{Y}_1$ be the sample average of observed outcomes. Assume $\bar{Y}_1 \sim N(\tau, 1)$ is normally distributed with an unknown mean $\tau \in \mathbb{R}$ and a known variance normalized to one, with the likelihood function

$$f(\bar{y}_1|\tau) = \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1 - \tau)^2\right]\right), \forall \bar{y}_1 \in \mathbb{R}. \tag{4.5}$$

Proposition 4.1. *In Example 4.1, consider the uniform (improper) prior π_f on τ. Then the Bayes treatment rule with respect to the mean square regret is*

$$\hat{\delta}_{\pi_f}(\bar{Y}_1) = \Phi\left(\bar{Y}_1\right)\left[1 + \bar{Y}_1 \cdot \Psi(\bar{Y}_1)\right],$$

where $\Psi(x) := \frac{\phi(x)}{\Phi(x)(1+x^2)} > 0$ for any $x \in \mathbb{R}$, and where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cdf and pdf of a standard normal random variable, respectively.

Proposition 4.1 is a direct application of Theorem 4.1. Since the prior is flat, the 'posterior density' is proportional to the likelihood (4.5). The form of the Bayes optimal rule then follows (4.4). The Bayes optimal rule $\hat{\delta}_{\pi_f}$ is a product of two terms. The first term, $\Phi(\bar{Y}_1)$, is the posterior probability that the treatment effect is positive given the uninformative prior, and corresponds to the posterior probability matching rule. The second term, $(1 + \bar{Y}_1 \cdot \Psi(\bar{Y}_1))$, adjusts the first term upwards if $\bar{Y}_1 > 0$, and adjusts it downwards if $\bar{Y}_1 < 0$ (note that $\Psi(x) > 0$). Therefore, this Bayes optimal rule tilts the posterior probability matching rule and assigns treatment with a probability closer to zero or one. Also see Table 1 and Figure 5.1 for the magnitudes of the probability assignment of the Bayes optimal rule and posterior probability matching rule with respect to the uniform prior.

4.2 Minimax optimality

As an alternative to Bayes rule, this section studies minimax optimal rule for nonlinear regret risk.

Definition 4.2 (Minimax optimal rule)**.** A minimax optimal rule $\hat{\delta}^*$ is such that

$$\sup_{P\in\mathcal{P}} R_g(\hat{\delta}^*, P) = \inf_{\hat{\delta}\in\mathcal{D}} \sup_{P\in\mathcal{P}} R_g(\hat{\delta}, P).$$

The following proposition characterizes the minimax optimal rule as a Bayes rule under a least favorable prior.

Proposition 4.2 (Lehmann and Casella (1998))**.** *Suppose π is a distribution on P such that*

$$r_g(\hat{\delta}_\pi, \pi) = \sup_{P\in\mathcal{P}} R_g(\hat{\delta}_\pi, P).$$

Then: (i) $\hat{\delta}_\pi$ is minimax; (ii) π is least favorable.

Proposition 4.2 is a direct result of Lehmann and Casella (1998, Theorem 5.1.4). Using Proposition 4.2, we can attempt to find the minimax optimal rule by adopting a 'guess-and-verify' approach: guess a least favorable prior and derive its associated Bayes optimal rule; verify that the resulting Bayes nonlinear regret risk equals the worst frequentist nonlinear regret risk of the Bayes optimal rule. In general, it can still be difficult to guess the least favorable distribution. However, in many parametric models, the support of the least favorable distribution is often discrete and finite, or the minimax optimal rule has a constant frequentist risk across its parameter space. See, for example, Kempthorne (1987). This greatly simplifies the problem. We now demonstrate the minimax optimal rule for Example 4.1.

Theorem 4.2. *In Example 4.1, a finite sample minimax treatment rule is*

$$\hat{\delta}^*(\bar{Y}_1) = \frac{\exp\left(2\tau^*\bar{Y}_1\right)}{\exp\left(2\tau^*\bar{Y}_1\right) + 1},$$

where $\tau^* \approx 1.23$*, which solves*

$$\sup_{\tau\in[0,\infty)} \frac{1}{2}\tau^2\mathbb{E}\left[\frac{1}{\exp\left(2\tau\bar{Y}_1\right) + 1}\right], \tag{4.6}$$

or, equivalently, solves

$$\sup_{\tau\in[0,\infty)} \tau^2\mathbb{E}\left[\left(\frac{1}{\exp\left(2\tau\bar{Y}_1\right) + 1}\right)^2\right], \tag{4.7}$$

where the expectation is with respect to $\bar{Y}_1 \sim N(\tau, 1)$*. Moreover, a least favorable prior* π^* *on* τ *is a two-point prior such that* $\pi^*(\tau^*) = \pi^*(-\tau^*) = \frac{1}{2}$*.*

Remark 4.3. The minimax optimal rule is a simple logistic transformation of the sample mean and is straightforward to calculate. Moreover, the minimax optimal rule agrees with the posterior probability matching rule, i.e., the treatment probability equals the posterior probability that the treatment effect is positive with respect to the least favorable prior, which is supported on two symmetric points around zero. In this way, we justify the posterior probability matching rule in a static environment without multiple exploration phases.

Remark 4.4. The treatment probability of our suggested minimax optimal rule is always between zero and one. As such, our rule naturally provides a degree of confidence on the hypothesis that $\tau > 0$, i.e., the treatment effect is positive. Given data from the single phase experiment, a larger value of $\hat{\delta}^*$ means we are more confident that $\tau > 0$, while a smaller value of $\hat{\delta}^*$ signals much less evidence supporting a positive treatment effect. Consider a scenario where $\bar{Y}_1$ is only slightly larger than zero. While the empirical success rule will assign the innovation treatment to everyone in the population, our rule only allocates a fraction (more than 50 percent) of the population to the innovation. The allocation probability can be interpreted as a measure of confidence in the performance of a treatment. Such an interpretation would not be possible if the optimal rule is deterministic rather than randomized.

Remark 4.5. On a more technical note, the proof of Theorem 4.2 relies on some new techniques that are absent from the existing treatment choice literature. For the mean regret criterion, deterministic threshold rules form an essential complete class, so the minimax optimal rule with respect to mean regret can be found by directly minimizing the worst-case regret with respect to the threshold, without figuring out a least favorable prior. See for example, Stoye (2009); Tetenov (2012). However, for mean square regret, deterministic rules are inadmissible, and it becomes essential to find the probability distribution of a least favorable prior. To find a least favorable prior, we adopt the 'guess-and-verify' approach. By observing the form of the mean square regret, we guess that a least favorable prior π^* is such that

$$\pi^*(\tau) = \frac{1}{2}, \pi^*(-\tau) = \frac{1}{2},$$

for some $0 < \tau < \infty$. Within this set of candidate least favorable priors π^*_τ indexed by τ, Theorem 4.1 implies the Bayes optimal rules admit the form $\hat{\delta}_{\pi^*_\tau}(\bar{Y}_1) = \frac{\exp\left(2\tau\bar{Y}_1\right)}{\exp\left(2\tau\bar{Y}_1\right)+1}$. Furthermore, $r_{sq}(\hat{\delta}_{\pi^*_\tau}, \pi^*_\tau)$ follows the form in (4.6), and is equivalent to the form in (4.7). Then, our guess for the least favorable prior is

$$\pi^*(\tau^*) = \frac{1}{2}, \pi^*(-\tau^*) = \frac{1}{2},$$

where τ^* solves (4.6) or (4.7). With this guess of the least favorable prior, we further establish that the following condition holds:

Condition 1. $r_{sq}(\hat{\delta}^*, \pi^*) = \sup_{\tau\in[0,\infty)} R_{sq}\left(\hat{\delta}^*, P_\tau\right)$.

The left-hand side of Condition 1 is the Bayes mean square regret of $\hat{\delta}^*$ with respect to our hypothesized least favorable prior π^*, and the right-hand side of Condition 1 is the worst mean square regret of $\hat{\delta}^*$. Thus, Proposition 4.2 implies that $\hat{\delta}^*$ is a minimax optimal rule and π^* is least favorable. See also Figure 4.1 for a graphical illustration. The full proof of Theorem 4.2 is left to Appendix B.

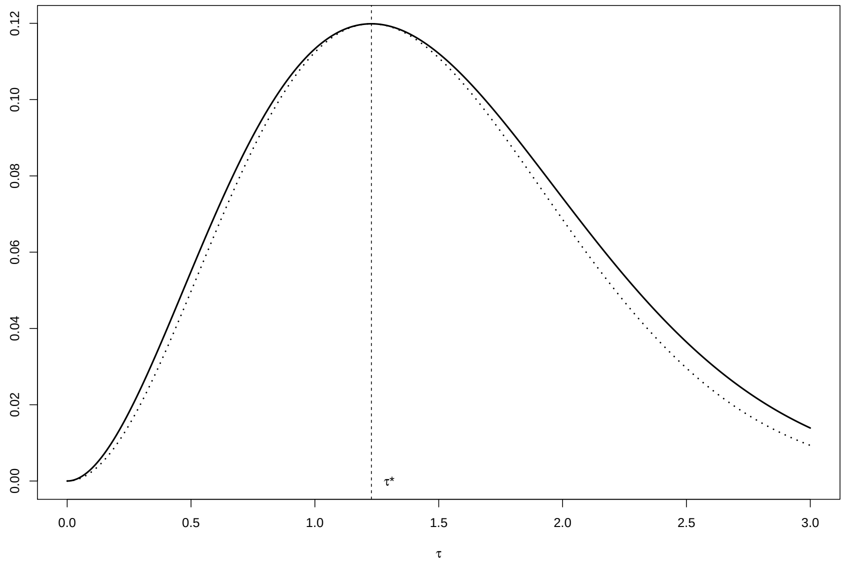

Figure 4.1: Illustration of Condition 1 for Theorem 4.2. The dotted line is $r_{sq}(\hat{\hat{\delta}}_{\pi_\tau^*}, \pi_\tau^*)$ as a function of τ; the solid line is $R_{sq}\left(\hat{\hat{\delta}}^*, P_\tau\right)$ as a function of τ.

5 Asymptotic optimality with mean square regret

In this section we derive asymptotically optimal rules via the limit experiment framework (Le Cam, 1986), following the approach taken by Hirano and Porter (2009). We first consider a local parametrization of the statistical model P so that, in large samples, the treatment choice problem is equivalent to a simpler problem in a Gaussian limit experiment. Then, we examine and normalize our nonlinear regret in the limit, and find the corresponding optimal treatment rule. A feasible and asymptotically optimal treatment rule also follows if there exists an efficient estimator of the parameters in the original statistical model P. For a review, see Hirano and Porter (2020).

5.1 Limit experiments

For simplicity, we focus on regular parametric models of $P \in \mathcal{P}$ with mean square regret R_{sq}. Semiparametric models and other nonlinear regret criteria can also be considered, albeit necessitating more technical analysis. Without loss of generality, consider a case where the distribution of $Y(0)$ is known and the mean of $Y(0)$ is zero. Suppose now the distribution of $Y(1)$, denoted by P, is parameterized by a finite dimensional parameter $\theta \in \Theta \subseteq \mathbb{R}^k$. Hence, the population average treatment effect is

$$\tau(\theta) = \int z dP_\theta(z).$$

Data $Z_n = \{Z_i\}_{i=1}^n$ is independently and identically drawn from P_θ. In particular, $Z_i \sim P_\theta$, where $Z_i \in \mathbf{Z}$ and $\mathbf{Z}$ is the support of Z_i. We now imagine a sequence of experiments $\mathcal{E}_n := \{P_\theta^n, \theta \in \Theta\}$ in which the sample size n grows. Let $\theta_0 \in \Theta$ satisfy $\tau(\theta_0) = 0$. We consider a sequence of local alternative parameters of the form $\theta_0 + \frac{h}{\sqrt{n}}$, $h \in \mathbb{R}^k$, the most challenging case in which to determine the optimal treatment rule, even in large samples.

Assumption DQM (Differentiability in Quadratic Mean)**.** There exists a function $s : \mathbf{Z} \to \mathbb{R}^k$ such that

$$\int \left[dP_{\theta_0+h}^{\frac{1}{2}}(z) - dP_{\theta_0}^{\frac{1}{2}}(z) - \frac{1}{2} h' s(z) dP_{\theta_0}^{\frac{1}{2}}(z) \right]^2 = o(\|h\|^2), \text{ as } h \to 0,$$

and $I_0 := \mathbb{E}_{\theta_0}[ss']$ is nonsingular.

Assumption DQM is a standard assumption in the limit experiment framework (e.g., Van der Vaart, 1998). The function s can usually be interpreted as the derivative of the loglikelihood function so that I_0 is the Fisher information under P_{θ_0}.

Assumption C (Convergence)**.** A sequence of treatment rules $\hat{\delta}_n$ in the experiments $\mathcal{E}_n$ is such that $\beta_n(h, 1) := \mathbb{E}_{P^n_{\theta_0 + \frac{h}{\sqrt{n}}}}[\hat{\delta}_n] \to \beta(h, 1)$ and $\beta_n(h, 2) := \mathbb{E}_{P^n_{\theta_0 + \frac{h}{\sqrt{n}}}}[(\hat{\delta}_n)^2] \to \beta(h, 2)$ for every h as $n \to \infty$.

Compared to mean regret criterion, our mean square regret additionally depends on the second moment of decision rules. Thus, Assumption C assumes convergence of both first and second moments of decision rules, differing from Hirano and Porter (2009), who only look at convergence of the first moment of decision rules. Under Assumptions DQM and C, we first establish the following result that allows us to simplify the original treatment problem to a Gaussian experiment in large samples.

Proposition 5.1 (Van der Vaart (1998))**.** *Suppose $\mathcal{E}_n$ satisfy Assumption DQM and a sequence of treatment rules $\hat{\delta}_n$ in $\mathcal{E}_n$ satisfy Assumption C. Then there exists a* *function* $\hat{\delta} : \mathbb{R}^k \to [0,1]$ *such that for every* $h \in \mathbb{R}^k$,

$$\beta(h,1) = \int \hat{\delta}(\Delta) dN(\Delta|h, I_0^{-1}), \ \beta(h,2) = \int \left(\hat{\delta}(\Delta)\right)^2 dN(\Delta|h, I_0^{-1}),$$

where $N(\Delta|h, I_0^{-1})$ *is a multivariate normal distribution with mean* h *and variance* I_0^{-1}.

Proposition 5.1 is a special case of Van der Vaart (1998, Theorem 13.1 and Theorem 7.10) applied to the mean square regret setup, following Hirano and Porter (2009, Proposition 3.1). To use Proposition 5.1, note for any treatment rule $\hat{\delta}_n$ in the experiments $\mathcal{E}_n$, the mean square regret is

$$\mathbb{E}_{P^n_{\theta_0+\frac{h}{\sqrt{n}}}} \left[\tau\left(\theta_0 + \frac{h}{\sqrt{n}}\right)^2 \left(1\left\{\tau\left(\theta_0 + \frac{h}{\sqrt{n}}\right) \geq 0\right\} - \hat{\delta}_n\right)^2 \right],$$

which depends on $\hat{\delta}_n$ only through $\mathbb{E}_{P^n_{\theta_0+\frac{h}{\sqrt{n}}}}[\hat{\delta}_n]$ and $\mathbb{E}_{P^n_{\theta_0+\frac{h}{\sqrt{n}}}}[\hat{\delta}_n^2]$, to which we can apply Proposition 5.1. Thus, in terms of the mean square regret, any converging sequence of treatment rules is matched by some treatment rule in a simpler Gaussian experiment with unknown mean h and known variance I_0^{-1}.

Let $\dot{\tau}$ be the partial derivative of $\tau(\theta)$ at θ_0. Since $\tau(\theta_0) = 0$, it follows that $\sqrt{n}\tau\left(\theta_0 + \frac{h}{\sqrt{n}}\right) \to \dot{\tau}'h$ as $n \to \infty$. Thus, for any rule δ,

$$\sqrt{n} Reg\left(\delta, \left(\theta_0 + \frac{h}{\sqrt{n}}\right)\right) \to \dot{\tau}'h\left[1\left\{\dot{\tau}'h \geq 0\right\} - \delta\right] := Reg_\infty(\delta, h),$$

and $n\left[Reg\left(\delta, \left(\theta_0 + \frac{h}{\sqrt{n}}\right)\right)\right]^2 \to \left(Reg_\infty(\delta,h)\right)^2$ as $n \to \infty$. Hence, normalizing by n, for any converging rule $\hat{\delta}_n$ in the sense of Proposition 5.1, we define the corresponding limit mean square regret as

$$\begin{aligned} R^\infty_{sq}(\hat{\delta}, h) &:= \int \left(Reg_\infty(\hat{\delta}(\Delta), h)\right)^2 dN(\Delta|h, I_0^{-1}) \qquad (5.1) \\ &= \mathbb{E}_{\Delta\sim N(h, I_0^{-1})}\left[Reg_\infty(\hat{\delta}(\Delta), h)\right]^2. \end{aligned}$$

With (5.1) as the mean square regret in the limit experiment, we can apply our finite sample results in Section 4 and derive a feasible and asymptotically optimal treatment rule via an efficient estimator of the parameters.

5.2 Feasible and asymptotically optimal rules

We first present results in terms of minimax optimality. Denote $\overset{h}{\rightsquigarrow}$ as convergence in distribution under the sequence of probability measures $P^n_{\theta_0+\frac{h}{\sqrt{n}}}$. Define $\sigma_\tau := \sqrt{\dot{\tau}' I_0^{-1} \dot{\tau}}$ to be the standard deviation of $\dot{\tau}'\Delta$, where $\Delta \sim N(h, I_0^{-1})$.

Theorem 5.1. *Suppose Proposition 5.1 holds, $\tau(\theta_0) = 0$, and $\tau(\theta)$ is differentiable at θ_0.*

(i) The minimax optimal rule in the limit experiment is

$$\hat{\delta}^*(\Delta) = \frac{\exp\left(\frac{2\tau^*}{\sigma_\tau}\dot{\tau}'\Delta\right)}{\exp\left(\frac{2\tau^*}{\sigma_\tau}\dot{\tau}'\Delta\right) + 1},$$

where $\tau^* \approx 1.23$, and which solves (4.6).

(ii) If, in addition, there exists a best regular estimator $\hat{\theta}$ such that

$$\sqrt{n}\left(\hat{\theta} - \theta_0 - \frac{h}{\sqrt{n}}\right) \overset{h}{\rightsquigarrow} N(0, I_0^{-1}), \text{ for all } h \in \mathbb{R}^k, \tag{5.2}$$

and there exists some estimator $\hat{\sigma}_\tau \xrightarrow{p} \sigma_\tau$ under θ_0, the feasible treatment rule

$$\hat{\delta}^*_F(Z_n) = \frac{\exp\left(\frac{2\tau^*}{\hat{\sigma}_\tau}\sqrt{n}\tau(\hat{\theta})\right)}{\exp\left(\frac{2\tau^*}{\hat{\sigma}_\tau}\sqrt{n}\tau(\hat{\theta})\right) + 1}$$

is locally asymptotically minimax optimal in terms of mean square regret:

$$\sup_J \liminf_{n\to\infty} \sup_{h\in J} nR_{sq}(\hat{\delta}^*_F, \theta_0 + \frac{h}{\sqrt{n}}) = \inf_{\hat{\delta}\in\mathcal{D}} \sup_J \liminf_{n\to\infty} \sup_{h\in J} nR_{sq}(\hat{\delta}, \theta_0 + \frac{h}{\sqrt{n}}),$$

where J is a finite subset of $\mathbb{R}^k$ and $\mathcal{D}$ is the set of all decision rules that satisfy Assumption C (slightly abusing notation).

Theorem 5.1 extends our finite sample results to a large sample setting. Given a regular parametric model, the maximum likelihood estimator (MLE) usually satisfies

(5.2). Thus, Theorem 5.1 suggests a simple way to construct an asymptotically minimax optimal rule in terms of mean square regret: estimate the parameters of P_θ via MLE, calculate a t-statistic for the mean, and then carry out a simple logit transformation for the t-statistic. This rule is always fractional and very easy to implement for practitioners. We expect that our result can also be extended to regular semiparametric models.

Next, we derive a feasible rule that is locally asymptotically Bayes optimal. Let $\pi(\theta)$ be a positive and continuous prior density on Θ (slightly abusing notation). For a treatment rule $\hat{\delta}_n$ that satisfies Assumption C, the normalized Bayes mean square regret is

$$nr_{sq}(\hat{\delta}_n, \pi) = \int nR_{sq}(\hat{\delta}_n, \theta_0 + \frac{h}{\sqrt{n}})\pi(\theta_0 + \frac{h}{\sqrt{n}})dh.$$

We define the Bayes mean square regret in the limit experiment when $n \to \infty$ as

$$r_{sq}^{\infty}(\hat{\delta}) := \pi(\theta_0) \int R_{sq}^{\infty}(\hat{\delta}, h)dh.$$

That is, as the Bayes mean square regret with respect to an uninformative prior. Then we can apply Theorem 4.1 to derive the Bayes optimal rule for the limit experiment. Given an MLE estimate of the parameters in P_θ, Theorem 5.2 further implies that a feasible and asymptotically optimal Bayes rule also follows with a simple transformation of the t-statistic for the mean.

Theorem 5.2. *Suppose Proposition 5.1 holds, $\tau(\theta_0) = 0$ and $\tau(\theta)$ is differentiable at θ_0. Let $\pi(\theta)$ be the density of a prior distribution on Θ that is continuous and positive at θ_0.*

(i) The Bayes optimal rule in terms of mean square regret in the limit experiment is

$$\hat{\delta}_B\left(\varDelta\right) = \Phi\left(\frac{\dot{\tau}'\varDelta}{\sigma_\tau}\right)\left(1 + \frac{\dot{\tau}'\varDelta}{\sigma_\tau}\Psi\left(\frac{\dot{\tau}'\varDelta}{\sigma_\tau}\right)\right). \tag{5.3}$$

That is, $r_{sq}^{\infty}(\hat{\delta}_B) = \inf_{\delta \in \mathcal{D}_\infty} r_{sq}^{\infty}(\delta)$, where $\mathcal{D}_\infty$ is the set of all treatment rules in the $N(h, I_0^{-1})$ limit experiment.

(ii) If, in addition, there exists a best regular estimator $\hat{\theta}$ such that

$$\sqrt{n}\left(\hat{\theta} - \theta_0 - \frac{h}{\sqrt{n}}\right) \overset{h}{\rightsquigarrow} N(0, I_0^{-1}), \text{ for all } h \in \mathbb{R}^k,$$

and there exists some estimator $\hat{\sigma}_\tau \xrightarrow{p} \sigma_\tau$ under θ_0, the feasible treatment rule

$$\hat{\delta}_{B,F}(Z_n) = \Phi\left(\frac{\sqrt{n}\tau(\hat{\theta})}{\hat{\sigma}_\tau}\right)\left[1 + \frac{\sqrt{n}\tau(\hat{\theta})}{\hat{\sigma}_\tau}\Psi\left(\frac{\sqrt{n}\tau(\hat{\theta})}{\hat{\sigma}_\tau}\right)\right]$$

is locally asymptotically Bayes optimal, i.e.,

$$\lim_{n\to\infty} nr_{sq}(\hat{\delta}_{B,F}, \pi) = \inf_{\hat{\delta}\in\mathcal{D}} \liminf_{n\to\infty} nr_{sq}(\hat{\delta}, \pi).$$

In the limit, the Bayes optimal rule is a *tilted* posterior probability matching rule with respect to the uninformative prior. Compared to the posterior probability matching rule, the Bayes optimal rule assigns treatment with a probability closer to zero or one. Compared to the limit minimax optimal rule, the Bayes optimal rule also assigns treatment with a probability close to zero or one. This contrasts with the case of *linear* regret risk, where it is known that the Bayes optimal and minimax optimal rules are the same empirical success rule. See Figure 5.1 and Table 1 for various rules in a Gaussian limit experiment with unit variance. It can be seen that all three randomized rules approach one as $\bar{Y}_1$ gets large. For sufficiently large positive values of $\bar{Y}_1$ (e.g., 2.33), the Bayes and minimax optimal rules are to effectively treat everyone. Even with a modest value of $\bar{Y}_1 = 0.84$, the Bayes optimal rule recommends a probability of treatment of 0.94, which is quite high when compared with the corresponding probability of 0.8 recommended by the posterior probability matching rule. Figures 5.2, 5.3 and 5.4 present the mean square regret, mean regret and standard deviation of regret of the optimal rules in the same Gaussian limit experiment with unit variance. We make several observations: firstly, although they admit different forms, our Bayes optimal and minimax optimal rules in the limit experiment exhibit a similar performance in terms of the mean square regret (Figure 5.2); secondly, the ES rule is minimax optimal in terms of mean regret (Figure 5.3), but its excessive variance (Figure 5.4) in those states where mean regret is high implies that it is not optimal in terms of mean square regret.

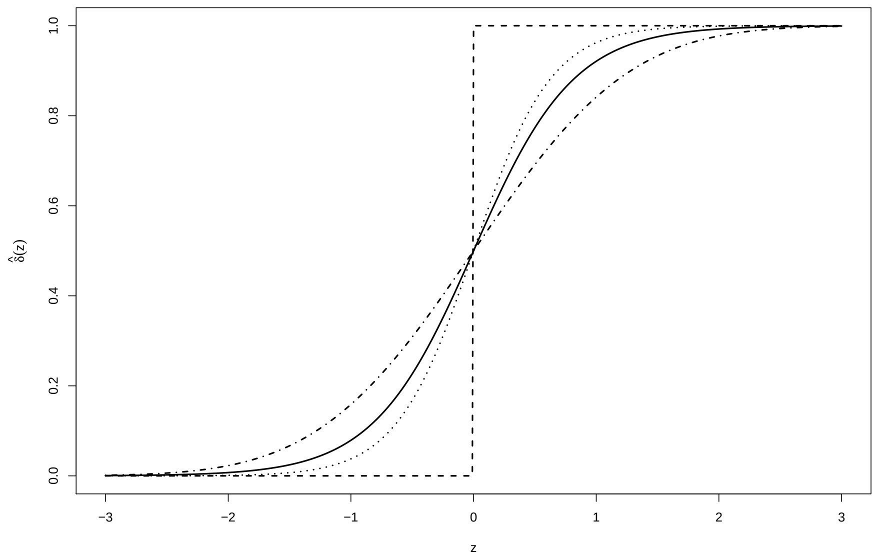

Figure 5.1: Optimal rules in the Gaussian limit experiment with unit variance. Solid line: minimax optimal rule for mean square regret; Dotted line: Bayes optimal rule for mean square regret; Dot-dashed line: posterior probability matching rule with respect to a flat prior; Dashed line: Empirical success (ES) rule.

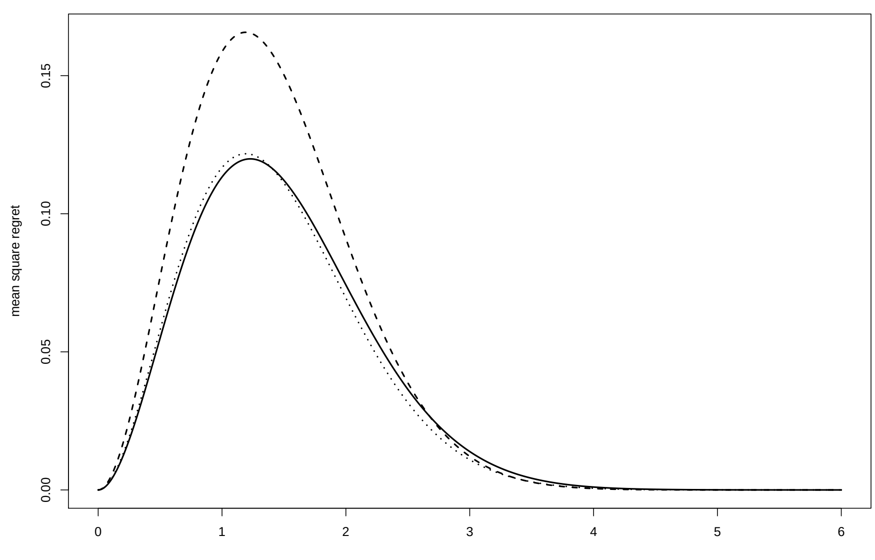

Figure 5.2: Mean square regret in the Gaussian limit experiment with unit variance. Solid line: minimax optimal rule; Dotted line: Bayes optimal rule with respect to a flat prior; Dashed line: ES rule.

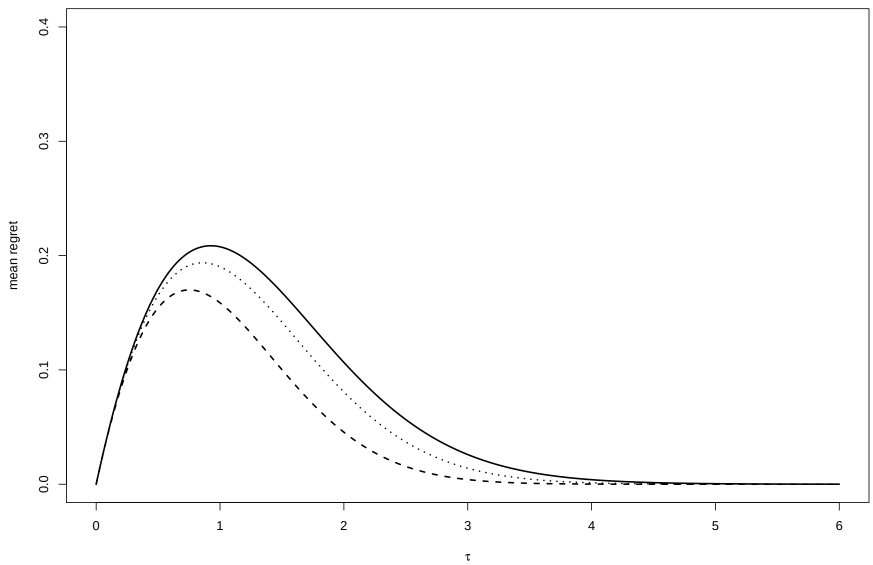

Figure 5.3: Mean regret in the Gaussian limit experiment with unit variance. Solid line: minimax optimal rule; Dotted line: Bayes optimal rule with respect to a flat prior; Dashed line: ES rule.

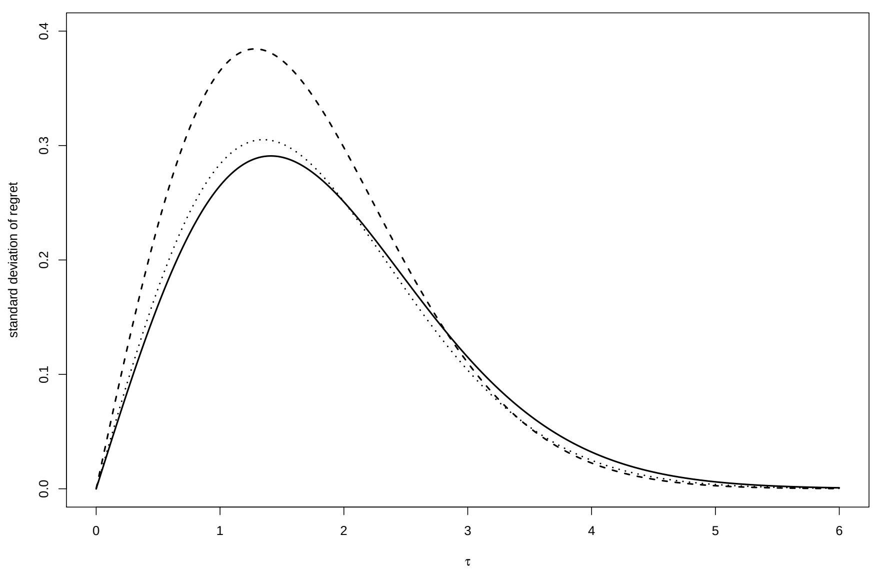

Figure 5.4: Standard deviation of regret in the Gaussian limit experiment with unit variance. Solid line: minimax optimal rule; Dotted line: Bayes optimal rule with respect to a flat prior; Dashed line: ES rule.

$\bar{Y}_1$	Minimax optimal rule	Bayes optimal rule	Posterior probability matching rule (flat prior)	ES rule
0	0.5	0.5	0.5	$[0, 1]$
0.2533	0.6507	0.6920	0.6	1
0.5244	0.7838	0.8430	0.7	1
0.8416	0.8877	0.9379	0.8	1
1.2816	0.9588	0.9851	0.9	1
1.6449	0.9827	0.9958	0.95	1
2.3263	0.9967	0.9997	0.99	1

Table 1: Treatment assignment probabilities in the Gaussian limit experiment with unit variance

6 Sample size calculations

In practice, the planner often has a preference for deterministic rules like the empirical success (ES) rule or the hypothesis testing (HT) rule, and calculates what is a sufficient sample size based on these deterministic rules. In this section we discuss the implications for the efficiency loss in terms of mean square regret if deterministic rules were implemented instead of our proposed minimax optimal rules. Compared to our minimax optimal rule, these deterministic rules often require significantly more data and thus are much less efficient. For instance, to guarantee the same mean square regret with our minimax optimal rule, ES rule and HT rule demand around 40% and 1100% more observations, respectively. A similar discussion can be had for the Bayes optimal rule, but we omit this for brevity.

Consider the Gaussian experiment in Example 4.1, but suppose now $\bar{Y}_1 \sim N(\tau, \frac{\sigma^2}{n})$ is the sample average calculated from experimental data with a sample size of n and known variance $\sigma^2 > 0$. In this case the minimax optimal rule in terms of mean square regret is

$$\hat{\delta}^*(\bar{Y}_1) = \frac{\exp(2\tau^* \frac{\sqrt{n}}{\sigma} \bar{Y}_1)}{\exp(2\tau^* \frac{\sqrt{n}}{\sigma} \bar{Y}_1) + 1},$$

where τ^* solves (4.6). Given each $\varepsilon > 0$, we can select n such that

$$\sqrt{\sup_{\tau \in [0, \infty)} R_{sq}(\hat{\delta}^*, P_\tau)} \leq \varepsilon,$$

i.e., the square root of the worst case mean square regret does not exceed ε. The worst case mean square regret can be calculated as

$$\sup_{\tau \in [0,\infty)} R_{sq}(\hat{\delta}^*, P_\tau) = \left(\frac{\sigma^2}{n}\right) R^*_{sq}(1),$$

where $R^*_{sq}(1) \approx 0.1199$ is the worst case mean square regret of the minimax optimal rule in Example 4.1. Thus, the worst case mean square regret shrinks to zero at a rate of $\frac{1}{n}$. In practice, we can choose ε to be proportional to σ, e.g., 0.01σ, so that the square root of the worst case mean square regret does not exceed 1% of the standard deviation.

Comparison with the ES rule

Manski and Tetenov (2016) choose a sufficient sample size for the ES rule via the $\varepsilon-$optimal approach: a policy $\hat{\delta}$ is $\varepsilon-$optimal if, for all states of the world,

$$W(\delta^*) - \mathbb{E}_{P^n}[W(\hat{\delta})] \leq \varepsilon,$$

where δ^* is the infeasible optimal treatment rule or, equivalently,

$$\mathbb{E}_{P^n}[Reg(\hat{\delta})] \leq \varepsilon, \tag{6.1}$$

for all states of the world. Given our Gaussian experiment $\bar{Y}_1 \sim N(\tau, \frac{\sigma^2}{n})$, the worst case mean regret of the ES rule $\widehat{\delta}_{ES} = \mathbf{1}\{\bar{Y}_1 \geq 0\}$ can be calculated exactly as

$$\sup_{\tau \in [0,\infty)} \tau \left(1 - \Phi\left(\frac{\sqrt{n}\tau}{\sigma}\right)\right) = \frac{\sigma}{\sqrt{n}} \sup_{\tau \in [0,\infty)} \tau \left(1 - \Phi\left(\tau\right)\right) = 0.1700 \frac{\sigma}{\sqrt{n}}.$$

If the planner has a preference for the ES rule and decides to choose the sample size so that (6.1) holds with some $\varepsilon > 0$, then the sample size should be at least

$$n_{ES} = 0.0289 \frac{\sigma^2}{\varepsilon^2}.$$

The worst case mean square regret of the ES rule, however, is

$$\sup_{\tau \in [0,\infty)} R_{sq}(\widehat{\delta}_{ES}, P_\tau) = \frac{\sigma^2}{n} R^{ES}_{sq}(1),$$

where $R_{sq}^{ES}(1) = \sup_{\tau \in [0,\infty)} \tau^2 \mathbb{E}_{\bar{Y}_1 \sim N(\tau,1)} \left[\left(1 - \mathbf{1}\{\bar{Y}_1 \geq 0\}\right)^2 \right] \approx 0.1657$. Hence, at n_{ES}, the worst case mean square regret of $\widehat{\delta}_{ES}$ is $\frac{\sigma^2}{n_{ES}} 0.1657 = 5.7355 \frac{\varepsilon^2}{\sigma^2}$. If, instead, the planner uses our minimax optimal rule, she only needs a sample size of $n^* = 0.0209 \frac{\sigma^2}{\varepsilon^2}$ for the worst case mean square regret not to exceed $5.7355 \frac{\varepsilon^2}{\sigma^2}$. Thus, to guarantee the same worst case mean square regret, the ES rule requires nearly 40% more observations than our minimax optimal rule.

Comparison with the HT rule

Practitioners who prefer the HT rule often select sample size by balancing Type I and II errors. In the Gaussian experiment $\bar{Y}_1 \sim N(\tau, \frac{\sigma^2}{n})$, if the planner uses a size α HT rule

$$\hat{\delta}_{HT} = \mathbf{1}\left\{ \frac{\sqrt{n}\bar{Y}_1}{\sigma} \geq z_{(1-\alpha)} \right\},$$

where $z_{(1-\alpha)}$ is the $(1-\alpha)$ quantile of a standard normal, then it is common for her to select sample size so that the power of the test is at least β, i.e., under the alternative $\tau > 0$, the probability of rejection is

$$\Pr\left\{ \frac{\bar{Y}_1 - \tau}{\frac{\sigma}{\sqrt{n}}} > z_{(1-\alpha)} - \frac{\tau}{\frac{\sigma}{\sqrt{n}}} \right\} = \beta.$$

Then the sample size should be at least

$$n_{HT} = \frac{\sigma^2}{\tau^2} (z_{(1-\alpha)} - z_{(1-\beta)})^2.$$

At this n_{HT}, we can also calculate the worst case mean square regret of the HT rule, which is approximately $\frac{\tau^2}{(z_{(1-\alpha)} - z_{(1-\beta)})^2} 1.4458$. However, at this n_{HT}, the worst case mean square regret of our minimax rule is only $0.1199 \frac{\sigma^2}{n_{HT}} = 0.1199 \frac{\tau^2}{(z_{(1-\alpha)} - z_{(1-\beta)})^2}$. That is to say, with the same sample size n_{HT}, our minimax optimal rule guarantees that the worst case mean square regret is only around 8.3% of the corresponding value for the HT rule. Equivalently, to guarantee the same worst case mean square regret, the HT rule requires around 11 times more observations than our minimax optimal rule.

7 Conclusions

Our paper proposes a novel approach to measure the performance of statistical decision rules by considering a nonlinear transformation of regret. Such a shift of criterion can incorporate other features of the regret distribution (e.g., second- and higher-order moments) into the decision-making process, and yields optimal rules that are drastically different from the existing literature. For a large class of nonlinear transformations, optimal rules are randomized, allocating only a fraction of the population to the treatment. For the mean square regret criterion, we also derive Bayes optimal and minimax optimal rules both for finite Gaussian samples and in asymptotic limit experiments. These rules have a simple and insightful form, and can be calculated easily by practitioners. Since our rules are always fractional, they naturally provide a degree of confidence in the performance of the treatment. Implementing our rules has the additional benefit of getting more data from randomized experiments that can be helpful for the inference of treatment effect, which would not be possible if deterministic rules were implemented.

Our treatment of the statistical decision via a nonlinear transformation of the regret suggests that there might exist an underlying notion of *regret aversion* of the planner toward sampling uncertainty. Our approach implicitly assumes that the planner dislikes large welfare loss when their acts depend on the source of uncertainty from random sampling. In that respect, the two-fold approach of Klibanoff et al. (2005); Denti and Pomatto (2022) on treating model uncertainty could be related. However, to the best of our knowledge, there seems no formal axiomatization of our nonlinear regret risk in the current literature of decision theory. How the selection of the nonlinear transformation is mapped to the preference of the planner also remains elusive. We leave these questions for future research.

A Comparison with Manski and Tetenov (2007)

In this section we clarify that our approach of treatment choice with nonlinear regret criteria fundamentally differs from the approach of risk averse welfare criteria taken by Manski and Tetenov (2007). To elaborate, let $f(\cdot) : \mathbb{R} \to \mathbb{R}$ be a concave function. A concave transformation of $W(\hat{\delta})$ is $f(W(\hat{\delta}))$. For the concave transformation f, the regret of treatment rule $\hat{\delta}$ defined in terms of *nonlinear welfare* is

$$Reg^f(\hat{\delta}) = f(W(\delta^*)) - f(W(\hat{\delta})) = f(\mu_0 + \delta^*\tau) - f(\mu_0 + \hat{\delta}\tau).$$

In contrast, our paper considers a nonlinear (possibly convex) transformation of regret measured in terms of the original welfare:

$$g(Reg(\hat{\delta})) = g\left((\mu_0 + \delta^*\tau) - (\mu_0 + \hat{\delta}\tau)\right),$$

where $g : \mathbb{R}^+ \to \mathbb{R}$ is a nonlinear function that does not depend on $\hat{\delta}$, μ_0 or μ_1. In other words, the loss function in Manski and Tetenov (2007) is $Reg^f(\hat{\delta})$ while in our paper the loss function is $g(Reg(\hat{\delta}))$.

Proposition A.1. *Consider the following statement*

$$\mathbb{E}_{P^n}[Reg^f(\hat{\delta})] = \mathbb{E}_{P^n}[g(Reg(\hat{\delta}))], \text{ for all } \hat{\delta}, \mu_0 \text{ and } \mu_1. \tag{A.1}$$

Then (A.1) holds for some concave function $f(\cdot) : \mathbb{R} \to \mathbb{R}$ and some function $g(\cdot) : \mathbb{R}^+ \to \mathbb{R}$ if and only if $f(x) = ax + b$ and $g(x) = ax$ for some constants a and b.

Proof. The *if* part is straightforward to show. We focus on the *only if* part. Let $\mathbb{F}(\cdot) := \mathbb{E}_{P^n}[f(\cdot)]$ and $\mathbb{G}(\cdot) := \mathbb{E}_{P^n}[g(\cdot)]$. Since convexity and concavity are preserved under the expectation operator, it holds that $\mathbb{F}(\cdot)$ is concave too. Then, by assumption,

$$\mathbb{F}(\mu_0 + \delta^*\tau) - \mathbb{F}(\mu_0 + \hat{\delta}\tau) = \mathbb{G}((\mu_0 + \delta^*\tau) - (\mu_0 + \hat{\delta}\tau)) \tag{A.2}$$

for all μ_0, μ_1 and $\hat{\delta}$, implying

$$\mathbb{F}(x) - \mathbb{F}(y) = \mathbb{G}(x - y), \forall x \geq y. \tag{A.3}$$

Fixing $y = 0$, (A.3) implies

$$\mathbb{F}(x) - \mathbb{F}(0) = \mathbb{G}(x), \forall x \geq 0. \tag{A.4}$$

Since $\mathbb{F}$ is concave, (A.4) implies $\mathbb{G}(x)$ is concave as well for all $x \geq 0$. Conversely, fixing $x = 0$, (A.3) implies

$$\mathbb{F}(0) - \mathbb{F}(y) = \mathbb{G}\left(-y\right), \forall y \leq 0, \tag{A.5}$$

or, equivalently,

$$\mathbb{F}(0) - \mathbb{F}(-x) = \mathbb{G}(x), \forall x \geq 0. \tag{A.6}$$

Since $\mathbb{F}$ is concave, (A.6) implies $\mathbb{G}(x)$ is convex for all $x \geq 0$. Thus, $\mathbb{G}(x)$ must be both concave and convex for $x \geq 0$, implying $\mathbb{G}(x)$ is an affine function for all $x \geq 0$. This implies g is affine and admits $g(x) = ax+t$ for some constants a and t. Since $\mathbb{G}(x)$ is affine, (A.4) implies that $\mathbb{F}(x)$ is affine for $x \geq 0$ and $f(x) = ax+t+\mathbb{F}(0)$ for $x \geq 0$. Furthermore, for all $y \leq 0$, (A.5) implies $\mathbb{F}(y) = \mathbb{F}(0) - \mathbb{G}(-y)$, i.e., $\mathbb{F}(y)$ is affine for $y \leq 0$ as well, and $f(y) = ay - t + \mathbb{F}(0)$ for $y \leq 0$. At $x = 0$, $t + \mathbb{F}(0) = -t + \mathbb{F}(0)$ must hold, implying $t = 0$. Thus, $g(x) = ax$ and $f(x) = ax + \mathbb{F}(0)$ must hold or, equivalently, $f(x) = ax + b$ for some constants a and b. □

Given a concave transformation f of the welfare considered in Manski and Tetenov (2007), Proposition A.1 shows that we cannot find a nonlinear transformation g of the original regret such that the regret of nonlinear welfare defined in Manski and Tetenov (2007) equals our nonlinear regret risk for all rules and all states of the world. The results of Proposition A.1 can be extended in several ways. Firstly, Proposition A.2 shows that even if we consider either f or g to be convex, or we restrict the domain of f to be positive, the results of Proposition A.1 continue to hold. For instance, suppose $g(r) = r^2$ and a nonlinear welfare transformation f were to exist so that

$$\mathbb{E}_{P^n}[f(W(\delta^*)) - f(W(\hat{\delta}))] = \mathbb{E}_{P^n}[(W(\delta^*) - W(\hat{\delta}))^2], \forall \hat{\delta}, \mu_0, \text{and } \mu_1. \tag{A.7}$$

Proposition A.2 shows that such an f does not exist. Secondly, one might argue that even though (A.1) does not hold, the risks of the two approaches could be affine transformations of each other, so that the optimal rules are the same. In Proposition A.3, we show that even in such a scenario, both f and g also have to be affine. Our approach in introducing nonlinear $g(\cdot)$ is inherently different from that of Manski and Tetenov (2007).

Proposition A.2. *(i) (A.1) holds for some convex function $f(\cdot) : \mathbb{R} \to \mathbb{R}$ and some function $g(\cdot) : \mathbb{R}^+ \to \mathbb{R}$ if and only if $f(x) = ax + b$ and $g(x) = ax$ for* *some constants a and b.*

(ii) (A.1) holds for some function $f(\cdot):\mathbb{R}\to\mathbb{R}$ and some convex function $g(\cdot):\mathbb{R}^+\to\mathbb{R}$ if and only if $f(x)=ax+b$ and $g(x)=ax$ for some constants a and b.

(iii) (A.1) holds for some concave function $f(\cdot):\mathcal{C}\to\mathbb{R}$, where $\mathcal{C}\subseteq\mathbb{R}^+$ is a compact interval, and some function $g(\cdot):\mathbb{R}^+\to\mathbb{R}$ if and only if $f(x)=ax+b$ and $g(x)=ax$ for some constants a and b.

Proof. Statement (i): the proof is the same as that of Proposition A.1.

Statement (ii): We only show the *only if* part. Note (A.3) still holds. Fix $T\in\mathbb{R}$. It holds that

$$\mathbb{F}(T)-\mathbb{F}(y)=\mathbb{G}(T-y),\forall y\leq T,$$

or, equivalently, that

$$\mathbb{F}(y)=\mathbb{F}(T)-\mathbb{G}(T-y),\forall y\leq T. \tag{A.8}$$

Since $\mathbb{G}$ is convex, (A.8) implies $\mathbb{F}(y)$ is concave for all $\forall y\leq T$. Letting $T\to\infty$ implies $\mathbb{F}(y)$ is a concave function in $\mathbb{R}$. The rest of the proof then follows that of Proposition A.1.

Statement (iii). We only show the *only if* part. Without loss of generality, suppose $f(\cdot):[0,1]\to\mathbb{R}$. Note (A.3) still holds for all $1\geq x\geq y\geq 0$, implying

$$\mathbb{F}(x)-\mathbb{F}(0.5)=\mathbb{G}(x-0.5),\forall 0.5\leq x\leq 1, \tag{A.9}$$

i.e., $\mathbb{G}$ is concave on the interval $[0,0.5]$. Conversely, (A.3) also implies

$$\mathbb{F}(0.5)-\mathbb{F}(y)=\mathbb{G}(0.5-y),\forall 0\leq y\leq 0.5,$$

which means $\mathbb{G}$ is convex on $[0,0.5]$. Thus, $\mathbb{G}$ must be affine on $[0,0.5]$, and $g(x)=ax+t$ for some constants a and t, for each $0\leq x\leq 0.5$. Further note $\mathbb{F}(x)-\mathbb{F}(0)=\mathbb{G}(x),\forall 0\leq x\leq 1$. Thus, combining this with (A.9), we find

$$\mathbb{G}(x)=\mathbb{F}(x)-\mathbb{F}(0)=\mathbb{G}(x-0.5)+\mathbb{F}(0.5)-\mathbb{F}(0),\forall 0.5\leq x\leq 1.$$

In particular, at $x = 0.5$, $\mathbb{G}(0.5) = \mathbb{G}(0) + \mathbb{F}(0.5) - \mathbb{F}(0)$. Hence,

$$\mathbb{G}(x) = \mathbb{G}(x - 0.5) + \mathbb{G}(0.5) - \mathbb{G}(0), \forall 0.5 \leq x \leq 1,$$

implying $\mathbb{G}(x)$ is affine and $g(x) = ax + t$ in $[0.5, 1]$ as well. Thus, $g(x) = ax + t$ for $0 \leq x \leq 1$. But then plugging $x = 0.5$ into (A.9) implies $t = 0$. Then, it is easy to see that $f(x) = ax + b$ for some constant b. □

Proposition A.3. *Let $A > 0$ and $B \in \mathbb{R}$ be some constants. Consider the following statement*

$$\mathbb{E}_{P^n}[Reg^f(\hat{\delta})] = A\mathbb{E}_{P^n}[g(Reg(\hat{\delta}))] + B, \text{ for all } \hat{\delta},\ \mu_0 \text{ and } \mu_1. \tag{A.10}$$

Then (A.10) holds for some concave function $f(\cdot) : \mathbb{R} \to \mathbb{R}$, some function $g(\cdot) : \mathbb{R}^+ \to \mathbb{R}$ and some constants $A > 0$ and B if and only if $f(x) = ax + b$ and $g(x) = \frac{a}{A}x - \frac{B}{A}$ for some constants a, b, $A > 0$ and B.

Proof. To see the *only if* part, let $\tilde{g}(x) = Ag(x) + B$. (A.10) implies

$$\mathbb{E}_{P^n}[Reg^f(\hat{\delta})] = \mathbb{E}_{P^n}[\tilde{g}(Reg(\hat{\delta}))], \text{ for all } \hat{\delta}, \mu_0 \text{ and } \mu_1.$$

Applying the results of Proposition A.1 yields that $Ag(x) + B = ax$ and $f(x) = ax + b$ for some constants a, b, $A > 0$ and B. That is, $g(x) = \frac{a}{A}x - \frac{B}{A}$. The *if* part is straightforward to show and omitted. □

B Proofs of main results

Proof of Proposition 4.1

The proof is similar to the proof of statement (i) of Theorem 5.2 and thus omitted.

Proof of Theorem 4.2

We split the proof into three steps by adopting the 'guess-and-verify' approach.

Step 1: Guess a least favorable prior. Note the worst case mean square regret of a minimax optimal rule is

$$\sup_{\tau \in \mathbb{R}} R_{sq}(\hat{\delta}^*, P_\tau), \tag{B.1}$$

where

$$\begin{aligned} R_{sq}(\hat{\delta}^*, P_\tau) &= \tau^2 \mathbb{E}[1\{\tau \geq 0\} - \hat{\delta}^*(\bar{Y}_1)]^2 \\ &= \begin{cases} \tau^2 \mathbb{E}\left[1 - \hat{\delta}^*(\bar{Y}_1)\right]^2 & \tau > 0, \\ 0 & \tau = 0, \\ \tau^2 \mathbb{E}\left[\hat{\delta}^*(\bar{Y}_1)\right]^2 & \tau < 0. \end{cases} \end{aligned}$$

By Lemma C.1, the support of the solution of (B.1) never contains zero. In Lemma C.2, we show that the support of the solution of (B.1) must be symmetric, i.e., if the support of the solution of (B.1) contains τ for some $0 < \tau < \infty$, it must also contain $-\tau$. Therefore, we conjecture that the least favorable prior π^* is two-point supported. Moreover, Lemma C.3 shows that for a symmetric two-point prior to be least favorable, each point is equally likely to be realised. Thus, our guess for the least favorable prior π^* is such that

$$\pi^*(\tau) = \frac{1}{2}, \pi^*(-\tau) = \frac{1}{2}, \text{ for some } 0 < \tau < \infty.$$

Step 2: Derive the Bayes optimal rule associated with the hypothesized least favorable prior. For each $0 < \tau < \infty$, let $\hat{\delta}_{\pi^*_\tau}$ be the Bayes optimal rule with respect to the two-point symmetric prior

$$\pi^*_\tau(\tau) = \frac{1}{2} \text{ and } \pi^*_\tau(-\tau) = \frac{1}{2}.$$

Within the above set of candidate least favorable priors, we show: (1) the Bayes optimal rules admit the form $\hat{\delta}_{\pi^*_\tau}(\bar{Y}_1) = \frac{\exp\left(2\tau\bar{Y}_1\right)}{\exp\left(2\tau\bar{Y}_1\right)+1}$; (2) $r_{sq}(\hat{\delta}_{\pi^*_\tau}, \pi^*_\tau)$ follows the form in (4.6), and is equivalent to the form in (4.7). Thus, our guess for the least favorable prior is

$$\pi^*(\tau^*) = \frac{1}{2}, \pi^*(-\tau^*) = \frac{1}{2},$$

where τ^* solves (4.6) or (4.7).

Indeed, the functional form of $\hat{\delta}_{\pi_\tau^*}$ is derived by applying Theorem 4.1,

$$\hat{\delta}_{\pi_\tau^*}(\bar{y}_1) = \frac{\int \tau^2 1\{\tau \geq 0\} d\pi_\tau^*(\tau|\bar{y}_1)}{\int \tau^2 d\pi_\tau^*(\tau|\bar{y}_1)},$$

where $\pi_\tau^*(\tau|\bar{y}_1)$ is the posterior distribution of π_τ^* conditional on $\bar{Y}_1 = \bar{y}_1$ and admits:

$$\pi_\tau^*\{\tau|\bar{y}_1\} = \frac{\frac{1}{2} f\{\bar{y}_1|\tau\}}{f\{\bar{y}_1\}} \quad \text{and} \quad \pi_\tau^*\{-\tau|\bar{y}_1\} = \frac{\frac{1}{2} f\{\bar{y}_1| - \tau\}}{f\{\bar{y}_1\}},$$

where $f\{\bar{y}_1|\tau\}$ is the likelihood of τ, $f\{\bar{y}_1| - \tau\}$ is the likelihood of $-\tau$, and $f\{\bar{y}_1\}$ is the marginal density of $\bar{Y}_1$. Note

$$f\{\bar{y}_1|\tau\} = \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1 - \tau)^2\right]\right) > 0,$$
$$f\{\bar{y}_1| - \tau\} = \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1 + \tau)^2\right]\right) > 0.$$

It follows that

$$\begin{aligned}
\hat{\delta}_{\pi_\tau^*}(\bar{y}_1) &= \frac{f\{\bar{y}_1|\tau\}}{f\{\bar{y}_1|\tau\} + f\{-\tau|\bar{y}_1\}} \\
&= \frac{\exp\left(-\frac{1}{2}\left[(\bar{y}_1 - \tau)^2\right]\right)}{\exp\left(-\frac{1}{2}\left[(\bar{y}_1 - \tau)^2\right]\right) + \exp\left(-\frac{1}{2}\left[(\bar{y}_1 + \tau)^2\right]\right)} \\
&= \frac{\exp(2\tau\bar{y}_1)}{\exp(2\tau\bar{y}_1) + 1}.
\end{aligned}$$

Therefore, the Bayes mean square regret of $\hat{\delta}_{\pi_\tau^*}$ admits the form in (4.6):

$$
\begin{aligned}
r_{sq}(\hat{\delta}_{\pi_\tau^*}, \pi_\tau^*) &= \frac{1}{2}\tau^2 \int \left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\} + f\{\bar{y}_1|-\tau\}} \right)^2 f\{\bar{y}_1|\tau\} d\bar{y}_1 \\
&\quad + \frac{1}{2}\tau^2 \int \left(\frac{f\{\bar{y}_1|\tau\}}{f\{\bar{y}_1|\tau\} + f\{\bar{y}_1|-\tau\}} \right)^2 f\{\bar{y}_1|-\tau\} d\bar{y}_1 \\
&= \frac{1}{2}\tau^2 \int \frac{f\{\bar{y}_1|-\tau\} f\{\bar{y}_1|\tau\}}{[f\{\bar{y}_1|\tau\} + f\{\bar{y}_1|-\tau\}]} d\bar{y}_1 \\
&= \frac{1}{2}\tau^2 \int \frac{\sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1-\tau)^2\right]\right)}{[\exp(2\tau\bar{y}_1) + 1]} d\bar{y}_1 \\
&= \frac{1}{2}\tau^2 \mathbb{E}\left[\frac{1}{\exp\left(2\tau\bar{Y}_1\right) + 1} \right].
\end{aligned}
$$

Since $\tau > 0$ and $R_{sq}(\hat{\delta}_{\pi_\tau^*}, P_\tau) = \tau^2 \mathbb{E}[1 - \hat{\delta}_{\pi_\tau^*}(\bar{Y}_1)]^2$, we see that (4.6) is equivalent to (4.7):

$$
\begin{aligned}
R_{sq}(\hat{\delta}_{\pi_\tau^*}, P_\tau) &= \tau^2 \int \left[\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\} + f\{\bar{y}_1|-\tau\}} \right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 \\
&= \tau^2 \int \left[\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\} + f\{\bar{y}_1|-\tau\}} \right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1-\tau)^2\right]\right) d\bar{y}_1 \\
&= \tau^2 \int \left[\frac{1}{\exp(2\tau\bar{y}_1) + 1} \right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1-\tau)^2\right]\right) d\bar{y}_1 \\
&= \tau^2 \mathbb{E}\left[\frac{1}{\exp\left(2\tau\bar{Y}_1\right) + 1} \right]^2,
\end{aligned}
$$

and by a change of variables,

$$
\begin{aligned}
R_{sq}(\hat{\delta}_{\pi_\tau^*}, P_{-\tau}) &= \tau^2 \int \left(\frac{f\{\bar{y}_1|\tau\}}{f\{\bar{y}_1|\tau\} + f\{\bar{y}_1|-\tau\}} \right)^2 f\{\bar{y}_1|-\tau\} d\bar{y}_1 \\
&= R_{sq}(\hat{\delta}_{\pi_\tau^*}, P_\tau).
\end{aligned}
$$

Step 3: For our guess of the least favorable prior, Lemma C.5 further establishes that Condition 1 holds. Thus, Proposition 4.2 implies that $\hat{\delta}^*$ is indeed a minimax optimal rule and the two-point prior $\pi^*(\tau^*) = \pi^*(-\tau^*) = \frac{1}{2}$ is indeed least favorable.

Proof of Theorem 5.1

Proof of statement (i)

Let $\hat{\delta}^*$ be a minimax optimal rule in the limit experiment. That is, $\hat{\delta}^*$ solves

$$\inf_{\hat{\delta}} \sup_{h} R^{\infty}_{sq}(\hat{\delta}, h) := R^*.$$

Following Hirano and Porter (2009), consider slicing the parameter space of h in the following way: define

$$h_1(b, h_0) = h_0 + \frac{b}{\dot{\tau}' I_0^{-1} \dot{\tau}} I_0^{-1} \dot{\tau},$$

where h_0 is such that $\dot{\tau}' h_0 = 0$ (without loss of generality) and $b \in \mathbb{R}$. Hence,

$$\dot{\tau}' h_1(b, h_0) = b.$$

Note for each $\hat{\delta} \in [0, 1]$, the limit regret $Reg_{\infty}(\hat{\delta}, h)$ only depends on h through $\dot{\tau}' h$. Thus, we can consider treatment rules of the form

$$\hat{\delta}(\varDelta) = \hat{\delta}\left(\dot{\tau}' \varDelta\right) := \hat{\delta}_{\tau}(\varDelta_{\tau}),$$

where $\varDelta_{\tau} := \dot{\tau}' \varDelta \sim N(\dot{\tau}' h, \dot{\tau}' I_0^{-1} \dot{\tau})$. Let $\hat{\delta}^*_{\tau}$ solve the simpler minimax exercise

$$\inf_{\hat{\delta}_{\tau}} \sup_{b} R^{\infty}_{sq}(\hat{\delta}_{\tau}, h_1(b, 0))$$

among rules of form $\hat{\delta}_{\tau}(\varDelta_{\tau})$. It follows by Lemma D.1 that $\hat{\delta}^*_{\tau}$ is a minimax optimal rule. Define

$$R^{\infty}_{sq}(\hat{\delta}_{\tau}, b) := b^2 \mathbb{E}_{\varDelta_{\tau} \sim N(b, \dot{\tau}' I_0^{-1} \dot{\tau})} \left[1\left\{b \geq 0\right\} - \hat{\delta}_{\tau}\left(\varDelta_{\tau}\right)\right]^2.$$

Lemma D.2 shows that $\hat{\delta}^*_{\tau}$ can be found by solving $\inf_{\hat{\delta}_{\tau}} \sup_b R^{\infty}_{sq}(\hat{\delta}_{\tau}, b)$, and Lemma D.3 establishes the the form of $\hat{\delta}^*_{\tau}$, which is a minimax optimal rule in the limit experiment.

Proof of statement (ii)

By Hirano and Porter (Lemma 3, 2009), $\sqrt{n}\frac{\tau(\hat{\theta})}{\hat{\sigma}_\tau} \overset{h}{\rightsquigarrow} N(\frac{\dot{\tau}'h}{\sigma_\tau}, 1)$. Furthermore, using the continuous mapping theorem,

$$\hat{\delta}_F^*(Z_n) \overset{h}{\rightsquigarrow} \frac{\exp\left(2\tau^* N(\frac{\dot{\tau}'h}{\sigma_\tau}, 1)\right)}{\exp\left(2\tau^* N(\frac{\dot{\tau}'h}{\sigma_\tau}, 1)\right) + 1}.$$

Therefore, $\hat{\delta}_F^*$ is matched with $\hat{\delta}^*$ in the limit experiment in the sense of Proposition 5.1. The desired conclusion follows via a similar argument to that in Hirano and Porter (Lemma 4, 2009).

Proof of Theorem 5.2

Proof of statement (i)

Applying Theorem 4.1 to the limit Bayes mean square criterion r_{sq}^∞ yields

$$\hat{\delta}_B(\Delta) = \frac{\int (\dot{\tau}'h)^2 \left(\mathbf{1}\left\{\dot{\tau}'h \geq 0\right\}\right) d\pi(h|\Delta)}{\int (\dot{\tau}'h)^2 \, d\pi(h|\Delta)}.$$

Notice in the limit experiment, h has a flat prior. It follows that the posterior distribution $\pi(h|\Delta)$ is proportional to a normal distribution with mean Δ and variance I_0^{-1}. Then

$$\begin{aligned}
\hat{\delta}_B(\Delta) &= \frac{\int (\dot{\tau}'h)^2 \left(\mathbf{1}\left\{\dot{\tau}'h \geq 0\right\}\right) dN(h|\Delta, I_0^{-1})}{\int (\dot{\tau}'h)^2 \, dN(h|\Delta, I_0^{-1})} \\
&= \frac{\int s^2 \left(\mathbf{1}\left\{s \geq 0\right\}\right) dN(s|\dot{\tau}'\Delta, \dot{\tau}' I_0^{-1}\dot{\tau})}{\int s^2 dN(s|\dot{\tau}'\Delta, \dot{\tau}' I_0^{-1}\dot{\tau})} \\
&= \frac{\int_{s\geq 0} s^2 dN(s|\dot{\tau}'\Delta, \sigma_\tau^2)}{\int s^2 dN(s|\dot{\tau}'\Delta, \sigma_\tau^2)} \\
&= \int \mathbf{1}\left\{s \geq 0\right\} dN(s|\dot{\tau}'\Delta, \sigma_\tau^2) \frac{\int s^2 dN(s|\dot{\tau}'\Delta, \sigma_\tau^2, S \geq 0)}{\int s^2 dN(s|\dot{\tau}'\Delta, \sigma_\tau^2)},
\end{aligned}$$

where $\int s^2 dN(s|\dot{\tau}'\Delta, \sigma_\tau^2, S \geq 0)$ denotes the conditional expectation of a normal random variable S with mean $\dot{\tau}'\Delta$ and variance σ_τ^2 given $S \geq 0$. By the properties of the normal distribution and truncated normal distribution,

$$\int \mathbf{1}\{s \geq 0\}\, dN(s|\dot{\tau}'\Delta, \sigma_\tau^2) = \Phi(\frac{\dot{\tau}'\Delta}{\sigma_\tau}),$$

$$\int s^2 dN(s|\dot{\tau}'\Delta, \sigma_\tau^2) = \sigma_\tau^2 + (\dot{\tau}'\Delta)^2\,,$$

$$\int s^2 dN(s|\dot{\tau}'\Delta, \sigma_\tau^2, S \geq 0) = \sigma_\tau^2 \left(1 - \frac{\dot{\tau}'\Delta}{\sigma_\tau}\frac{\phi(\frac{\dot{\tau}'\Delta}{\sigma_\tau})}{\Phi(\frac{\dot{\tau}'\Delta}{\sigma_\tau})} - \frac{\phi^2(\frac{\dot{\tau}'\Delta}{\sigma_\tau})}{\Phi^2(\frac{\dot{\tau}'\Delta}{\sigma_\tau})}\right) + \left((\dot{\tau}'\Delta) + \sigma_\tau \frac{\phi(\frac{\dot{\tau}'\Delta}{\sigma_\tau})}{\Phi(\frac{\dot{\tau}'\Delta}{\sigma_\tau})}\right)^2.$$

Statement (i) follows.

Proof of statement (ii)

Similar to the argument in the proof of statement (ii) of Theorem 5.1, $\hat{\delta}^*_{B,F}$ is matched with $\hat{\delta}_B$ in the limit experiment in the sense of Proposition 5.1. The conclusion follows via a similar argument to that in Hirano and Porter (Lemma 1, 2009).

References

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

Bhattacharya, D. and Dupas, P. (2012). Inferring welfare maximizing treatment assignment under budget constraints. *Journal of Econometrics*, 167(1):168–196.

Cassidy, R. and Manski, C. F. (2019). Tuberculosis diagnosis and treatment under uncertainty. *Proceedings of the National Academy of Sciences*, 116(46):22990–22997.

Chamberlain, G. (2011). Bayesian aspects of treatment choice. In Geweke, J., Koop, G., and van Dijk, H., editors, *The Oxford Handbook of Bayesian Econometrics*, pages 11–39. Oxford University Press.

Christensen, T., Moon, H., and Schorfheide, F. (2020). Robust forecasting. arXiv:2011.03153 [econ.EM], https://doi.org/10.48550/arXiv.2011.03153.

Dehejia (2005). Program evaluation as a decision problem. *Journal of Econometrics*, 125:141–173.

Denti, T. and Pomatto, L. (2022). Model and predictive uncertainty: A foundation for smooth ambiguity preferences. *Econometrica*, 90(2):551–584.

Giacomini, R., Kitagawa, T., and Read, M. (2021). Robust Bayesian analysis for econometrics. Federal Board of Chicago Working Paper, No. 2021-11, https://doi.org/10.21033/wp-2021-11.

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

Hirano, K. and Porter, J. R. (2020). Asymptotic analysis of statistical decision rules in econometrics. In Durlauf, S. N., Hansen, L. P., Heckman, J. J., and Matzkin, R. L., editors, *Handbook of Econometrics, Volume 7A*, volume 7 of *Handbook of Econometrics*, pages 283–354. Elsevier.

Ishihara, T. and Kitagawa, T. (2021). Evidence aggregation for treatment choice. arXiv:2108.06473 [econ.EM], https://doi.org/10.48550/arXiv.2108.06473.

Karlin, S. (1957). Pólya type distributions, ii. *The Annals of Mathematical Statistics*, 28(2):281–308.

Karlin, S. and Rubin, H. (1956). The theory of decision procedures for distributions with monotone likelihood ratio. *The Annals of Mathematical Statistics*, pages 272–299.

Kempthorne, P. J. (1987). Numerical specification of discrete least favorable prior distributions. *SIAM Journal on Scientific and Statistical Computing*, 8(2):171–184.

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

Kitagawa, T. and Tetenov, A. (2021). Equality-minded treatment choice. *Journal of Business & Economic Statistics*, 39(2):561–574.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

Le Cam, L. (1986). *Asymptotic methods in statistical decision theory*. Springer Science & Business Media.

Lehmann, E. L. and Casella, G. (1998). *Theory of point estimation*. Springer Science & Business Media, second edition.

Manski, C. F. (1988). Ordinal utility models of decision making under uncertainty. *Theory and Decision*, 25(1):79–104.

Manski, C. F. (2000). Identification problems and decisions under ambiguity: empirical analysis of treatment response and normative analysis of treatment choice. *Journal of Econometrics*, 95:415–442.

Manski, C. F. (2002). Treatment choice under ambiguity induced by inferential problems. *Journal of Statistical Planning and Inference*, 105(1):67–82.

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

Manski, C. F. (2005). *Social choice with partial knowledge of treatment response*. Princeton University Press.

Manski, C. F. (2007a). *Identification for prediction and decision*. Harvard University Press.

Manski, C. F. (2007b). Minimax-regret treatment choice with missing outcome data. *Journal of Econometrics*, 139:105–115.

Manski, C. F. (2009). The 2009 Lawrence R. Klein lecture: Diversified treatment under ambiguity. *International Economic Review*, 50(4):1013–1041.

Manski, C. F. (2011). Actualist rationality. *Theory and Decision*, 71(2):195–210.

Manski, C. F. (2013). *Public policy in an uncertain world: analysis and decisions*. Harvard University Press.

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

Manski, C. F. (2021b). Probabilistic prediction for binary treatment choice: with focus on personalized medicine. Technical report, National Bureau of Economic Research.

Manski, C. F. (2022). Identification and statistical decision theory. *arXiv preprint arXiv:2204.11318*.

Manski, C. F. and Tetenov, A. (2007). Admissible treatment rules for a risk-averse planner with experimental data on an innovation. *Journal of Statistical Planning and Inference*, 137(6):1998–2010.

Manski, C. F. and Tetenov, A. (2014). The quantile performance of statistical treatment rules using hypothesis tests to allocate a population to two treatments. Technical report, cemmap working paper (CWP44/14). https://doi.org/10.1920/wp.cem.2014.4414.

Manski, C. F. and Tetenov, A. (2016). Sufficient trial size to inform clinical practice. *Proceedings of the National Academy of Sciences*, 113(38):10518–10523.

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

Savage, L. (1951). The theory of statistical decision. *Journal of the American Statistical Association*, 46:55–67.

Schlag, K. H. (2006). ELEVEN - tests needed for a recommendation. Technical report, European University Institute Working Paper, ECO No. 2006/2. https://cadmus.eui.eu/bitstream/handle/1814/3937/ECO2006-2.pdf.

Stoye, J. (2009). Minimax regret treatment choice with finite samples. *Journal of Econometrics*, 151(1):70–81.

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

Tetenov, A. (2012). Statistical treatment choice based on asymmetric minimax regret criteria. *Journal of Econometrics*, 166(1):157–165.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294.

Van der Vaart, A. W. (1998). *Asymptotic statistics*. Cambridge university press.

Wald, A. (1950). *Statistical Decision Functions*. Wiley, New York.

Yata, K. (2021). Optimal decision rules under partial identification. arXiv:2111.04926 [econ.EM], https://doi.org/10.48550/arXiv.2111.04926.

Online Appendix

C Lemmas for Section 4

Lemma C.1. $\tau = 0$ is never a solution of (B.1).

Proof. Note at $\tau = 0$, the squared regret is 0. Suppose it is a solution of (B.1), then it must hold that

$$\mathbb{E}\left[1 - \hat{\delta}^*(\bar{Y}_1)\right]^2 = 0 \text{ and } \mathbb{E}\left[\hat{\delta}^*(\bar{Y}_1)\right]^2 = 0. \tag{C.1}$$

Since $\hat{\delta}^*(\bar{y}_1) \in [0,1]$ for all $\bar{y}_1$, (C.1) implies

$$1 - \hat{\delta}^*(\bar{y}_1) = 0 \text{ and } \hat{\delta}^*(\bar{y}_1) = 0, \text{ for all } \bar{y}_1 \text{ a.s.},$$

which cannot be true. This implies $\tau = 0$ is never a solution of (B.1). $\square$

Lemma C.2. The solution of (B.1) is symmetric, i.e., if some $\tau \in (0, \infty)$ solves (B.1), then it also holds that $-\tau$ solves (B.1).

Proof. Suppose $\tau \in (0, \infty)$ solves (B.1) but $-\tau$ does not. Note the mean square regret of $\hat{\delta}^*$ at τ is

$$\begin{aligned} R_{sq}(\hat{\delta}^*, P_\tau) &= \tau^2 \mathbb{E}\left[1 - \hat{\delta}^*(\bar{Y}_1)\right]^2 \\ &= \tau^2 \int \left[1 - \hat{\delta}^*(\bar{y}_1)\right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1 - \tau)^2\right]\right) d\bar{y}_1, \end{aligned}$$

while the mean square regret at $-\tau$ is

$$\begin{aligned} R_{sq}(\hat{\delta}^*, P_{-\tau}) =& \tau^2 \mathbb{E}\left[\hat{\delta}^*(\bar{Y}_1)\right]^2 \\ =& \tau^2 \int \left[\hat{\delta}^*(\bar{y}_1)\right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1 + \tau)^2\right]\right) d\bar{y}_1 \\ =& \tau^2 \int \left[\hat{\delta}^*(-\tilde{y}_1)\right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\tau - \tilde{y}_1)^2\right]\right) d\tilde{y}_1 \\ =& \tau^2 \int \left[\hat{\delta}^*(-\bar{y}_1)\right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\tau - \bar{y}_1)^2\right]\right) d\bar{y}_1, \end{aligned}$$

where the third equality uses the change of variable $\tilde{y}_1 = -\bar{y}_1$, and the fourth equality changes the variable of integration from $\tilde{y}_1$ to $\bar{y}_1$.

If τ solves (B.1) but $-\tau$ does not, then there must exist some $\bar{y}_1 \in \mathbb{R}$ such that $1 - \hat{\delta}^*(\bar{y}_1) \neq \hat{\delta}^*(-\bar{y}_1)$. Let

$$\mathbf{S} = \left\{ \bar{y}_1 \in \mathbb{R} : 1 - \hat{\delta}^*(\bar{y}_1) \neq \hat{\delta}^*(-\bar{y}_1) \right\}$$

be the collection of all $\bar{y}_1$ such that $1 - \hat{\delta}^*(\bar{y}_1) \neq \hat{\delta}^*(-\bar{y}_1)$.[1] The contribution of the elements of $\mathbf{S}$ to the mean square regret at τ is

$$\tau^2 \int_{\mathbf{S}} \left[1 - \hat{\delta}^*(\bar{y}_1)\right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1 - \tau)^2\right]\right) d\bar{y}_1$$

while the contribution of the elements in $\mathbf{S}$ to the mean square regret at $-\tau$ is

$$\tau^2 \int_{\mathbf{S}} \left[\hat{\delta}^*(-\bar{y}_1)\right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\tau - \bar{y}_1)^2\right]\right) d\bar{y}_1.$$

Since τ solves (B.1) but not $-\tau$, it holds that

$$\begin{aligned} &\tau^2 \int_{\mathbf{S}} \left[1 - \hat{\delta}^*(\bar{y}_1)\right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\bar{y}_1 - \tau)^2\right]\right) d\bar{y}_1 \\ >&\tau^2 \int_{\mathbf{S}} \left[\hat{\delta}^*(-\bar{y}_1)\right]^2 \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}\left[(\tau - \bar{y}_1)^2\right]\right) d\bar{y}_1. \end{aligned} \tag{C.2}$$

If (C.2) holds though, we can strictly reduce the mean square regret for τ by switching to an alternative policy $\bar{\delta}$, where

$$\bar{\delta}(\bar{y}_1) = \begin{cases} \hat{\delta}^*(\bar{y}_1) & \text{if } \bar{y}_1 \notin \mathbf{S}, \\ 1 - \hat{\delta}^*(-\bar{y}_1) & \text{if } \bar{y}_1 \in \mathbf{S}. \end{cases}$$

[1] Notice the set $\mathbf{S}$ must be symmetric, that is, if

$$1 - \hat{\delta}^*(\bar{y}_1) \neq \hat{\delta}^*(-\bar{y}_1)$$

holds then

$$1 - \hat{\delta}^*(-\bar{y}_1) \neq \hat{\delta}^*(\bar{y}_1)$$

also holds.

This contradicts the assumption that $\hat{\delta}^*$ is a minimax optimal rule, i.e.,

$$R_{sq}(\hat{\delta}^*, P_\tau) = \inf_{\hat{\delta} \in \mathcal{D}} R_{sq}(\hat{\delta}, P_\tau).$$

□

Lemma C.3. A least favorable prior distribution π^* is such that

$$\pi^*(\tau) = \frac{1}{2}, \pi^*(-\tau) = \frac{1}{2},$$

for some $\tau \in (0, \infty)$.

Proof. For each $\tau \in (0, \infty)$, consider the symmetric prior

$$\pi^*(\tau) = p_\tau, \pi^*(-\tau) = 1 - p_\tau, \text{ where } p_\tau \in [0, 1]. \tag{C.3}$$

If (C.3) is indeed the least favorable prior, then $\hat{\delta}^*(\bar{y}_1) = \frac{(1-p_\tau) f\{\bar{y}_1|-\tau\}}{p_\tau f\{\bar{y}_1|\tau\} + (1-p_\tau) f\{\bar{y}_1|-\tau\}}$, and the mean square regret of $\hat{\delta}^*$ at P_τ is

$$R_{sq}(\hat{\delta}^*, P_\tau) = \tau^2 \int \frac{(1-p_\tau)^2 f^2\{\bar{y}_1| - \tau\} f\{\bar{y}_1|\tau\}}{\left[p_\tau f\{\bar{y}_1|\tau\} + (1-p_\tau) f\{\bar{y}_1| - \tau\}\right]^2} d\bar{y}_1. \tag{C.4}$$

The mean square regret of $\hat{\delta}^*$ at $P_{-\tau}$ is

$$R_{sq}(\hat{\delta}^*, P_{-\tau}) = \tau^2 \int \frac{p_\tau^2 f^2\{\bar{y}_1|\tau\} f\{\bar{y}_1| - \tau\}}{\left[p_\tau f\{\bar{y}_1|\tau\} + (1-p_\tau) f\{\bar{y}_1| - \tau\}\right]^2} d\bar{y}_1. \tag{C.5}$$

By Lemma C.2, τ and $-\tau$ yield the same mean square regret at $\hat{\delta}^*$, so p_τ must be such that

$$(C.4) = (C.5).$$

For each $\bar{y}_1$, the numerator of the integrand in (C.4) is

$$\begin{aligned}
&(1-p_\tau)^2 f^2\{\bar{y}_1| - \tau\} f\{\bar{y}_1|\tau\} \\
=&(1-p_\tau)^2 \left(\frac{1}{2\pi}\right)^{\frac{3}{2}} \exp\left(-\left[(\bar{y}_1+\tau)^2\right] - \frac{1}{2}\left[(\bar{y}_1-\tau)^2\right]\right)
\end{aligned}$$

while for each $\bar{y}_1$, the numerator of the integrand in (C.5) is

$$p_\tau^2 \left(\frac{1}{2\pi}\right)^{\frac{3}{2}} \exp\left(-\left[(\bar{y}_1-\tau)^2\right] - \frac{1}{2}\left[(\bar{y}_1+\tau)^2\right]\right).$$

Therefore, (C.5) can be written as

$$\begin{aligned} & R_{sq}(\hat{\delta}^*, P_{-\tau}) \\ =& \tau^2 \int \frac{p_\tau^2 \left(\frac{1}{2\pi}\right)^{\frac{3}{2}} \exp\left(-\left[(\bar{y}_1-\tau)^2\right] - \frac{1}{2}\left[(\bar{y}_1+\tau)^2\right]\right)}{\left[p_\tau f\{\bar{y}_1|\tau\} + (1-p_\tau) f\{\bar{y}_1|-\tau\}\right]^2} d\bar{y}_1 \\ =& \tau^2 \int \frac{p_\tau^2 \left(\frac{1}{2\pi}\right)^{\frac{3}{2}} \exp\left(-\left[(\bar{y}_1+\tau)^2\right] - \frac{1}{2}\left[(\bar{y}_1-\tau)^2\right]\right)}{\left[p_\tau f\{-\bar{y}_1|\tau\} + (1-p_\tau) f\{-\bar{y}_1|-\tau\}\right]^2} d\bar{y}_1, \end{aligned} \tag{C.6}$$

where the second equality follows from a change of variable. (C.4) admits

$$R_{sq}(\hat{\delta}^*, P_\tau) = \tau^2 \int \frac{(1-p_\tau)^2 \left(\frac{1}{2\pi}\right)^{\frac{3}{2}} \exp\left(-\left[(\bar{y}_1+\tau)^2\right] - \frac{1}{2}\left[(\bar{y}_1-\tau)^2\right]\right)}{\left[p_\tau f\{\bar{y}_1|\tau\} + (1-p_\tau) f\{\bar{y}_1|-\tau\}\right]^2} d\bar{y}_1. \tag{C.7}$$

Hence, p_τ must be such that

$$(C.6) = (C.7),$$

which is satisfied if $p_\tau = \frac{1}{2}$. Indeed, when $p_\tau = \frac{1}{2}$, (C.6) and (C.7) only differ in their denominators. Furthermore, for (C.7), the denominator of the integrand can be written as

$$\begin{aligned} & \left[\frac{1}{2} f\{\bar{y}_1|\tau\} + \frac{1}{2} f\{\bar{y}_1|-\tau\}\right]^2 \\ =& \left[\frac{1}{2}\sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}(\bar{y}_1-\tau)^2\right) + \frac{1}{2}\sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}(\bar{y}_1+\tau)^2\right)\right]^2 \end{aligned}$$

while for (C.6), the corresponding restatement is

$$\begin{aligned}
&\left[\frac{1}{2}f\{-\bar{y}_1|\tau\}+\frac{1}{2}f\{-\bar{y}_1|-\tau\}\right]^2\\
=&\left[\frac{1}{2}\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left(-\bar{y}_1-\tau\right)^2\right)+\frac{1}{2}\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left(-\bar{y}_1+\tau\right)^2\right)\right]^2\\
=&\left[\frac{1}{2}\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left(\bar{y}_1+\tau\right)^2\right)+\frac{1}{2}\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left(\bar{y}_1-\tau\right)^2\right)\right]^2
\end{aligned}$$

which is equivalent. □

Lemma C.4. If

$$\tau^*\in\arg\sup_{\tau\in[0,\infty)}\tau^2\int\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)^2f\{\bar{y}_1|\tau\}d\bar{y}_1,$$

then $\left(\frac{\partial}{\partial\tau}R_{sq}(\hat{\delta}^*,P_\tau)\right)|_{\tau=\tau^*}=0$.

Proof. Since $\tau^*\in\arg\sup_{\tau\in[0,\infty)}\tau^2\int\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)^2f\{\bar{y}_1|\tau\}d\bar{y}_1$ and the objective function is continuously differentiable, it holds that

$$\left[\frac{\partial}{\partial\tau}\left(\tau^2\int\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)^2f\{\bar{y}_1|\tau\}d\bar{y}_1\right)\right]|_{\tau=\tau^*}=0. \tag{C.8}$$

On the other hand,

$$\begin{aligned}
&\left(\frac{\partial}{\partial\tau}R_{sq}(\hat{\delta}^*,P_\tau)\right)\\
=&\frac{\partial}{\partial\tau}\left(\tau^2\int\left[\frac{f\{\bar{y}_1|-\tau^*\}}{f\{\bar{y}_1|\tau^*\}+f\{\bar{y}_1|-\tau^*\}}\right]^2f(\bar{y}_1|\tau)d\bar{y}_1\right).
\end{aligned}\tag{C.9}$$

Observing the objective function in (C.8) and (C.9), $\left(\frac{\partial}{\partial\tau}R_{sq}(\hat{\delta}^*,P_\tau)\right)|_{\tau=\tau^*}=0$ holds if

$$\left[\frac{\partial}{\partial\tau}\left((\tau^*)^2\int\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)^2f\{\bar{y}_1|\tau^*\}d\bar{y}_1\right)\right]|_{\tau=\tau^*}=0. \tag{C.10}$$

In what follows, we verify that (C.10) indeed holds. Note

$$
\begin{aligned}
&\frac{\partial}{\partial\tau}\left((\tau^*)^2\int\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)^2 f\{\bar{y}_1|\tau^*\}d\bar{y}_1\right)\\
=&(\tau^*)^2\int 2\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)\frac{\partial}{\partial\tau}\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)f\{\bar{y}_1|\tau^*\}d\bar{y}_1\\
=&2(\tau^*)^2\int\left(\frac{f\{\bar{y}_1|-\tau\}f\{\bar{y}_1|\tau^*\}}{[f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}]^3}\right)\left(\frac{\partial f\{\bar{y}_1|-\tau\}}{\partial\tau}f\{\bar{y}_1|\tau\}-f\{\bar{y}_1|-\tau\}\frac{\partial f\{\bar{y}_1|\tau\}}{\partial\tau}\right)d\bar{y}_1.
\end{aligned}
$$

Since $f\{\bar{y}_1|\tau\} = \sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1-\tau)^2\right]\right)$ and $f\{\bar{y}_1|-\tau\} = \sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1+\tau)^2\right]\right)$, we have that

$$
\frac{\partial f\{\bar{y}_1|\tau\}}{\partial\tau} = f\{\bar{y}_1|\tau\}(\bar{y}_1-\tau) \quad \text{and} \quad \frac{\partial f\{\bar{y}_1|-\tau\}}{\partial\tau} = -f\{\bar{y}_1|-\tau\}(\bar{y}_1+\tau).
$$

It follows that

$$
\begin{aligned}
&\frac{\partial}{\partial\tau}\left((\tau^*)^2\int\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)^2 f\{\bar{y}_1|\tau^*\}d\bar{y}_1\right)\\
=&-4(\tau^*)^2\int\left(\frac{f\{\bar{y}_1|-\tau\}f\{\bar{y}_1|\tau^*\}f\{\bar{y}_1|-\tau\}f\{\bar{y}_1|\tau\}}{[f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}]^3}\right)\bar{y}_1 d\bar{y}_1. \qquad \text{(C.11)}
\end{aligned}
$$

Evaluating (C.11) at $\tau=\tau^*$ yields

$$
\begin{aligned}
&\left[\frac{\partial}{\partial\tau}\left((\tau^*)^2\int\left(\frac{f\{\bar{y}_1|-\tau\}}{f\{\bar{y}_1|\tau\}+f\{\bar{y}_1|-\tau\}}\right)^2 f\{\bar{y}_1|\tau^*\}d\bar{y}_1\right)\right]|_{\tau=\tau^*}\\
=&-4(\tau^*)^2\int\left[\frac{(f\{\bar{y}_1|-\tau^*\}f\{\bar{y}_1|\tau^*\})^2}{[f\{\bar{y}_1|\tau^*\}+f\{\bar{y}_1|-\tau^*\}]^3}\right]\bar{y}_1 d\bar{y}_1\\
=&-4(\tau^*)^2\int w(\bar{y}_1)\bar{y}_1 d\bar{y}_1, \qquad \text{(C.12)}
\end{aligned}
$$

where $w(\bar{y}_1)=\frac{(f\{\bar{y}_1|-\tau^*\}f\{\bar{y}_1|\tau^*\})^2}{[f\{\bar{y}_1|\tau^*\}+f\{\bar{y}_1|-\tau^*\}]^3}$. However, notice for each $\bar{y}_1\in\mathbb{R}$:

$$
w(\bar{y}_1)=\frac{\left(\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1+\tau^*)^2\right]\right)\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1-\tau^*)^2\right]\right)\right)^2}{\left[\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1-\tau^*)^2\right]\right)+\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1+\tau^*)^2\right]\right)\right]^3}
$$

while

$$
\begin{aligned}
w(-\bar{y}_1) &= \frac{\left(\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(-\bar{y}_1+\tau^*)^2\right]\right)\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(-\bar{y}_1-\tau^*)^2\right]\right)\right)^2}{\left[\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(-\bar{y}_1-\tau^*)^2\right]\right)+\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(-\bar{y}_1+\tau^*)^2\right]\right)\right]^3} \\
&= \frac{\left(\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1-\tau^*)^2\right]\right)\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1+\tau^*)^2\right]\right)\right)^2}{\left[\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1+\tau^*)^2\right]\right)+\sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left[(\bar{y}_1-\tau^*)^2\right]\right)\right]^3} \\
&= \frac{\left(f\{\bar{y}_1|\tau^*\}f\{\bar{y}_1|-\tau^*\}\right)^2}{\left[f\{\bar{y}_1|-\tau^*\}+f\{\bar{y}_1|\tau^*\}\right]^3} \\
&= w(\bar{y}_1).
\end{aligned}
$$

Therefore

$$
(C.12) = -4(\tau^*)^2\int w(\bar{y}_1)\bar{y}_1 d\bar{y}_1 = 0
$$

and the conclusion of the lemma follows. □

Lemma C.5. τ^* is the unique solution of $\sup_{\tau\in[0,\infty)} R_{sq}(\hat{\delta}^*, P_\tau)$.

Proof. Write $\omega^*(\bar{y}_1) := (1-\hat{\delta}^*(\bar{y}_1))^2$. We evaluate the first derivative of

$$
R_{sq}(\hat{\delta}^*, P_\tau) = \tau^2\int \left[\omega^*(\bar{y}_1)\right]^2 f\{\bar{y}_1|\tau\}d\bar{y}_1
$$

as a function of $\tau\in[0,\infty)$. Notice for each $\bar{y}_1\in\mathbb{R}$ and each $\tau\in[0,\infty)$,

$$
\frac{\partial}{\partial\tau}f\{\bar{y}_1|\tau\} = f\{\bar{y}_1|\tau\}\left(\bar{y}-\tau\right) = -\frac{\partial}{\partial\bar{y}_1}f\{\bar{y}_1|\tau\}.
$$

Therefore, using integration by parts twice yields

$$
\begin{aligned}
R_{sq}^{(1)}(\tau) &:= \frac{\partial}{\partial\tau} R_{sq}(\hat{\delta}^*, P_\tau) \\
&= 2\tau \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 + \tau^2 \int \left[\omega^*(\bar{y}_1)\right]^2 \frac{\partial}{\partial\tau} \left(f(\bar{y}_1|\tau)\right) d\bar{y}_1 \\
&= 2\tau \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 - \tau^2 \int \left[\omega^*(\bar{y}_1)\right]^2 \frac{\partial}{\partial\bar{y}_1} f(\bar{y}_1|\tau) d\bar{y}_1 \\
&= 2\tau \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 - \tau^2 \int \left[\omega^*(\bar{y}_1)\right]^2 df(\bar{y}_1|\tau) \\
&= 2\tau \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 + 2\tau^2 \left(\int \omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} f(\bar{y}_1|\tau) d\bar{y}_1 \right) \\
&= 2\tau \left\{ \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 + \int \omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} \tau f(\bar{y}_1|\tau) d\bar{y}_1 \right\} \\
&= 2\tau \Bigg\{ \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 + \int \omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} \left(\tau - \bar{y}_1\right) f(\bar{y}_1|\tau) d\bar{y}_1 \\
&\qquad + \int \omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} \bar{y}_1 f(\bar{y}_1|\tau) d\bar{y}_1 \Bigg\} \\
&= 2\tau \Bigg\{ \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 + \int \omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} df(\bar{y}_1|\tau) \\
&\qquad + \int \omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} \bar{y}_1 f(\bar{y}_1|\tau) d\bar{y}_1 \Bigg\} \\
&= 2\tau \Bigg\{ \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 - \int \frac{\partial}{\partial\bar{y}_1} \left[\omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} \right] f(\bar{y}_1|\tau) d\bar{y}_1 \\
&\qquad + \int \omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} \bar{y}_1 f(\bar{y}_1|\tau) d\bar{y}_1 \Bigg\} \\
&= 2\tau \Bigg\{ \int \Bigg\{ \left[\omega^*(\bar{y}_1)\right]^2 - \left(\frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} \right)^2 - \omega^*(\bar{y}_1) \frac{\partial^2\left(\omega^*(\bar{y}_1)\right)}{\partial\left(\bar{y}_1\right)^2} \\
&\qquad\quad + \omega^*(\bar{y}_1) \frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial\bar{y}_1} \bar{y}_1 \Bigg\} f(\bar{y}_1|\tau) d\bar{y}_1 \Bigg\}.
\end{aligned}
$$

The sign of $R_{sq}^{(1)}(\tau)$ is determined by

$$
\mathbf{R}(\tau) := \int \mathbf{w}(\bar{y}_1) f\{\bar{y}_1|\tau\} d\bar{y}_1,
$$

where

$$\mathbf{w}(\bar{y}_1) = [\omega^*(\bar{y}_1)]^2 - \left(\frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial \bar{y}_1}\right)^2 - \omega^*(\bar{y}_1)\frac{\partial^2\left(\omega^*(\bar{y}_1)\right)}{\partial\left(\bar{y}_1\right)^2} + \omega^*(\bar{y}_1)\frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial \bar{y}_1}\bar{y}_1.$$

We aim to show that $\mathbf{R}(\tau)$ has a unique sign change from $+$ to $-$ at τ^*, with the conclusion immediately following.

Step 1: we show $\mathbf{R}(\tau)$ has at most one sign change from $+$ to $-$. Notice $\omega^*(\bar{y}_1) = \frac{1}{\exp(2\tau^*\bar{y}_1)+1}$. Therefore,

$$\begin{aligned}
\frac{\partial\left(\omega^*(\bar{y}_1)\right)}{\partial \bar{y}_1} &= -\left[\omega^*(\bar{y}_1)\right]^2 \exp\left(2\tau^*\bar{y}_1\right) 2\tau^*, \\
\frac{\partial^2\left(\omega^*(\bar{y}_1)\right)}{\partial\left(\bar{y}_1\right)^2} &= 2\left(\exp\left(2\tau^*\bar{y}_1\right) 2\tau^*\right)^2 \left[\omega^*(\bar{y}_1)\right]^3 - \left[\omega^*(\bar{y}_1)\right]^2 \exp\left(2\tau^*\bar{y}_1\right)\left(2\tau^*\right)^2
\end{aligned}$$

Plugging in $\mathbf{w}(\bar{y}_1)$ yields

$$\begin{aligned}
\mathbf{w}(\bar{y}_1) = \left[\omega^*(\bar{y}_1)\right]^2 \Big\{ & 1 - 3\left(\omega^*(\bar{y}_1)\exp\left(2\tau^*\bar{y}_1\right) 2\tau^*\right)^2 + \omega^*(\bar{y}_1)\exp\left(2\tau^*\bar{y}_1\right)\left(2\tau^*\right)^2 \\
& - \omega^*(\bar{y}_1)\exp\left(2\tau^*\bar{y}_1\right) 2\tau^*\bar{y}_1 \Big\} \\
= \left[\omega^*(\bar{y}_1)\right]^2 & \left\{ 1 - 3\left(\hat{\delta}^*(\bar{y}_1)2\tau^*\right)^2 + \hat{\delta}^*(\bar{y}_1)\left(2\tau^*\right)^2 - \hat{\delta}^* 2\tau^*\bar{y}_1 \right\}.
\end{aligned}$$

Since $\left[\omega^*(\bar{y}_1)\right]^2 > 0$ for all $\bar{y}_1$, the sign of $\mathbf{w}(\bar{y}_1)$ is determined by

$$\tilde{\mathbf{w}}(\bar{y}_1) = 1 - 3\left(\hat{\delta}^*(\bar{y}_1)2\tau^*\right)^2 + \hat{\delta}^*(\bar{y}_1)\left(2\tau^*\right)^2 - 2\tau^*\hat{\delta}^*(\bar{y}_1)\bar{y}_1$$

Since $\hat{\delta}^*(\bar{y}_1) > 0$, $\tilde{\mathbf{w}}(\bar{y}_1) = 0$ if and only if

$$\frac{1}{\hat{\delta}^*(\bar{y}_1)} - 3\left(2\tau^*\right)^2 \hat{\delta}^*(\bar{y}_1) + \left(2\tau^*\right)^2 = 2\tau^*\bar{y}_1. \tag{C.13}$$

Moreover, it is straightforward to check that $\frac{\partial}{\partial \bar{y}_1}\hat{\delta}^*(\bar{y}_1) > 0$. It follows the first derivative of the left hand side (LHS) of (C.13) is

$$\frac{\partial \text{LHS}}{\partial \bar{y}_1} = \left(-\frac{1}{\left(\hat{\delta}^*(\bar{y}_1)\right)^2} - 3\left(2\tau^*\right)^2 \right) \frac{\partial}{\partial \bar{y}_1} \hat{\delta}^*(\bar{y}_1) < 0,$$

which implies the LHS of (C.13) is a decreasing function. Also, the right hand side of (C.13) is an increasing function. Thus, (C.13) has at most one sign change from $+$ to $-$. Furthermore, note $\lim_{\bar{y}_1 \to -\infty} \tilde{\mathbf{w}}(\bar{y}_1) = 1$ and $\lim_{\bar{y}_1 \to \infty} \tilde{\mathbf{w}}(\bar{y}_1) = -\infty$, implying (C.13) indeed has one and only one sign change from $+$ to $-$. It follows from Theorem C.1 (i) that $\mathbf{R}(\tau)$ also has at most one sign change.

Step 2: we show $\mathbf{R}(\tau)$ indeed has one sign change. Note it also holds that

$$\mathbf{R}(\tau) = \int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau) d\bar{y}_1 + \frac{1}{2}\tau \int \left[\omega^*(\bar{y}_1)\right]^2 \frac{\partial}{\partial \tau} \left(f\{\bar{y}_1|\tau\}\right) d\bar{y}_1.$$

Hence,

$$\frac{\partial}{\partial \tau}\mathbf{R}(\tau) = \frac{3}{2}\mathbf{R}_1(\tau) + \frac{1}{2}\tau \mathbf{R}_2(\tau)$$

where

$$\begin{aligned}
\mathbf{R}_1(\tau) &= \int \left[\omega^*(\bar{y}_1)\right]^2 \frac{\partial}{\partial \tau} \left(f\{\bar{y}_1|\tau\}\right) d\bar{y}_1, \\
\mathbf{R}_2(\tau) &= \int \left[\omega^*(\bar{y}_1)\right]^2 \frac{\partial^2}{\partial \tau^2} \left(f\{\bar{y}_1|\tau\}\right) d\bar{y}_1.
\end{aligned}$$

By Lemma C.4, $\mathbf{R}(\tau^*) = 0$. Since $\int \left[\omega^*(\bar{y}_1)\right]^2 f(\bar{y}_1|\tau^*) d\bar{y}_1 > 0$, it holds that $\mathbf{R}_1(\tau^*) < 0$. Moreover, note

$$\frac{\partial^2}{\partial \tau^2} \left(f\{\bar{y}_1|\tau\}\right) = f\{\bar{y}_1|\tau\}(\bar{y}_1 - \tau)^2 - f\{\bar{y}_1|\tau\}.$$

Hence,

$$\begin{aligned}\mathbf{R}_2(\tau) &= \int \left[\omega^*(\bar{y}_1)\right]^2 f\{\bar{y}_1|\tau\}(\bar{y}_1-\tau)^2 d\bar{y}_1 - \int \left[\omega^*(\bar{y}_1)\right]^2 f\{\bar{y}_1|\tau\} d\bar{y}_1 \\ &= \int \left[\omega^*(t+\tau)\right]^2 f\{t|0\} t^2 dt - \int \left[\omega^*(t+\tau)\right]^2 f\{t|0\} dt \\ &= \int \left[\omega^*(t+\tau)\right]^2 f\{t|0\} \left(t^2-1\right) dt < 0\end{aligned}$$

for all $\tau > 0$ since $f\{t|0\}\left(t^2-1\right)$ as a function of t is symmetric around zero, and

$$\left[\omega^*(t+\tau)\right]^2 = \left[\frac{1}{\exp\left(2\tau^*(t+\tau)\right)+1}\right]^2$$

is a decreasing function of t. Therefore, we conclude that

$$\left[\frac{\partial}{\partial\tau}\mathbf{R}(\tau)\right]_{\tau=\tau^*} = \frac{3}{2}\mathbf{R}_1(\tau^*) + \frac{1}{2}\tau^*\mathbf{R}_2(\tau^*) < 0,$$

implying τ^* is indeed a point of sign change. Thus, $\mathbf{R}(\tau)$ indeed has one and only one sign change by Theorem C.1 (i).

Step 3: From Steps 1 and 2, Theorem C.1 (ii) further implies that $\mathbf{R}(\tau)$ and $\tilde{\mathbf{w}}(\bar{y}_1)$ changes sign in the same order. Hence, we conclude that $\mathbf{R}(\tau)$ only has one sign change at τ^* from + to −, i.e., τ^* is indeed a unique maximum of $\sup_{\tau\in[0,\infty)} R_{sq}(\hat{\delta}^*, P_\tau)$.

□

Theorem C.1 (Theorem 3 and Corollary 2, Karlin (1957))**.** *Let p be strictly Pólya type ∞ and assume that p can be differentiated n times with respect to x for all t. Let F be a measure on the real line, and let h be a function of t which changes sign n times.*

(i) If

$$g(x) = \int p(x,t)h(t)dF(t)$$

can be differentiated n times with respect to x inside the integral sign, then g changes sign at most n times and has at most n zeroes counting multiplicities, or is identically zero. The function g is identically zero if and only if the spectrum of F is contained in the set of zeros of h.

(ii) If the number of sign changes of g is n, then g and h change signs in the same order.

D Lemmas for Section 5

Lemma D.1. Treatment rule $\hat{\delta}^*_\tau$ is a minimax optimal rule in the limit experiment, i.e., $\sup_h R^\infty_{sq}(\hat{\delta}^*_\tau, h) = R^*$.

Proof. The mean square regret of a treatment rule $\hat{\delta}_\tau$ for each $h_1(b, h_0)$ is

$$\begin{aligned} R^\infty_{sq}(\hat{\delta}_\tau, h_1(b, h_0)) &= [\dot{\tau}' h_1(b, h_0)]^2 \, \mathbb{E}_{\Delta \sim N(h_1(b,h_0), I_0^{-1})} \left[1\left\{\dot{\tau}' h_1(b, h_0) \geq 0\right\} - \hat{\delta}\left(\dot{\tau}' \Delta\right)\right]^2 \\ &= [\dot{\tau}' h_0 + b]^2 \, \mathbb{E}_{\Delta \sim N(h_1(b,h_0), I_0^{-1})} \left[1\left\{\dot{\tau}' h_0 + b \geq 0\right\} - \hat{\delta}\left(\dot{\tau}' \Delta\right)\right]^2 \\ &= R^\infty_{sq}(\hat{\delta}_\tau, h_1(b, 0)), \end{aligned} \tag{D.1}$$

where the last relation follows from $\dot{\tau}' h_1(b, h_0) = \dot{\tau}' h_0 + b$ and $\dot{\tau}' h_0 = 0$ by construction. Thus,

$$\begin{aligned} R^* &\leq \sup_{h_1} R^\infty_{sq}(\hat{\delta}^*_\tau, h_1) = \sup_{h_0} \sup_{b} R^\infty_{sq}(\hat{\delta}^*_\tau, h_1(b, h_0)) \\ &= \sup_{b} R^\infty_{sq}(\hat{\delta}^*_\tau, h_1(b, 0)) \leq \sup_{b} R^\infty_{sq}(\hat{\delta}^*, h_1(b, 0)) \leq R^*, \end{aligned}$$

where the first relation follows from the definition of R^*, the second relation follows from definition of h_1, the third relation follows from (D.1), the fourth relation follows by definition of $\hat{\delta}^*_\tau$, and the final relation follows because R^* is the worst case mean square regret of $\hat{\delta}^*$ and so must be no smaller than $\sup_b R^\infty_{sq}(\hat{\delta}^*, h_1(b, 0))$. □

Lemma D.2. $\hat{\delta}^*_\tau$ can be found by solving

$$\inf_{\hat{\delta}_\tau} \sup_{b} R^\infty_{sq}(\hat{\delta}_\tau, b), \tag{D.2}$$

where we recall that

$$R^\infty_{sq}(\hat{\delta}_\tau, b) = b^2 \mathbb{E}_{\Delta_\tau \sim N(b, \dot{\tau}' I_0^{-1} \dot{\tau})} \left[1\left\{b \geq 0\right\} - \hat{\delta}_\tau\left(\Delta_\tau\right)\right]^2.$$

Proof. Note for each $b \in \mathbb{R}$ and $\hat{\delta}_\tau = \hat{\delta}(\dot{\tau}'\Delta)$,

$$
\begin{aligned}
R_{sq}^{\infty}(\hat{\delta}_\tau, h_1(b,0)) &= b^2 \mathbb{E}_{\Delta \sim N(h_1(b,h_0), I_0^{-1})} \left[1\{b \geq 0\} - \hat{\delta}\left(\dot{\tau}'\Delta\right) \right]^2 \\
&= b^2 \mathbb{E}_{\Delta_\tau \sim N(b, \dot{\tau}' I_0^{-1} \dot{\tau})} \left[1\{b \geq 0\} - \hat{\delta}_\tau\left(\Delta_\tau\right) \right]^2 \\
&= R_{sq}^{\infty}(\hat{\delta}_\tau, b)
\end{aligned}
$$

where the second equality follows from

$$
\Delta_\tau = \dot{\tau}'\Delta \sim N(\dot{\tau}' h_1(b, h_0), \dot{\tau}' I_0^{-1} \dot{\tau}),
$$

$\dot{\tau}' h_1(b, h_0) = b$ and we defined $\hat{\delta}_\tau(\Delta_\tau) = \hat{\delta}\left(\dot{\tau}'\Delta\right)$. □

Lemma D.3. Under assumptions of Theorem 5.1, the minimax optimal policy in the limit experiment is

$$
\hat{\delta}^*(\Delta) = \frac{\exp\left(\frac{2\tau^*}{\sqrt{\dot{\tau}' I_0^{-1} \dot{\tau}}} \dot{\tau}'\Delta \right)}{\exp\left(\frac{2\tau^*}{\sqrt{\dot{\tau}' I_0^{-1} \dot{\tau}}} \dot{\tau}'\Delta \right) + 1},
$$

where $\tau^* \approx 1.23$ and solves (4.6) or (4.7).

Proof. By Lemmas D.1 and D.2, it suffices to find $\hat{\delta}_\tau^*$. Recall $\sigma_\tau^2 = \dot{\tau}' I_0^{-1} \dot{\tau}$. Thus,

$$
\begin{aligned}
R_{sq}^{\infty}(\hat{\delta}_\tau, b) &= b^2 \mathbb{E}_{\Delta_\tau \sim N(b, \sigma_\tau^2)} \left[1\{b \geq 0\} - \hat{\delta}_\tau\left(\Delta_\tau\right) \right]^2 \\
&= b^2 \int \left[1\{b \geq 0\} - \hat{\delta}_\tau\left(\Delta_\tau\right) \right]^2 \frac{1}{\sqrt{2\pi\sigma_\tau^2}} \exp\left(-\frac{(\Delta_\tau - b)^2}{2\sigma_\tau^2} \right) d\Delta_\tau \\
&= b^2 \int \left[1\{b \geq 0\} - \hat{\delta}_\tau\left(\sigma_\tau z\right) \right]^2 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(z - \frac{b}{\sigma_\tau})^2}{2} \right) dz \\
&= \sigma_\tau^2 \left(b_\tau\right)^2 \int \left[1\{b_\tau \geq 0\} - \hat{\delta}_1(z) \right]^2 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(z - b_\tau)^2}{2} \right) dz \\
&= \sigma_\tau^2 \left(b_\tau\right)^2 \mathbb{E}_{Z \sim N(b_\tau, 1)} \left[1\{b_\tau \geq 0\} - \hat{\delta}_1(Z) \right]^2
\end{aligned}
$$

where the third line follows from the change of variable $z = \frac{\Delta_\tau}{\sigma_\tau}$, and the fourth line follows by letting $b_\tau := \frac{b}{\sigma_\tau}$ and $\hat{\delta}_1(z) := \hat{\delta}_\tau\left(\sigma_\tau z\right)$. Therefore, the minimax optimal rule $\hat{\delta}_\tau^*(\mathit{\Delta}_\tau) = \hat{\delta}_1^*(\frac{\mathit{\Delta}_\tau}{\sigma_\tau})$, where $\hat{\delta}_1^*(z)$ solves

$$\min_{\hat{\delta}_1} \sup_{b_\tau} R_{sq}^\infty(\hat{\delta}_1, b_\tau), \tag{D.3}$$

and where $R_{sq}^\infty(\hat{\delta}_1, b_\tau) = (b_\tau)^2 \mathbb{E}_{Z\sim N(b_\tau,1)} \left[1\{b_\tau \geq 0\} - \hat{\delta}_1(Z)\right]^2$. By Theorem 4.2, we know the solution of (D.3) is $\hat{\delta}_1(z) = \frac{\exp(2\tau^* z)}{\exp(2\tau^* z)+1}$, where τ^* solves (4.6) or (4.7). Hence, $\hat{\delta}_\tau^*(\mathit{\Delta}_\tau) = \frac{\exp\left(2\tau^* \frac{\mathit{\Delta}_\tau}{\sigma_\tau}\right)}{\exp\left(2\tau^* \frac{\mathit{\Delta}_\tau}{\sigma_\tau}\right)+1}$. Finally, note $\mathit{\Delta}_\tau = \dot{\tau}'\mathit{\Delta}$. Thus, the minimax optimal policy in the limit is

$$\hat{\delta}^*(\mathit{\Delta}) = \frac{\exp\left(\frac{2\tau^*}{\sigma_\tau}\dot{\tau}'\mathit{\Delta}\right)}{\exp\left(\frac{2\tau^*}{\sigma_\tau}\dot{\tau}'\mathit{\Delta}\right)+1}.$$

□