Finite-Sample Minimax-Regret Treatment Choice with Interval-Observed Outcomes

JEL classification: C44, C21, C15.

Nikita Pavlov
Penn State University

October 2026

Abstract

Consider a setup where a decision maker has to choose whether to apply treatment to a population after observing a finite sample. Stoye (2009) characterises the solution to the minimax regret problem in this setup when a finite sample consists of actual outcome realizations. However, in many applied problems, instead of the realized outcomes, the decision maker observes only the intervals containing them. We study a finite-sample minimax regret problem in this setting. With no restriction on interval width, assigning treatment with probability 1/2 yields minimax regret. With widths bounded by h_0 and h_1 in the two arms, we first show that minimax regret converges to $(h_0 + h_1)/4$ in the limiting experiment and derive a corresponding asymptotically optimal plug-in treatment rule. Then, for $h_0 + h_1 \geq 1$, we derive an explicit minimax regret rule which attains this value in finite samples. For $h_0 + h_1 < 1$, we establish the existence of an optimal treatment rule and characterize it implicitly through a well-behaved finite-dimensional optimization problem. A simulation experiment then shows that the derived analytic rule produces superior finite-sample minimax regret relative to several heuristic alternatives, with the plug-in rule coming very close when the sample size is sufficiently large.

Introduction

In many applied problems the outcomes realizations can not be observed exactly and are reported in intervals instead. As such, different social surveys often report respondent characteristics (e.g., income) in brackets. Interval observations also arise naturally when the outcomes are censored (e.g., unemployment records) or measurements are recorded with known limits of precision (e.g., medical records). Moreover, depending on application, the widths of reported intervals may differ across observations and depend on the outcome itself.

Within such setting, suppose a decision maker has to choose whether to apply treatment to a population after observing a finite sample of interval reports. Generally, the solution to this problem depends on the decision criterion used. In an absence of a prior of the outcome distribution, one natural approach is to adopt the minimax regret criterion, which minimizes the worst potenital welfare loss relative to the oracle treatment choice. When the sample consists of actual outcome realizations, finite sample minimax regret treatment rules have been extensively characterized by Stoye (2007a, 2007b, 2009). However, when the outcomes are presented in intervals relatively little is known about finite sample minimax regret treatment rules.

In this paper, we study finite sample minimax regret treatment rules when the outcomes are interval-reported. In particular, we assume that a decision maker is presented with a fixed sample of N_0 intervals from the control group and N_1 intervals from the treatment group, $N_0, N_1 \geq 1$. Random outcomes are assumed to be bounded with normalized support $[0, 1]$. The endpoints of reported intervals are also allowed to be stochastic with respective widths no greater than h_0 and h_1 in the two treatment arms, $h_0, h_1 \in [0, 1]$. Apart from the bounded outcome support, independent sampling, and the width restriction, we impose no restrictions on joint distribution of outcomes and interval endpoints.

The analysis proceeds as follows. First, we show that with no restriction on interval width, assigning treatment with probability $1/2$ yields minimax regret. Next, with widths bounded by h_0 and h_1 in the two arms, we demonstrate that minimax regret converges to $(h_0 + h_1)/4$ in the limiting experiment. We also derive corresponding asymptotically optimal plug-in treatment rule and characterize upper bound of its efficiency loss in finite samples. Then, for $h_0 + h_1 \geq 1$, we derive an explicit finite-sample minimax regret rule which also attains $(h_0 + h_1)/4$ minimax value. For $h_0 + h_1 < 1$, we establish the existence of an optimal treatment rule and characterize it implicitly through a finite-dimensional optimization problem. A simulation experiment then examines the performance of the analytic finite sample minimax regret rule relatively to the plug-in rule and some simple heuristic alternatives. Derived analytic rule produces superior finite sample minimax regret, with plug-in rule coming very close when the sample size is sufficiently large.

Related literature

This paper contributes to the literature on statistical treatment choice under the minimax regret criterion, following the framework of statistical decision theory developed by Wald (1950). Manski (2004) studies statistical treatment rules based on randomized experimental data and derives finite-sample bounds on the maximal regret of empirical success rules. Stoye (2009) instead works directly with exact finite-sample regret and characterizes minimax regret treatment rules under several sampling designs, including matched pairs, random assignment, and testing an innovation. He finds that empirical success rules are often close to minimax regret optimal, but that substantial differences can arise for small sample sizes and certain sampling designs. Hirano and Porter (2009) study the asymptotic behavior of statistical treatment rules, while Kitagawa and Tetenov (2018) develop empirical welfare maximization methods for treatment choice with covariates. All these

papers assume that the outcomes are observed exactly.

Related work considers treatment choice when the objective is a quantile rather than expected welfare. Guggenberger, Mehta, and Pavlov (2026) characterize finite-sample minimax regret treatment rules when welfare is evaluated using a quantile of the outcome distribution. More recently, Manski and Mullahy (2025) study quantile-welfare evaluation of social policies as an alternative to utilitarian welfare, with an application to health cost-effectiveness analysis. These contributions illustrate how the choice of welfare objective affects the evaluation of alternative treatment policies.

Another closely related strand of literature considers treatment choice when the outcome distributions are only partially identified. Manski (2000) discusses identification problems and treatment choice under ambiguity. Manski (2007) and Stoye (2007a, 2007b) study minimax regret treatment choice when some outcomes are missing, including settings with finite-sample uncertainty. The information structure in the present paper is different. Rather than assuming that some outcomes are entirely unobservable, I assume that each sampled outcome is reported through an interval containing its realization.

More recent work studies statistical treatment choice under partial identification in broader settings. Yata (2023) derives exact finite-sample minimax regret treatment rules under Gaussian errors and restrictions on the underlying parameter space. Montiel Olea, Qiu, and Stoye (2026a) consider a general class of treatment choice problems with partial identification and characterize minimax regret rules in Gaussian experiments. They show, among other results, that sufficiently severe partial identification can lead to multiple minimax regret optimal rules. Christensen, Moon, and Schorfheide (2026) develop asymptotically optimal decision rules when payoffs are partially identified, with applications to treatment choice and pricing. Montiel Olea, Qiu, and Stoye (2026b) study treatment choice in an asymptotic framework that allows for both sampling uncertainty and partial identification. These contributions examine different forms of partial identification and statistical uncertainty. In the present paper, I instead focus on exact finite-sample treatment choice when the observed data consist of intervals containing the outcomes.

1 Formulation

Suppose a decision maker has to choose between two treatment regimes on the basis of a finite sample of signals about the outcomes. In Stoye (2009) these signals constitute random draws from the respective outcome distributions. Here we instead assume that she only observes outcome intervals which contain the respective actual outcomes almost surely.

Formally, let the potential treatment regime be denoted by $t \in \{0, 1\}$ with respective potential outcomes $Y_t \in [0, 1]$. The sample consists of $N_1 \geq 1$ treated and $N_0 \geq 1$ non-treated reports. Each report is a draw from a random tuple of endpoints of an interval containing the actual outcome:

$$Z_t = (L_t, U_t), \qquad 0 \leq L_t \leq Y_t \leq U_t \leq 1, \qquad t \in \{0, 1\}. \tag{1.1}$$

Let $\mathcal{Z} = \{(l, u) \in [0, 1]^2 : 0 \leq l \leq u \leq 1\}$ denote the space of interval reports. Let a state s determine the joint distribution of potential outcomes and their interval reports for the two treatment regimes. The pairs $(Y_{t,i}, Z_{t,i}), i \in \{1, ..., N_t\}$ are assumed to be i.i.d. within each treatment arm, and the two arm samples are independent. However, random vectors (Y_0, Z_0) and (Y_1, Z_1) corresponding to the same individual do not have to be independent.

Given the sample, decision maker then chooses a measurable treatment rule

$$\delta : \mathcal{Z}^{N_0+N_1} \longrightarrow [0, 1], \qquad \delta(W) = \mathbb{P}(\text{assign treatment } 1 \mid W), \tag{1.2}$$

where $W = (Z_{0,1}, \ldots, Z_{0,N_0}, Z_{1,1}, \ldots, Z_{1,N_1})$ is the sample of observed reports.

For a state s, define

$$\mu_t(s) = \mathbb{E}_s[Y_t], \qquad \Delta(s) = \mu_1(s) - \mu_0(s), \qquad \pi_\delta(s) = \mathbb{E}_s[\delta(W)]. \tag{1.3}$$

By the law of iterated expectation, expected welfare under rule δ is

$$\mathbb{E}_s[Y_\delta] = (1 - \pi_\delta(s))\mu_0(s) + \pi_\delta(s)\mu_1(s). \tag{1.4}$$

If the state s was known, it would therefore be optimal to select an arm with a larger mean. Relative to this hypothetical optimal choice, we define regret, which is a difference in expected welfare under oracle rule and under $\delta(W)$:

$$\begin{aligned} R(\delta, s) &= \max\{\mu_0(s), \mu_1(s)\} - \mathbb{E}_s[Y_\delta] \\ &= [\Delta(s)]_+ \left(1 - \pi_\delta(s)\right) + [-\Delta(s)]_+ \pi_\delta(s). \end{aligned} \tag{1.5}$$

where $[x]_+ = \max\{x, 0\}$.

Let $\mathcal{S}$ denote the set of all possible joint distributions of $\{(Y_0, Z_0), (Y_1, Z_1)\}$ that satisfy (1.1), and $\mathcal{D}$ denote the set of all measurable treatment rules the decision maker could choose from. Given the sample consisting of N_1 treated and N_0 non-treated reports, the minimax regret value is then

$$V_{N_0,N_1}(\mathcal{S}) = \inf_{\delta \in \mathcal{D}} \sup_{s \in \mathcal{S}} R(\delta, s). \tag{1.6}$$

Additionally, let $\mathcal{S}_{h_0,h_1}$ denote the set of all states that also satisfy the width bounds

$$U_t - L_t \leq h_t \qquad \text{a.s. for } t = 0, 1, \tag{1.7}$$

where $h_t \in [0, 1]$. Write $\mathcal{Z}_{h_t} = \{(l, u) \in \mathcal{Z} : u - l \leq h_t\}$ for the corresponding report space. At $h_0 = h_1 = 1$, both treated and non-treated reports are unrestricted in width. Throughout the paper, we use $V_{N_0,N_1}(h_0, h_1) = V_{N_0,N_1}(\mathcal{S}_{h_0,h_1})$ for notational convenience.

2 Unrestricted interval width

Let us first start with no restriction on the interval width. Under this regime, nature could always choose to produce reports that are completely uninformative about the outcome distribution. As a result, the no data rule $\delta = \frac{1}{2}$ achieves minimax regret.

Theorem 2.1 (No width restriction on the reports)**.** *If nature is restricted only by* (1.1)*, then, for all* $N_0, N_1 \geq 1$*,*

$$V_{N_0,N_1}(1, 1) = \frac{1}{2}. \tag{2.1}$$

Moreover, the no-data rule $\delta(W) \equiv 1/2$ *is minimax regret.*

3 Restricted interval widths

To make the problem more interesting, for the remainder of the analysis we impose (1.7) restriction on the nature's state space with $h_0 + h_1 < 2$.

3.1 Known distributions of the reported intervals

We start with an instructive case where the reporting distribution of $Z_t = (L_t, U_t)$ for each arm is assumed to be known by decision maker. Note that although the reporting distribution is known, the conditional distribution $Y_t|Z_t$ remains free. Analysis of this case serves two purposes. First, it provides intuition to the general approach of how minimax regret rules are constructed under restricted width, which is helpful in some of the later proofs. Second, the known distribution minimax treatment rule is then used to construct a plug-in treatment rule which is asymptotically minimax regret. As we would see, finite sample minimax regret rules under some of the width bounds require numerical optimization, which could be computationally demanding when the sample sizes are large. For such cases, a plug-in rule serves as a useful alternative.

Let Q_t denote the reporting distribution of $Z_t = (L_t, U_t)$. Define

$$\ell_t = \mathbb{E}_{Q_t}[L_t], \qquad u_t = \mathbb{E}_{Q_t}[U_t], \qquad a = \ell_1 - u_0, \quad b = u_1 - \ell_0. \tag{3.1}$$

Note that $a \leq b$ since $u_t \geq \ell_t$. Given the notation, we start with a simple technical lemma.

Lemma 3.1. *Given the known reporting distributions (Q_0, Q_1), possible treatment means and treatment effects chosen by nature are restricted as follows:*

$$\mu_t \in [\ell_t, u_t], \qquad \Delta \in [a, b]. \tag{3.2}$$

Given this lemma, we establish minimax results for the known reporting distribution case. Since instead of observing outcomes we only observe intervals that contain outcomes, and the distribution of endpoints of said intervals is known, decision maker could again completely disregard the data.

Theorem 3.2. *Suppose that the reporting distributions Q_0, Q_1 are known. Then there exists a no-data minimax rule that chooses treatment with constant probability depending on known means of reporting distributions. In particular, any rule of the form*

$$d^*(a, b; d) = \begin{cases} \dfrac{b_+}{b_+ + (-a)_+}, & b_+ + (-a)_+ > 0, \\ d, & a = b = 0. \end{cases} \tag{3.3}$$

where $x_+ = \max\{0, x\}$ and $d \in [0, 1]$, is minimax regret. Thus, $d^* = 1$ if $a \geq 0$ and $b > 0$, $d^* = 0$ if $b \leq 0$ and $a < 0$, and $d^* = b/(b - a)$ if $a < 0 < b$.

The corresponding minimax regret value is

$$v(a, b) = \begin{cases} \dfrac{b \cdot (-a)}{b - a}, & b > 0 > a, \\ 0, & otherwise. \end{cases} \tag{3.4}$$

Moreover,

$$v(a, b) \leq \frac{b - a}{4} \leq \frac{h_0 + h_1}{4}. \tag{3.5}$$

To better understand the rationale behind the optimal (3.3) rule, first suppose that $[a, b]$ lies entirely on one side of zero. Since by Lemma 3.1 nature's choice of treatment effect is restricted by $[a, b]$, it follows that treatment effect is either positive or negative with certainty, which yields $d^* = 1$ and $d^* = 0$, respectively. Conversely, when $a < 0 < b$, nature can choose any sign of treatment effect. The optimal response is thus to equalize the regrets at the two ends of the $[a, b]$ interval, which is done by setting $d^* = b/(b - a)$.

3.2 Unknown reporting distributions: a limiting minimax result

We now remove the restriction on reporting distributions, thus assuming that nature is allowed to choose freely both them and the conditional outcome laws. First, define the sample endpoint means

$$\widehat{\ell}_t = \frac{1}{N_t}\sum_{i=1}^{N_t} L_{t,i}, \qquad \widehat{u}_t = \frac{1}{N_t}\sum_{i=1}^{N_t} U_{t,i}.$$

Also define the corresponding sample endpoints of the treatment-effect interval

$$\widehat{a} = \widehat{\ell}_1 - \widehat{u}_0, \qquad \widehat{b} = \widehat{u}_1 - \widehat{\ell}_0. \tag{3.6}$$

By applying the known reporting distributions rule (3.3) on (3.6) with $d = \frac{1}{2}$ we obtain an intuitive plug-in rule

$$\widehat{\delta}(W) = d^*(\widehat{a}, \widehat{b}; d = \frac{1}{2}), \tag{3.7}$$

Note that since each individual report $Z_{t,i} = (L_{t,i}, U_{t,i})$ satisfies (1.1) and (1.7), it follows that $\widehat{b}, \widehat{a} \in [-1, 1]$ and so $\delta(W) = d^*(\widehat{a}, \widehat{b})$ is well-defined. The next theorem shows that this plug-in rule is asymptotically minimax regret.

Theorem 3.4. *For every fixed $h_0, h_1 \in [0, 1]$ and $\widehat{\delta}(W)$ as defined in (3.7),*

$$\lim_{N_0\to\infty, N_1\to\infty} V_{N_0,N_1}(h_0, h_1) = \lim_{N_0\to\infty, N_1\to\infty} \sup_{s\in\mathcal{S}_{h_0,h_1}} R(\widehat{\delta}(W), s) = \frac{h_0 + h_1}{4}. \tag{3.8}$$

Moreover, for every $N_0, N_1 \geq 1$,

$$\frac{h_0 + h_1}{4} \leq V_{N_0,N_1}(h_0, h_1) \leq \sup_{s\in\mathcal{S}_{h_0,h_1}} R(\widehat{\delta}, s) \leq \frac{h_0 + h_1}{4} + \frac{1}{\sqrt{2}}\sqrt{\frac{1}{N_0} + \frac{1}{N_1}} \tag{3.9}$$

Comments.

1. When the reporting distributions are unknown and restricted only by width, the nature could always choose symmetric states with respective treatment effects of $(h_0 + h_1)/2$ and $-(h_0 + h_1)/2$ and identical reports that yield the same treatment probability across both states. This produces regret of at least $\frac{h_0+h_1}{4}$. On the other hand, proof of Theorem 3.4 shows that in every possible state of nature the regret from plug-in rule consists of two terms. First term is a minimax regret value which would've occurred if the reporting distribution was known and had parameters $a = \widehat{a}, b = \widehat{b}$. As was before in Theorem 3.2, this term is bounded by $\frac{h_0+h_1}{4}$. Second term is the cost of sampling uncertainty, which disappears in the limit. Combining the upper and lower bounds yields the result of Theorem 3.4.

2. Theorem 3.4 provides a finite sample upper bound for minimax regret which also bounds the regret from the plug-in rule 3.7, in particular

$$V_{N_0,N_1}(h_0, h_1) \leq \sup_{s\in\mathcal{S}_{h_0,h_1}} R(\widehat{\delta}, s) \leq \frac{h_0 + h_1}{4} + \frac{1}{\sqrt{2}}\sqrt{\frac{1}{N_0} + \frac{1}{N_1}}. \tag{3.10}$$

   This gives an idea of potential efficiency losses from using the plug-in rule instead of the exact finite sample minimax regret rules derived later in this section.

Next, we derive the exact finite-sample minimax regret rules when the reporting distribution is unknown and the width is restricted. Two important cases arise here. Under $h_0 + h_1 \geq 1$, the nature can always use reports that are uninformative enough so that decision maker cannot deduce the sign of treatment effect. Under $h_0 + h_1 < 1$, the data can sometimes help to overcome the ambiguity. This distinction makes the finite sample treatment rule dependent on reporting width, and we consider the two cases separately.

3.3 An explicit finite-sample minimax characterisation when $h_0 + h_1 \geq 1$

We start with the wide interval case. The analysis in finite samples is substantially complicated by the fact that, in principle, nature does not have to choose reporting width to be exactly h_t in each arm. To overcome this issue, the next two lemmas show that: 1) restricting nature to the usage of canonical reports (reports of width exactly h_t in arm t) does not change the value of minimax regret; 2) any rule that is minimax regret when the nature is restricted to canonical reports is also minimax regret in the original problem.

Lemma 3.5 (Canonical enlargement)**.** *Let $0 \leq h_t \leq 1$ and $[L, U] \subseteq [0, 1]$ satisfy $U - L \leq h_t$. Define*

$$X_t(L, U) = \min\left\{1 - h_t, \max\left\{0, M - \frac{h_t}{2}\right\}\right\}, \tag{3.11}$$

where $M = \frac{L+U}{2}$. Then $X_t(L, U) \in [0, 1 - h_t]$ and

$$[L, U] \subseteq \big[X_t(L, U),\ X_t(L, U) + h_t\big]. \tag{3.12}$$

Throughout, for convenience, we assume that when the nature is restricted to canonical reports, the chosen report is a random variable X_t instead of the tuple $(X_t, X_t + h_t)$. Note that under restriction, the true outcome could be expressed in terms of X_t

$$Y_t = X_t + h_t A_t, \qquad A_t = \begin{cases} \frac{(Y_t - X_t)}{h_t} \in [0, 1], & h_t \in (0, 1], \\ 0, & h_t = 0. \end{cases} \tag{3.13}$$

Conversely, any joint distribution of (X_t, A_t) on $[0, 1 - h_t] \times [0, 1]$ yields a valid canonical report and outcome.

Lemma 3.6 (Minimax equivalence of canonical reports)**.** *For fixed widths $h_0, h_1 \in [0, 1]$, define a canonical restricted minimax problem as*

$$\inf_{\delta \in \mathcal{D}^{can}_{h_0,h_1}} \sup_{s \in \mathcal{S}^{can}_{h_0,h_1}} R(\delta, s), \tag{3.14}$$

where $\mathcal{D}^{can}_{h_0,h_1}$ consists of all measurable maps $\delta : [0, 1 - h_0]^{N_0} \times [0, 1 - h_1]^{N_1} \to [0, 1]$, and $\mathcal{S}^{can}_{h_0,h_1}$ is the set of all states with scalar reports $X_t \in [0, 1 - h_t]$ satisfying $X_t \leq Y_t \leq X_t + h_t$ almost surely.

Restricting nature to the reports of canonical form (or to the state space $\mathcal{S}^{can}_{h_0,h_1}$) does not change the minimax value. Moreover, in the original problem by applying $X_t(L, U)$ as defined in (3.11) to every report realization, any rule that is minimax in the restricted problem induces a minimax rule in the original problem.

Comments.

1. Lemma 3.6 claims that by restricting nature to canonical reports the minimax value does not change and we can transfer the restricted minimax rules to the original problem. It does not claim that the two state spaces are equivalent in any other regard.

2. For the intuition behind the proof, note that the restricted state space of canonical reports can be embedded into the original space. Hence, regret under the restricted space cannot be larger than under unrestricted space. For the other direction, note that arbitrary original report can be transformed into a canonical report by applying the map from Lemma 3.5. For an arbitrary canonical rule, we then construct an original rule such that for every $s \in \mathcal{S}_{h_0,h_1}$ there exists $T(s) \in \mathcal{S}^{can}_{h_0,h_1}$ that produces the same regret. Since by construction $T(\mathcal{S}_{h_0,h_1}) \subseteq \mathcal{S}^{can}_{h_0,h_1}$, the other direction follows.

3. Although lemmas 3.5 and 3.6 were initially derived to prove the minimax result under the wide interval case ($h_0 + h_1 \geq 1$), their result holds irrespective of the values of h_0, h_1 and are later used in the subsequent proofs.

   Given A_t as specified in (3.13), for $h_0, h_1 \in [0, 1]$, define

$$m_t = \mathbb{E}[X_t], \qquad a_t = \mathbb{E}[A_t], \qquad q = m_1 - m_0. \tag{3.15}$$

Then

$$\mu_t = m_t + h_t a_t, \qquad \Delta = q + h_1 a_1 - h_0 a_0 \in [q - h_0, q + h_1]. \tag{3.16}$$

Note that for a fixed q, since a_0, a_1 can vary over $[0, 1]$, every value in the treatment effect interval is attainable. With that said, we are ready to characterise the central minimax result for wide intervals.

Theorem 3.7 (Exact minimax rule for wide intervals)**.** *Suppose $N_0, N_1 \geq 1$ and $h_0 + h_1 \geq 1$. Transform each reported interval in the original sample W to $X_{t,i}$ by (3.11), and let*

$$\overline{X}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} X_{t,i}.$$

Then the rule

$$\boxed{\delta^*(W) = \frac{1}{2} + \frac{\overline{X}_1 - \overline{X}_0 + (h_1 - h_0)/2}{h_0 + h_1}} \tag{3.17}$$

is minimax regret. Moreover,

$$\boxed{V_{N_0,N_1}(h_0, h_1) = \frac{h_0 + h_1}{4}.} \tag{3.18}$$

Comments.

1. The rule (3.17) increases the treatment probability when the mean transformed report is higher in the treated arm, accounting for the difference in width bounds.

2. Note that at $h_0 = h_1 = 1$ each canonical signal is deterministically zero and so, as expected, the minimax regret rule and minimax regret value coincide with those of the unrestricted case from Theorem 2.1.

3. For $h_0 + h_1 < 1$, the rule (3.17) is no longer well defined as it may take values outside of $[0, 1]$. We study this case separately in the next section.

4. Note that it was never stated that the minimax rule 3.17 is unique. In fact, next corollary shows that the plug-in rule we devised in (3.7) is also minimax when the the intervals are wide.

Corollary 3.8. *Suppose $N_0, N_1 \geq 1$ and $H = h_0 + h_1 \geq 1$. Then the plug-in rule $\widehat{\delta}(W)$ defined in (3.7) is minimax regret in the original interval-report problem. In particular,*

$$\sup_{s \in \mathcal{S}_{h_0,h_1}} R(\widehat{\delta}, s) = V_{N_0,N_1}(h_0, h_1) = \frac{H}{4}. \tag{3.19}$$

3.4 An implicit finite-sample minimax characterisation when $h_0 + h_1 < 1$

Since the previously found minimax treatment rule is no longer valid in this region, we need a different one. Unfortunately, we were unable to obtain an explicit characterization of the resulting

minimax treatment rule for narrow intervals in closed form. Instead, we prove its existence and characterize it implicitly as a solution to a well-defined finite-dimensional optimization problem.

In order to do so, we start with a dimension reduction lemma, which establishes that restricting nature further to canonical reports that have Bernoulli distribution does not change the minimax regret value and allows treatment rule transfer from canonical Bernoulli problem to original problem. The proof is heavily inspired by an analogous result of Stoye (2009).

Lemma 3.10 (Bernoulli coarsening)**.** *For fixed widths $h_0, h_1 \in [0,1)$, define a canonical Bernoulli-restricted minimax problem as*

$$\inf_{\delta \in \mathcal{D}^{ber}_{h_0,h_1}} \sup_{s \in \mathcal{S}^{ber}_{h_0,h_1}} R(\delta, s), \tag{3.20}$$

where $\mathcal{D}^{ber}_{h_0,h_1}$ consists of all measurable maps $\delta : \{0,1\}^{N_0+N_1} \to [0,1]$, and $\mathcal{S}^{ber}_{h_0,h_1} \subseteq \mathcal{S}^{can}_{h_0,h_1}$ is a set of all canonical states in which the canonical reports are independent scalings of Bernoulli distributed random variables,

$$X_{t,i} = (1-h_t)B_{t,i}, \qquad B_{t,i} \sim \text{Bernoulli}(p_t), \qquad p_t \in [0,1], \qquad t \in \{0,1\},$$

with unrestricted latent components $A_t \in [0,1]$.

Restricting nature to the Bernoulli canonical reports (or to the state space $\mathcal{S}^{ber}_{h_0,h_1}$) does not change the minimax value. Moreover, any minimax rule for the Bernoulli problem can be transformed and implemented as a treatment rule in the original problem yielding the same minimax regret.

Comments.

1. Thus, strikingly, arbitrary reports in the original problem could be replaced by reports of fixed width taking only two values $[0, h_t]$ and $[1-h_t, 1]$ within an arm without changing the minimax regret value. Moreover, the decision rules are transferable.

2. The idea behind the Bernoulli equivalence for interval reports is very similar to the one behind Bernoulli equivalence established in Stoye (2009) for exactly observed outcomes. Restricting nature to Bernoulli reports cannot produce greater regret, so one direction is simple. For the other one, we proceed as follows. First, for a fixed canonical state s and every sample canonical report realization $X_{t,i} \in [0, 1-h_t]$, we construct an auxiliary binary variable with conditional success probability $X_{t,i}/(1-h_t)$. Averaging over the canonical report realizations then produces independent Bernoulli variables with success probabilities $p_t = \mathbb{E}[X_{t,i}]/(1-h_t)$. Using this construction, for any treatment rule of the Bernoulli canonical problem, we define a canonical treatment rule with the same treatment probability. Then, we construct a Bernoulli canonical state with the same treatment effect as s. Since regret function depends only on treatment effect and treatment assignment probability, regret from such canonical state and rule is equal to regret from the corresponding Bernoulli state and rule. Then, by using the fact that any canonical state could be transformed into the induced Bernoulli state, and by the fact that any Bernoulli rule can be transformed into the induced canonical rule, the other direction follows.

3. Note that similarly to the canonical equivalence Lemma 3.6 the result does not depend on a particular value of $h_0 + h_1$ as long as $h_0, h_1 < 1$. Thus, it is valid for the wide interval case as well.

Restricting nature to Bernoulli canonical problem as defined in Lemma 3.10 allows for dimension reduction on the side of nature's state space. However, to characterize minimax problem in terms of a finite number of parameters, we would also need a finite-dimensional characterization of the corresponding space of Bernoulli canonical treatment rules. Let $J_t = \sum_{i=1}^{N_t} B_{t,i}$ denote the success count in arm t, where $X_{t,i} = (1-h_t)B_{t,i}$. The next lemma shows that an arbitrary Bernoulli canonical rule $d : \{0,1\}^{N_0+N_1} \to [0,1]$ may be replaced with a symmetrized rule $\overline{d}$ which depends only on the success count J_t within each arm.

Lemma 3.11 (Treatment rule symmetrization)**.** *Suppose the nature is restricted to a Bernoulli canonical space $\mathcal{S}^{ber}_{h_0,h_1}$. Let $G = S_{N_0} \times S_{N_1}$ denote the finite group of within-arm permutations. For a sample of binary observations $x^n \in \{0,1\}^{N_0} \times \{0,1\}^{N_1}$, where $n = N_0 + N_1$, let $g \cdot x^n$ denote the corresponding permuted sample. In the Bernoulli canonical problem, for every treatment rule $d \in \mathcal{D}^{ber}_h$ there exists a symmetrized rule*

$$\overline{d}(x^n) = \frac{1}{|G|} \sum_{g \in G} d(g \cdot x^n),$$

which yields the same regret under every Bernoulli canonical state $s \in \mathcal{S}^{ber}_{h_0,h_1}$, and depends only on the number of successes J_t within each arm.

Lemma 3.11 shows that in the Bernoulli canonical problem we can restrict the treatment rule space to a matrix indexed treatment rules which depend only on the success count tuple (J_0, J_1):

$$\mathbf{d} = (d_{jk})_{\substack{0 \le j \le N_0 \\ 0 \le k \le N_1}}, \qquad 0 \le d_{jk} \le 1, \tag{3.21}$$

where each d_{jk} is the conditional treatment probability after j control successes and k treatment successes. Hence, the unconditional treatment probability becomes

$$D_{\mathbf{d}}(p_0, p_1) = \sum_{j=0}^{N_0} \sum_{k=0}^{N_1} d_{jk}\, \beta_{N_0,j}(p_0)\, \beta_{N_1,k}(p_1), \tag{3.22}$$

where $\beta_{n,j}(p) = \binom{n}{j} p^j (1-p)^{n-j}$ is a standard binomial probability of j successes.

Now we are ready to formulate the central minimax result for narrow intervals.

Theorem 3.12 (Implicit minimax result for narrow intervals)**.** *Fix $N_0, N_1 \ge 1$ and $h_0 + h_1 < 1$. Define*

$$\begin{aligned} b_+(p_0, p_1) &= [(1-h_1)p_1 - (1-h_0)p_0 + h_1]_+ , \\ b_-(p_0, p_1) &= [h_0 - (1-h_1)p_1 + (1-h_0)p_0]_+ . \end{aligned} \tag{3.23}$$

The value of the original minimax problem can be expressed as

$$\boxed{\begin{aligned} V_{N_0,N_1}(h_0, h_1) = \min_{\mathbf{d} \in [0,1]^{(N_0+1)\times(N_1+1)}} \max_{(p_0,p_1) \in [0,1]^2} \\ \max\{b_+(p_0,p_1)[1 - D_{\mathbf{d}}(p_0,p_1)], \\ b_-(p_0,p_1) D_{\mathbf{d}}(p_0,p_1)\}, \end{aligned}} \tag{3.24}$$

where $D_{\mathbf{d}}(p_0, p_1)$ is defined by (3.22). Moreover, the minimum in (3.24) is attained.

In particular, let $\mathbf{d}^*$ be any minimizing matrix. For an original sample of interval reports

$$W = \big((L_{0,i}, U_{0,i})_{i=1}^{N_0}, (L_{1,i}, U_{1,i})_{i=1}^{N_1}\big),$$

define the normalized canonical sample

$$\widetilde{x}^n = \big((\widetilde{x}_{0,i})_{i=1}^{N_0}, (\widetilde{x}_{1,i})_{i=1}^{N_1}\big) \in [0,1]^{N_0+N_1},$$

by applying the map $X_t(l,u)$ as defined in (3.11) to the original sample W component-wise and scaling it by $\frac{1}{1-h_t}$ in arm t, $\widetilde{x}_{t,i} = \frac{X_t(L_{t,i}, U_{t,i})}{1-h_t}$

Additionally, given the $\widetilde{x}^n$, for each treatment arm $t = 0, 1$ and each $j = 0, \ldots, N_t$, define the conditional probability of obtaining exactly j successes from N_t independent Bernoulli draws with success probabilities $\widetilde{x}_{t,1}, \ldots, \widetilde{x}_{t,N_t}$ as

$$q_{t,j}(W) = \sum_{\substack{I \subseteq \{1,\ldots,N_t\} \\ |I|=j}} \left(\prod_{i \in I} \widetilde{x}_{t,i} \right) \left(\prod_{i \notin I} (1 - \widetilde{x}_{t,i}) \right). \tag{3.25}$$

Then

$$\boxed{\delta_{\mathbf{d}^*}(W) = \sum_{j=0}^{N_0} \sum_{k=0}^{N_1} d^*_{jk}\, q_{0,j}(W) q_{1,k}(W)} \tag{3.26}$$

is a well-defined minimax-regret treatment rule for the original problem.

Comments.

1. The latent components in the Bernoulli canonical problem as described in Lemma 3.10 could be tuned by nature without changing the reporting distribution. As a result, for any treatment assignment rule (3.21) and fixed $p \in [0, 1]^2$, first term inside the max in (3.24) is the maximal possible regret from not assigning treatment, while the second is the maximal possible regret from assigning one. Nature's problem is then reduced to choosing p, while by Lemma 3.11 decision maker could be restricted to treatment rules of the form (3.21). Theorem 3.12 establishes that the corresponding optimal treatment rule always exists and attained. Then the resulting Bernoulli canonical matrix rule is transferred to the original problem by mechanisms established in lemmas 3.10 and 3.6, yielding 3.26.

2. Expressing the corresponding optimal treatment rule in a closed form is very non-trivial. Numerical exercises show that its entries follow different non-linear trajectories in response to changes in sample size and interval width, even upon assuming that those are equal between arms.

3. Numerical optimization of (3.24) could be demanding when the sample size is large. Luckily, for such cases the efficiency losses from using the plug-in rule (3.7) become very small, which is studiend in more detail in the next section.

4 Finite-sample simulation

In this section, we produce a finite-sample simulation to examine the performance of analytic minimax regret rule as derived in Theorems 3.12 and 3.7 relatively to the plug-in rule as defined in (3.7) and several heuristic alternatives. Recall that by Lemmas 3.6 and 3.10 restricting nature to the Bernoulli canonical problem does not change the minimax regret. Therefore we consider the Bernoulli restricted problem throughout the simulation. For simplicity, we focus on balanced samples with $N_0 = N_1 = N$, where $N \in \{1, 2, 4, 8, 16, 32\}$. However, we allow for different width across arms. In particular, to explore the impact of asymmetric reporting precision, we consider combined interval widths $H = h_0 + h_1 \in \{0, 0.1, \ldots, 2\}$. Then, for each value of H, we compare three allocations of interval widths across the treatment arms: equal widths, moderate asymmetry, and maximum asymmetry. Formally, for $a(H) = \min\{H, 2 - H\}$, these specifications correspond to

$$(h_0, h_1) = \begin{cases} (H/2, H/2), & \text{equal widths,} \\ (H/2 - a(H)/4,\ H/2 + a(H)/4), & \text{moderate asymmetry,} \\ (H/2 - a(H)/2,\ H/2 + a(H)/2), & \text{maximum asymmetry.} \end{cases} \tag{4.1}$$

This allows us to compare different levels of width asymmetry between arms while holding the combined width fixed.

For the wide intervals case ($H \geq 1$), we use the explicit minimax result of Theorem 3.7. For narrow intervals ($H < 1$), we calculate minimax regret numerically as follows.

1. First, we restrict nature to a finite grid of Bernoulli parameters $(p_0, p_1) \in [0,1]^2$ and solve the resulting optimisation problem on a grid. This yields a candidate treatment rule.
2. Holding the candidate treatment rule fixed, we search for Bernoulli parameter pairs outside the initial grid at which regret exceeds the value obtained in the first step.
3. We then add these Bernoulli parameter pairs to nature's grid and solve the original program again, obtaining an updated treatment rule. The procedure is repeated until no further improvement is found.

Note that the procedure above yields a minimax regret estimate which in principle can be lower than the actual value because the nature is restricted to a grid while the treatment rule calculation is exact. To bound the actual minimax value from above and verify the result, we first obtain the polynomial representation of regret

$$\max\{[q + h_1]_+[1 - D_{\mathbf{d}}(p_0, p_1)],\ [h_0 - q]_+ D_{\mathbf{d}}(p_0, p_1)\}, \tag{4.2}$$

where $q = (1 - h_1)p_1 - (1 - h_0)p_0$. Then we plug in the optimal treatment rule obtained from the iterative algorithm on a grid and maximize regret again over the whole $(p_0, p_1) \in [0,1]^2$. This yields an upper bounds since now nature's problem is exact while the decision maker is restricted to its best response to a grid-restricted nature. Then, comparing the upper bound with the lower bound from the iterative algorithm provides a measure of numerical accuracy. The maximal difference between the bounds across all configurations is approximately 3.3×10^{-4}.

4.1 Finite-sample performance of the analytic minimax regret rule

First, we examine the performance of the exact minimax regret treatment rule. Figure 1 plots minimax regret $V_{N,N}(h_0, h_1)$ against the combined width H for $N \in \{1, 2, 4, 8, 16, 32\}$. The three panels then correspond to equal widths, moderately asymmetric widths, and maximum feasible asymmetry, respectively. For comparison, we also include the identification lower bound $H/4$, which, by Theorem 3.4, is a minimax regret attained in an absence of sampling uncertainty.

Interestingly, for $H < 1$ minimax regret is generally above $H/4$. To explain that, note that, as demonstrated in Theorem 3.4, nature can always use observationally equivalent reports in two symmetric states yielding $H/4$. However, this forces the differences in outcome means to lie in the fixed interval width bound. Consequently, when intervals are narrow, nature prefers to rely more on sampling uncertainty by choosing difference between outcome means to be larger than observationally equivalent reports allow. However, for $H \geq 1$, Theorem 3.7 shows $V_{N,N}(h_0, h_1) = H/4$ for every sample size and width combination. As a result, all curves coincide with the identification lower bound. In particular, at $H = 2$ minimax regret problem corresponds to the unrestricted case covered in Theorem 2.1 with minimax value $\frac{1}{2}$ and optimal no-data rule $\delta = \frac{1}{2}$.

Comparing the three panels, we observe that for the minimax case the distribution of a fixed combined width between treatment arms has a relatively small effect on minimax regret for $H < 1$ and no difference at all for $H \geq 1$.

Note that the minimax regret values displayed in Figure 1 apply to the original interval-report problem by Lemmas 3.6 and 3.10. In the next subsections we report maximal regret of alternative rules relative to these values in the Bernoulli canonical problem. However, Lemmas 3.6 and 3.10 only establish minimax regret equivalence and optimal rule transferability. Consequently, non-minimax

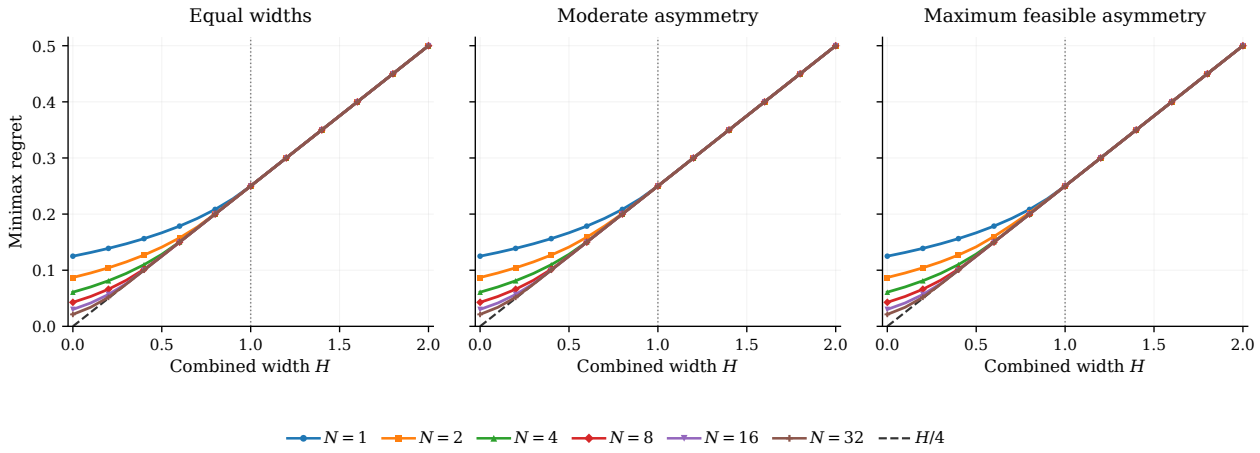


Figure 1: Finite-sample minimax regret for $N \in \{1, 2, 4, 8, 16, 32\}$ as a function of combined width $H = h_0 + h_1$. The panels correspond to the three width specifications in (4.1). The dashed line shows the identification lower bound $H/4$.

rules need not to maintain the same regret values between problems, and since nature is more restricted under Bernoulli canonical setting their maximal regret over all states in $\mathcal{S}_{h_0,h_1}$ could be larger.

4.2 Comparison with the plug-in rule

Next, I compare the minimax rules with the plug-in rule $\widehat{\delta}$ as defined in (3.7). Figure 2 reports its excess maximal regret relative to $V_{N,N}(h_0, h_1)$.

Under equal widths, the plug-in rule performs close to minimax throughout the considered configurations. With unequal widths, the difference can be considerably larger for narrow intervals, particularly when the sample is small. For example, at $N = 1$ and $H = 0.1$, excess maximal regret is approximately 0.061 under moderate asymmetry and 0.118 under maximum asymmetry. However, the difference becomes much smaller as the sample size increases. At $H = h_0 + h_1 \geq 1$, the performance of the plug-in rule coincides with the explicit minimax rule, as was demonstrated in Corollary 3.8.

4.3 Comparison with other treatment rules

Finally, I consider three heuristic alternatives: midpoint choice, interval dominance, and the no-data rule. For each alternative, I report excess maximal regret relative to the minimax value attained by the analytic minimax rule.

First alternative is the midpoint rule. Define the average midpoint as

$$\overline{M}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{L_{t,i} + U_{t,i}}{2}, \qquad t \in \{0, 1\}.$$

Then the midpoint rule chooses treatment if the midpoint of the treatment rule is larger.

$$\delta^{\text{mid}}(W) = \begin{cases} 1, & \overline{M}_1 > \overline{M}_0, \\ 1/2, & \overline{M}_1 = \overline{M}_0, \\ 0, & \overline{M}_1 < \overline{M}_0. \end{cases} \tag{4.3}$$

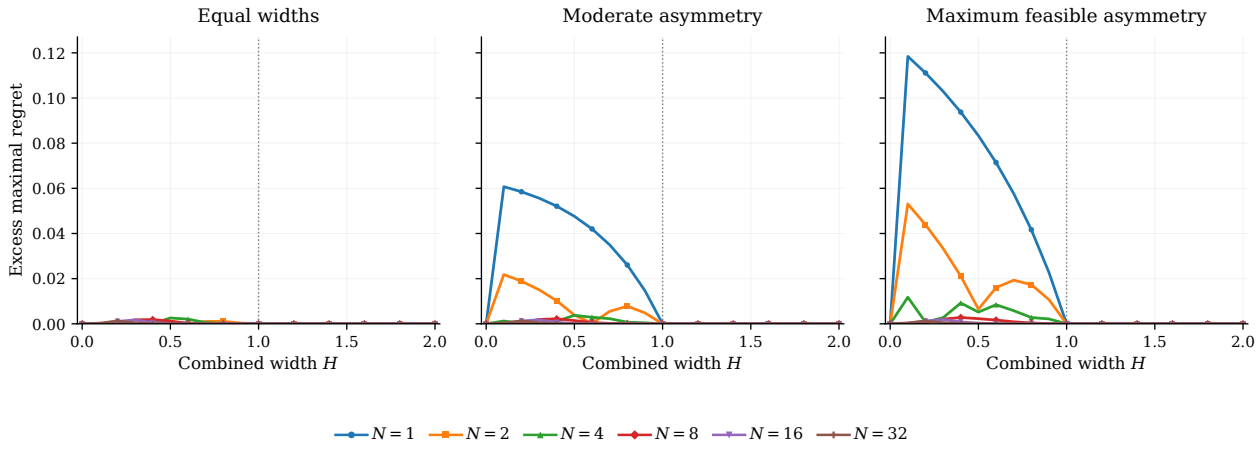


Figure 2: Excess maximal regret of the robust plug-in rule relative to $V_{N,N}(h_0, h_1)$ for $N \in \{1, 2, 4, 8, 16, 32\}$ and the three width specifications in (4.1).

Figure 3 shows that the rule performs similarly to the exact minimax rule when intervals are sufficiently narrow. However, its excess maximal regret increases considerably with interval width and sample size. Unlike the plug-in rule which explicitly addresses sampling uncertainty, more observations do not improve the worst-case performance of midpoint rule. However, at $H = 2$ every canonical report is the interval $[0, 1]$, so the midpoint rule assigns treatment with probability $1/2$ which coincides with the optimal minimax regret treatment rule at the unrestricted case.

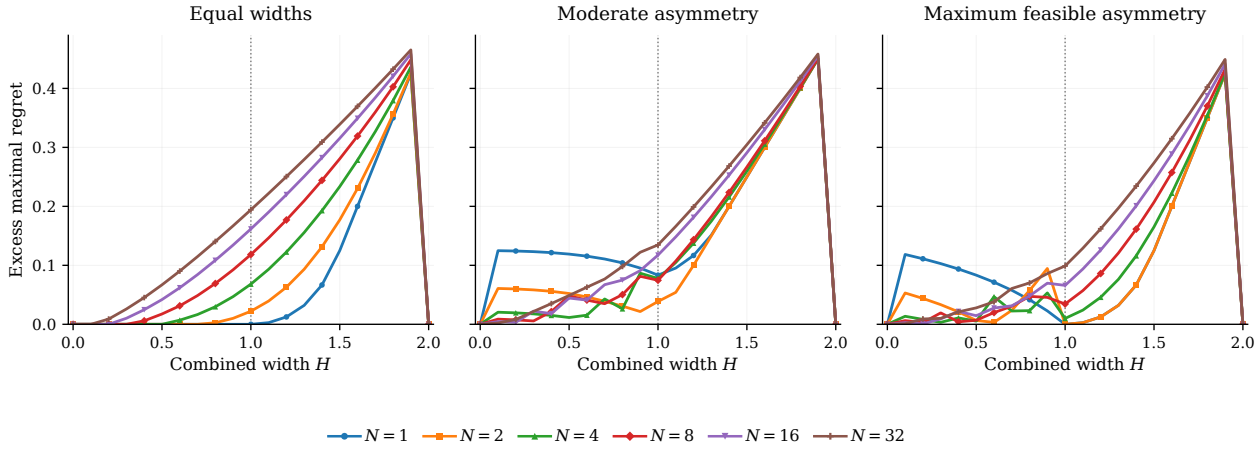


Figure 3: Excess maximal regret of midpoint choice relative to $V_{N,N}(h_0, h_1)$ for $N \in \{1, 2, 4, 8, 16, 32\}$ and the three width specifications in (4.1).

Second alternative is the interval dominance rule, which makes a decision with certainty when the estimated treatment effect interval is on the one side of zero, and randomizes otherwise. Formally,

$$\delta^{\text{dom}}(W) = \begin{cases} 1, & \widehat{a} > 0, \\ 0, & \widehat{b} < 0, \\ 1/2, & \widehat{a} \leq 0 \leq \widehat{b}. \end{cases} \tag{4.4}$$

Figure 4 shows that under narrow intervals its excess maximal regret can increase with combined width, particularly under asymmetric reporting precision. However, note that when $H \geq 1$ the

canonical estimated treatment-effect interval always contains zero. As a result, interval dominance rule is then reduced to equal randomization irrespective of sample size or width allocation. Consequently, all curves coincide in the $H \geq 1$ region and converge to zero at $H = 2$.

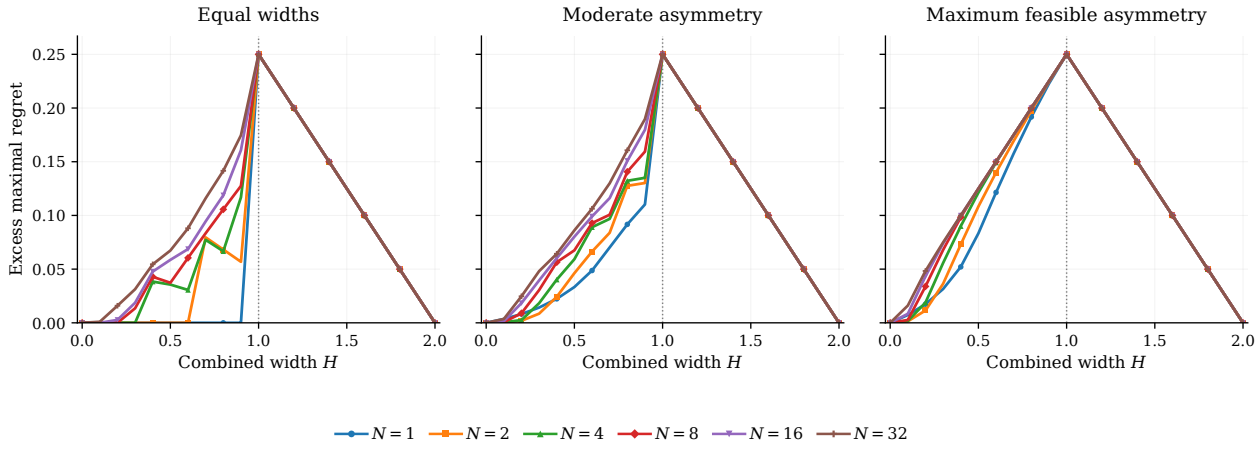


Figure 4: Excess maximal regret of the interval dominance rule relative to $V_{N,N}(h_0, h_1)$ for $N \in \{1, 2, 4, 8, 16, 32\}$ and the three width specifications in (4.1).

The final alternative is the no-data rule, which assigns treatment with probability $1/2$ irrespective of the sample. Note that its maximal regret is $1/2$ for every pair of widths, since in response nature can always choose a state with treatment effect 1. Consequently, difference in regret with the minimax rule decreases as H increases, while for a fixed narrow width it generally increases with N. The three panels are nearly identical for $H < 1$, since they are nearly identical for the minimax rule and here regret is fixed. Finally, for $H \geq 1$, excess regret is fixed at $1/2 - H/4$ in every specification and when the two rules coincide at $H = 2$.

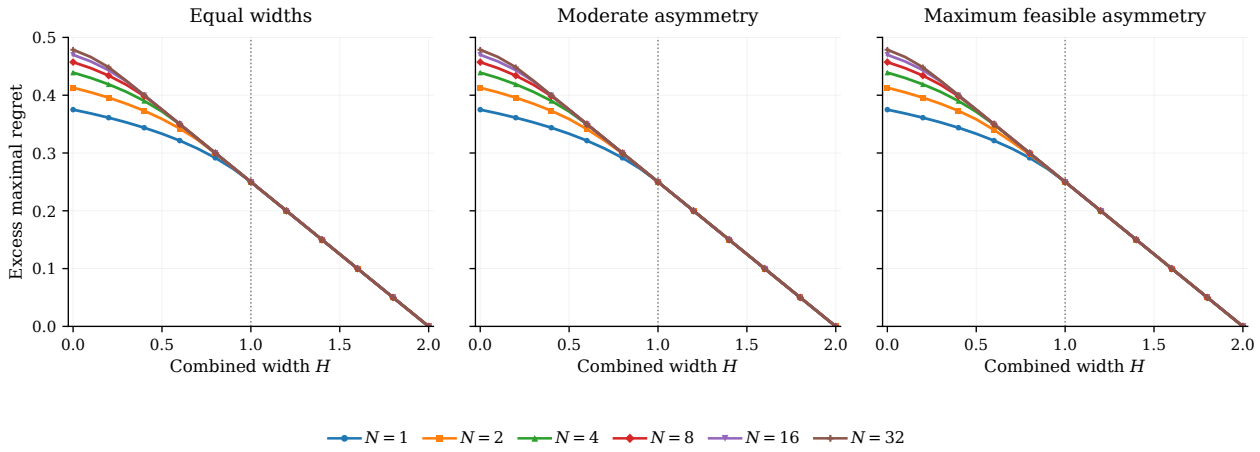


Figure 5: Excess maximal regret of the no-data rule $\delta(W) \equiv 1/2$ relative to $V_{N,N}(h_0, h_1)$ for $N \in \{1, 2, 4, 8, 16, 32\}$ and the three width specifications in (4.1).

Conclusion

This paper characterizes the minimax regret treatment rules in finite samples when the sample consists of interval reports rather than exact outcome realizations. First, we characterized the minimax regret treatment rule in a simplified problem when the reporting distribution is known, and then constructed its plug-in counterpart which yields minimax regret in the limiting experiment in the original problem. Then we have shown that in finite samples nature could be restricted to a canonical report problem, where the reports across both arms all have fixed maximal width, without changing the minimax regret. We also specified how minimax rules from the restricted problem could be mapped back to the original problem. This allowed us to characterize the finite sample minimax regret treatment rule when $h_0 + h_1 \geq 1$ explicitly. To characterize the minimax regret treatment rule when $h_0 + h_1 < 1$, we have shown that the nature can be further restricted to Bernoulli canonical intervals, where each arm produces just two possible interval values of maximal width, without changing the minimax regret value. Moreover, we also specified the mechanism of how the optimal rule from the Bernoulli canonical problem could be transferred back to the original problem. These two results allowed us to characterize the minimax regret treatment rule as a function of outcome of a finite-dimensional well-behaved optimization problem, thus yielding an implicit characterization of minimax regret in that region. Finally, we conducted a simulation experiment showing a clear superiority (in terms of minimax regret criterion) of the derived analytic treatment rule to a number of heuristic alternatives, with a plug-in rule being close behind.

As for the future work, an interesting extension of the original design would be to allow the outcomes to fall outside of the reported interval with some fixed probability. Such setting may arise naturally for example when outcomes are reported in confidence bands. Another promising direction would be to allow decision maker to costly improve interval precision and sample sizes, which creates an interesting sample size/measurement precision tradeoff.

References

[1] Christensen, T., Moon, H. R., and Schorfheide, F. (2026). Optimal decision rules when payoffs are partially identified. *Review of Economic Studies*, forthcoming. https://doi.org/10.1093/restud/rdag017.

[2] Guggenberger, P., Mehta, N., and Pavlov, N. (2026). Minimax regret treatment rules with finite samples when a quantile is the object of interest. *Econometric Theory*. https://doi.org/10.1017/S0266466626100498.

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

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

[5] 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(2), 415–442.

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

[7] Manski, C. F. (2007). Minimax-regret treatment choice with missing outcome data. *Journal of Econometrics*, 139(1), 105–115.

[8] Manski, C. F., and Mullahy, J. (2025). Utilitarian or quantile-welfare evaluation of health policy? *NBER Working Paper* No. 34247. https://doi.org/10.3386/w34247.

[9] Montiel Olea, J. L., Qiu, C., and Stoye, J. (2026a). Decision theory for treatment choice problems with partial identification. *Review of Economic Studies*, forthcoming. https://doi.org/10.1093/restud/rdag015.

[10] Montiel Olea, J. L., Qiu, C., and Stoye, J. (2026b). Asymptotics for treatment choice with partial identification. Working paper, arXiv:2608.09027.

[11] Stoye, J. (2007a). Minimax regret treatment choice with incomplete data and many treatments. *Econometric Theory*, 23(1), 190–199.

[12] Stoye, J. (2007b). Minimax regret treatment choice with finite samples and missing outcome data. *Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications (ISIPTA '07)*, Prague.

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

[14] Wald, A. (1950). *Statistical Decision Functions*. Wiley.

[15] Yata, K. (2023). Optimal decision rules under partial identification. Working paper, arXiv:2111.04926.

A Proofs

A.1 Unrestricted interval width

Proof of Theorem 2.1. For the upper bound, consider the no-data rule $\delta_0(W) = 1/2$. Under arbitrary $s \in \mathcal{S}_{1,1}$, regret (1.5) becomes

$$R(\delta_0, s) = \tfrac{1}{2}\left[\Delta(s)\right]_+ + \tfrac{1}{2}\left[-\Delta(s)\right]_+ = \tfrac{1}{2}|\Delta(s)|.$$

Since $\mu_t = \mathbb{E}_s[Y_t] \in [0,1]$, it follows that $|\Delta(s)| = |\mu_1 - \mu_0| \leq 1$. Therefore, nature would pick $|\Delta(s)| = 1$ which entails $R(\delta_0, s) = 1/2$. Thus, $V_{N_0,N_1}(1) \leq 1/2$

For the lower bound, take two symmetric states s^+ and s^-. Under s^+, let the outcomes be $Y_0 = 0$ and $Y_1 = 1$ almost surely. Under s^-, let the outcomes be $Y_0 = 1$ and $Y_1 = 0$ almost surely. Moreover, suppose that under both s^+ and s^- the reports are completely uninformative, $Z_t = (0,1)$ almost surely.

Now, take any rule $\delta \in \mathcal{D}$. Note that the sample of reports W has the same distribution under both states. Hence, two states are observationally equivalent and lead to the same treatment probability under δ. Denote it by

$$q = \mathbb{E}_{s^+}[\delta(W)] = \mathbb{E}_{s^-}[\delta(W)] \in [0,1].$$

Note that $\Delta(s^+) = 1$ and $\Delta(s^-) = -1$. Thus, the regret under two states is

$$R(\delta, s^+) = 1 - q, \qquad R(\delta, s^-) = q.$$

Thus, $\sup_{s \in \mathcal{S}_{1,1}} R(\delta, s) \geq \max\{1-q, q\} \geq \frac{1}{2}$.

Since the rule δ was arbitrary, it follows that $V_{N_0,N_1}(1) \geq 1/2$. Combining the upper and lower bounds of $V_{N_0,N_1}(1)$ yields the result. □

A.2 Known distributions of the reported intervals

Proof of Lemma 3.1. Note that since $L_t \leq Y_t \leq U_t$ almost surely, it follows that $\ell_t \leq \mu_t \leq u_t$. To show feasibility of all $\mu_t \in [\ell_t, u_t]$, take $\lambda_t \in [0,1]$ and construct

$$Y_t^{(\lambda_t)} = L_t + \lambda_t(U_t - L_t).$$

Clearly, such random outcome satisfies the coverage assumption while leaving the marginal Q_t laws unchanged. By varying λ_t, all possible means in the $[\ell_t, u_t]$ are attainable.

Next, since there is no restriction on the joint distribution of conditional outcomes, all pairs $(\mu_0, \mu_1) \in [\ell_0, u_0] \times [\ell_1, u_1]$ are feasible. Hence, any treatment effect $\Delta = \mu_1 - \mu_0$ in the $[a, b]$ interval is attainable. □

Proof of Theorem 3.2. The key observation is that the observed sample has fixed distribution $Q_0^{\otimes N_0} \otimes Q_1^{\otimes N_1}$ that cannot be affected by nature. In particular, while nature is free to choose the conditional distributions of Y_0 given Z_0 and Y_1 given Z_1, this does not change the known reporting distribution of (Q_0, Q_1). As a result, treatment probability induced by arbitrary rule $\delta \in \mathcal{D}$ is invariant under any distribution of $(Y_0|Q_0, Y_1|Q_1)$:

$$d = \mathbb{E}_{Q_0^{\otimes N_0} \otimes Q_1^{\otimes N_1}}[\delta(W)] \in [0,1].$$

Thus, the decision problem could be reduced to a choice of scalar d.

Conversely, since the sample distribution is fixed, nature can affect regret only through the treatment effect $\Delta = \mu_1 - \mu_0$. For a fixed d, regret at Δ is

$$r_d(\Delta) = [\Delta]_+ (1-d) + [-\Delta]_+ d.$$

This function is increasing when $\Delta \geq 0$ and decreasing when $\Delta \leq 0$. By Lemma 3.1, any treatment effect Δ in the $[a, b]$ interval is feasible. Hence, the worst-case regret becomes

$$\sup_{\Delta \in [a,b]} r_d(\Delta) = \max\{b_+(1-d), (-a)_+ d\}. \tag{A.1}$$

First, suppose $a = b = 0$. In that case, the treatment effect becomes deterministically zero, so any rule is minimax regret. Conversely, when $b > a \geq 0$ choosing $d = 1$ yields zero regret, and when $a < 0, b = 0$ choosing $d = 0$ gives zero regret.

Next, consider the case when $a < 0, b > 0$. Note that $b(1-d)$ is strictly decreasing in d, while $(-a) \cdot d$ is strictly increasing in d. Hence, $\max\{b(1-d), (-a) \cdot d\}$ attains minimum at the intersection

$$b(1-d) = (-a) \cdot d \quad \iff \quad d = \frac{b}{-a+b}.$$

Combining all cases together yields (3.3). Plugging the resulting optimal treatment rule into the maximal regret (A.1) then yields (3.4).

Finally, to prove (3.5), note that when $b \geq a \geq 0$ or $0 \geq b \geq a$ minimax regret is zero and so (3.5) holds trivially. For the remaining case when $b > 0 > a$, write $x = b$, $y = -a$. Since $(x-y)^2 \geq 0$,

$$\frac{xy}{x+y} \leq \frac{x+y}{4} = \frac{b-a}{4}.$$

Moreover,

$$\begin{aligned} b - a &= (u_1 - \ell_0) - (\ell_1 - u_0) \\ &= (u_1 - \ell_1) + (u_0 - \ell_0) \\ &= \mathbb{E}_{Q_1}[U_1 - L_1] + \mathbb{E}_{Q_0}[U_0 - L_0] \leq h_0 + h_1. \end{aligned}$$

Combining the two inequalities then gives $v(a, b) \leq \frac{(b-a)}{4} \leq \frac{h_0+h_1}{4}$. □

A.3 Unknown reporting distributions: a limiting minimax result

Proof of Theorem 3.4. We proceed in two steps: first, we find the finite sample minimax regret lower bound; second, we characterize the finite sample maximal regret of the plug-in rule, which gives the finite sample minimax regret upper bound. Combining the two yields (3.9). Then, by squeezing both sides, we obtain (3.8).

The proof of lower bound is mechanically the same as for the case with unrestricted width. First, take two symmetric states s^+ and s^-. Under both states, let $Z_t = ((1-h_t)/2, (1+h_t)/2)$ almost surely. Under s^+, let $Y_0 = (1-h_0)/2$ and $Y_1 = (1+h_1)/2$ almost surely. Under s^-, let $Y_0 = (1+h_0)/2$ and $Y_1 = (1-h_1)/2$ almost surely.

Now, take any rule $\delta \in \mathcal{D}$. Note that the sample of reports W has the same distribution under both states. Hence, two states are observationally equivalent and lead to the same treatment probability under δ. Denote it by

$$\pi = \mathbb{E}_{s^+}[\delta(W)] = \mathbb{E}_{s^-}[\delta(W)] \in [0, 1].$$

Note that $\Delta(s^+) = (h_0 + h_1)/2$ and $\Delta(s^-) = -(h_0 + h_1)/2$. Thus, the regret under two states is

$$R(\delta, s^+) = (1-\pi)\frac{h_0+h_1}{2}, \qquad R(\delta, s^-) = \pi \cdot \frac{h_0+h_1}{2}.$$

Therefore, $\sup_{s\in\mathcal{S}_{h_0,h_1}} R(\delta, s) \geq \max\{(1-\pi)\frac{h_0+h_1}{2}, \pi \cdot \frac{h_0+h_1}{2}\} \geq \frac{h_0+h_1}{4}$, which, given that δ was arbitrary, yields the lower bound.

For the upper bound, consider the $\widehat{\delta}(W)$ plug-in rule and fix an arbitrary state $s \in \mathcal{S}_{h_0,h_1}$ with treatment effect Δ, and an arbitrary sample realization W.

First, by Lemma 3.1, $\Delta \in [a, b]$. Also, note that $b \leq \widehat{b} + |b - \widehat{b}|$. Combining these two facts together yields

$$[\Delta]_+(1 - \widehat{\delta}(W)) \leq [b]_+ \cdot (1 - \widehat{\delta}(W)) \leq \left[\widehat{b}\right]_+ (1 - \widehat{\delta}(W)) + |b - \widehat{b}|.$$

and symmetrically

$$[-\Delta]_+ \cdot \widehat{\delta}(W) \leq [-a]_+ \cdot \widehat{\delta}(W) \leq [-\widehat{a}]_+ \, \widehat{\delta}(W) + |a - \widehat{a}|.$$

Combining the two cases then gives

$$[\Delta]_+ (1 - \widehat{\delta}(W)) + [-\Delta]_+ \cdot \widehat{\delta}(W) \leq \max\left\{ \left[\widehat{b}\right]_+ (1 - \widehat{\delta}(W)) + |b - \widehat{b}|, [-\widehat{a}]_+ \cdot \widehat{\delta}(W) + |a - \widehat{a}| \right\} \tag{A.2}$$

$$\leq \max\left\{ \left[\widehat{b}\right]_+ (1 - \widehat{\delta}(W)), [-\widehat{a}]_+ \cdot \widehat{\delta}(W) \right\} + \max\left\{|b - \widehat{b}|, |a - \widehat{a}|\right\} \tag{A.3}$$

Denote the first term in (A.3) by $v(\widehat{a}, \widehat{b})$. Then, plugging in $\widehat{\delta}(W) = d^*(\widehat{a}, \widehat{b}; d = \frac{1}{2})$ as defined in Theorem 3.2 gives

$$v(\widehat{a}, \widehat{b}) = \begin{cases} \dfrac{\widehat{b} \cdot (-\widehat{a})}{\widehat{b} - \widehat{a}}, & \widehat{b} > 0 > \widehat{a}, \\ 0, & otherwise. \end{cases}$$

Moreover,

$$v(\widehat{a}, \widehat{b}) \leq \frac{\widehat{b} - \widehat{a}}{4} \leq \frac{h_0 + h_1}{4},$$

where the first inequality uses the fact $(\widehat{b} + \widehat{a})^2 \geq 0$, and the second follows from

$$\widehat{b} - \widehat{a} = \frac{1}{N_1}\sum_{i=1}^{N_1}(U_{1,i} - L_{1,i}) + \frac{1}{N_0}\sum_{i=1}^{N_0}(U_{0,i} - L_{0,i}) \leq h_0 + h_1.$$

Thus, $v(\widehat{a}, \widehat{b}) \leq \frac{h_0+h_1}{4}$ for every sample. Taking expectation in (A.3) then yields

$$R(\widehat{\delta}, s) \leq \frac{h_0 + h_1}{4} + \mathbb{E}_s \max\{|a - \widehat{a}|, |b - \widehat{b}|\}. \tag{A.4}$$

To bound the remaining term, note that

$$\mathbb{E}_s \max\{|a - \widehat{a}|, |b - \widehat{b}|\} \leq \mathbb{E}_s \sqrt{(a - \widehat{a})^2 + (b - \widehat{b})^2} \leq \sqrt{Var(\widehat{a}) + Var(\widehat{b})},$$

where the first inequality uses the fact that $\max\{x, y\} \leq \sqrt{x^2 + y^2}$, and the second is Jensen's inequality. Furthermore,

$$Var(\widehat{a}) = Var(\widehat{\ell}_1) + Var(\widehat{u}_0) = \frac{Var(L_1)}{N_1} + \frac{Var(U_0)}{N_0} \leq \frac{1}{4}(\frac{1}{N_1} + \frac{1}{N_0}),$$

where the first two equalities use sampling independence, and the last inequality uses the fact that $U_0, L_1 \in [0, 1]$. Similarly,

$$Var(\widehat{b}) \leq \frac{1}{4}(\frac{1}{N_0} + \frac{1}{N_1}).$$

Thus,

$$\mathbb{E}_s \max\{|a - \widehat{a}|, |b - \widehat{b}|\} \leq \frac{1}{\sqrt{2}}\sqrt{\frac{1}{N_0} + \frac{1}{N_1}}.$$

Therefore, by using it in (A.4), it follows that

$$R(\widehat{\delta}, s) \leq \frac{h_0 + h_1}{4} + \frac{1}{\sqrt{2}}\sqrt{\frac{1}{N_0} + \frac{1}{N_1}},$$

which gives an upper bound for minimax regret since s was arbitrary.

Finally, by combining the upper and lower bounds and taking the limit, the result follows. □

A.4 Canonical reports and the explicit finite-sample minimax result

Proof of Lemma 3.5. Let $w = U - L \in [0, h_t]$. Note that

$$L = M - \frac{w}{2}, \quad U = M + \frac{w}{2}.$$

Consider four cases. First, let $h_t/2 \leq M \leq 1 - h_t/2$, then

$$0 \leq M - h_t/2 \leq 1 - h_t,$$

so no clipping occurs. Hence, $X_t(L, U) = M - \frac{h_t}{2}$, and The canonical interval has endpoints $M - h_t/2$ and $M + h_t/2$. Since $w \leq h_t$, it follows that

$$M - \frac{h_t}{2} \leq M - \frac{w}{2} = L \leq U = M + \frac{w}{2} \leq M + \frac{h_t}{2}.$$

Second, suppose that $M < h_t/2$, then (3.11) yields $X_t = 0$. The canonical interval has endpoints 0 and h_t. Note that the lower reporting endpoint $L \geq 0$, while

$$U = M + \frac{w}{2} \leq M + \frac{h_t}{2} < h_t.$$

Third, if $M > 1 - h_t/2$, then (3.11) yields $X_t = 1 - h_t$. The canonical interval has endpoints $1 - h_t$ and 1. Note that the upper endpoint $U \leq 1$, while

$$L = M - \frac{w}{2} \geq M - \frac{h_t}{2} > 1 - h_t.$$

Finally, when $h_t = 0$ the width restriction forces $L = U$, and (3.11) yields $X_t = L = U$. This exhausts all cases and the (3.12) follows. □

Proof of Lemma 3.6. First, we define measurable maps between the two observation spaces and their states. Define $n = N_0 + N_1$ and

$$j_t(x) = (x, x + h_t), \qquad x \in [0, 1 - h_t], \quad t = 0, 1.$$

Next, let J apply j_t componentwise to the N_t observations in arm t. Similarly, let G_{h_0,h_1} apply the map $X_t(L, U)$ from Lemma 3.5 componentwise. Note that,

$$G_{h_0,h_1}(J(x^n)) = x^n \qquad \text{for every } x^n \in [0, 1 - h_0]^{N_0} \times [0, 1 - h_1]^{N_1}.$$

Let V_{orig} be the finite sample minimax value of the original problem, and V_{can} be the finite sample minimax value for the problem where nature is restricted to canonical reports.

First, we show that $V_{\text{orig}} \geq V_{\text{can}}$. Take any canonical state s'. Each distribution of (Y_t, X_t) satisfies $X_t \leq Y_t \leq X_t + h_t$ almost surely. Fix Y_t, and replace the X_t by original report $Z_t = j_t(X_t)$. This report has width h_t, lies within $[0, 1]$, and by construction contains the outcome almost surely. Thus, it is a valid original state. Call it $\iota(s')$.

Next, take any original problem treatment rule δ. By restricting its use to the samples of the form $J(x^n)$ we obtain a valid canonical rule

$$d_\delta(x^n) = \delta(J(x^n)).$$

By construction, under the transformed state $\iota(s')$, W is distributed exactly as $J(X^n)$ under the canonical state s', which gives

$$\mathbb{E}_{\iota(s')}[\delta(W)] = \mathbb{E}_{s'}[\delta(J(X^n))] = \mathbb{E}_{s'}[d_\delta(X^n)].$$

Note that this transformation leaves Y_t unchanged, and hence the treatment effect Δ stays the same. Thus, every input in the regret formula (1.5) agrees, yielding

$$R(\delta, \iota(s')) = R(d_\delta, s'), \qquad \forall s' \in \mathcal{S}^{can}_{h_0,h_1}.$$

Since the canonical states embed into $\mathcal{S}_{h_0,h_1}$, it follows that

$$\begin{aligned}\sup_{s \in \mathcal{S}_{h_0,h_1}} R(\delta, s) &\geq \sup_{s' \in \mathcal{S}^{can}_{h_0,h_1}} R(\delta, \iota(s')) \\ &= \sup_{s' \in \mathcal{S}^{can}_{h_0,h_1}} R(d_\delta, s') \geq V_{\text{can}}.\end{aligned}$$

Since the original rule was arbitrary, it follows that $V_{\text{orig}} \geq V_{\text{can}}$.

To prove the opposite inequality $V_{\text{orig}} \leq V_{\text{can}}$, fix any canonical rule d and apply it to transformed original samples $G_{h_0,h_1}(W)$. This creates a valid original-data rule

$$\widetilde{\delta}(W) = d(G_{h_0,h_1}(W)). \tag{A.5}$$

Next, take any original state $s \in \mathcal{S}_{h_0,h_1}$, and let $P_{t,s}$ denote the corresponding arm-t joint distribution of (Y_t, L_t, U_t), and let $Q_{t,s}$ be its marginal distribution of (L_t, U_t). Define the deterministic function

$$\Phi_t(y, l, u) = (y, X_t(l, u)).$$

Applying it to (Y_t, L_t, U_t) creates a pushforward distribution for arm t

$$P'_{t,s} = P_{t,s} \circ \Phi_t^{-1}.$$

By Lemma 3.5, $[L_t, U_t] \subseteq j_t(X_t(L_t, U_t))$ and hence $Y_t \in j_t(X_t(L_t, U_t))$. Thus, $X_t(L_t, U_t)$ is an admissible canonical report with distribution $Q'_{t,s} = Q_{t,s} \circ X_t^{-1}$. Call the corresponding canonical state $T(s)$, with joint law $P'_{t,s}$.

Note that Φ_t does not change the outcome distribution. Therefore, treatment effect Δ stays the same. At the same time, applying the transformation componentwise preserves the product sampling law:

$$\mathcal{L}_s(G_{h_0,h_1}(W)) = \mathcal{L}_{T(s)}(X^n).$$

Hence, by using (A.5),

$$\mathbb{E}_s[\widetilde{\delta}(W)] = \mathbb{E}_s[d(G_{h_0,h_1}(W))] = \mathbb{E}_{T(s)}[d(X^n)].$$

Thus, since under the transformation both treatment effect and treatment probability stay the same, it follows that regret is unchanged,

$$R(\widetilde{\delta}, s) = R(d, T(s)), \qquad \forall s \in \mathcal{S}_{h_0,h_1}.$$

Consequently,

$$V_{\text{orig}} \leq \sup_{s \in \mathcal{S}_{h_0,h_1}} R(\widetilde{\delta}, s) \tag{A.6}$$

$$= \sup_{s \in \mathcal{S}_{h_0,h_1}} R(d, T(s)) \tag{A.7}$$

$$\leq \sup_{s' \in \mathcal{S}^{can}_{h_0,h_1}} R(d, s'). \tag{A.8}$$

Taking the infimum over canonical rules yields $V_{\text{orig}} \leq V_{\text{can}}$. Combining the two inequalities establishes equality of the minimax values. Finally, if the canonical rule d is minimax, then the original rule $\widetilde{\delta}$ defined in (A.5) is also minimax. □

Corollary A.1 (Canonical transfer of treatment rules)**.** *Fix $h_0, h_1 \in [0, 1]$ and an original state $s \in \mathcal{S}_{h_0,h_1}$. The componentwise transformation G_{h_0,h_1} defined in the proof of Lemma 3.6 induces a canonical state $T(s) \in \mathcal{S}^{can}_{h_0,h_1}$ satisfying*

$$\mathcal{L}_s(G_{h_0,h_1}(W)) = \mathcal{L}_{T(s)}(X^n), \qquad \Delta(s) = \Delta(T(s)).$$

Consequently, for every canonical treatment rule d, the original-data rule $\widetilde{\delta}(W) = d(G_{h_0,h_1}(W))$ satisfies

$$\pi_{\widetilde{\delta}}(s) = \pi_d(T(s)), \qquad R(\widetilde{\delta}, s) = R(d, T(s)).$$

Proof. Take the original arm-t joint distribution $P_{t,s}$ of (Y_t, L_t, U_t) and define

$$\Phi_t(y, l, u) = (y, X_t(l, u)), \qquad P'_{t,s} = P_{t,s} \circ \Phi_t^{-1}.$$

By Lemma 3.5, the interval $[X_t(l, u), X_t(l, u) + h_t]$ contains $[l, u]$ and hence the unchanged outcome y. Thus $P'_{t,s}$ is a valid canonical distribution. The transformation leaves the outcome distribution unchanged, so the treatment effect is the same. Moreover, applying the same measurable transformation componentwise to i.i.d. observations within each arm and to the independent arm samples preserves their product sampling distribution. Therefore,

$$\mathcal{L}_s(G_{h_0,h_1}(W)) = \mathcal{L}_{T(s)}(X^n).$$

For the rule $\widetilde{\delta}(W) = d(G_{h_0,h_1}(W))$, this gives

$$\mathbb{E}_s[\widetilde{\delta}(W)] = \mathbb{E}_{T(s)}[d(X^n)].$$

Since both treatment effect and assignment probability are unchanged, equality of regret follows. □

Proof of Theorem 3.7. As was demonstrated in Lemma 3.6, it suffices to show that there exists a canonical rule that yields $(h_0 + h_1)/4$ minimax regret in the restricted problem, and that the rule can be transferred to the original problem by applying the canonical maps componentwise.

First, take any canonical state $s' \in \mathcal{S}^{can}_{h_0,h_1}$, fix $n = N_0 + N_1$, and let $H = h_0 + h_1$. Define a function d^* on $[0, 1 - h_0]^{N_0} \times [0, 1 - h_1]^{N_1}$ as follows:

$$d^*(X^n) = \frac{1}{2} + \frac{\overline{X}_1 - \overline{X}_0 + (h_1 - h_0)/2}{H}. \tag{A.9}$$

Note that d^* is a valid canonical rule. To see it, observe that each $X_{t,i} \in [0, 1 - h_t]$, and so are their respective sample averages. Hence,

$$-1 + \frac{H}{2} \leq \overline{X}_1 - \overline{X}_0 + \frac{h_1 - h_0}{2} \leq 1 - \frac{H}{2}.$$

Since $H \geq 1$, it follows that $1 - H/2 \leq H/2$, and therefore $0 \leq d^*(X^n) \leq 1$ for every sample realization.

Second, we show that d^* as defined in (A.9) attains the $(h_0 + h_1)/4$ regret. Let

$$z = q + \frac{h_1 - h_0}{2}.$$

Note that

$$\pi_{d^*} = \mathbb{E}[d^*(X^n)] = \frac{1}{2} + \frac{z}{H}. \tag{A.10}$$

Moreover, by (3.16), $\Delta \in [z - H/2, z + H/2]$. Since $|z| \leq 1 - H/2 \leq H/2$, both endpoints cross zero. Hence, for an arbitrary canonical state s', the regret is

$$\begin{aligned} R(d^*, s') &\leq \max\left\{\left(z + \frac{H}{2}\right)\left(\frac{1}{2} - \frac{z}{H}\right), \left(\frac{H}{2} - z\right)\left(\frac{1}{2} + \frac{z}{H}\right)\right\} \\ &= \frac{H}{4} - \frac{z^2}{H} \leq \frac{H}{4}. \end{aligned}$$

Since s' was arbitrary, the proposed canonical rule has maximal regret at most $H/4$. Recall that by Lemma 3.6 and Theorem 3.4 the minimax value is at least $H/4$. Hence, d^* is minimax in the canonical problem and $V_{N_0,N_1}(h_0, h_1) = H/4$. Finally, Lemma 3.6 transfers it to the original problem. □

Proof of Corollary 3.8. The proof uses the same approach as the second part of the proof of Theorem 3.7. First, note that by construction, for every sample realization W,

$$-1 \leq \widehat{a} \leq \widehat{b} \leq 1, \qquad \widehat{b} - \widehat{a} \leq H.$$

We first show that when $H \geq 1$ the plug-in rule satisfies

$$\widehat{\delta}(W) \geq \frac{\widehat{b}}{H}, \qquad 1 - \widehat{\delta}(W) \geq -\frac{\widehat{a}}{H}. \tag{A.11}$$

First, suppose that $\widehat{a} < 0 < \widehat{b}$. By (3.3) and the fact that $\widehat{b} - \widehat{a} \leq H$,

$$\widehat{\delta}(W) = \frac{\widehat{b}}{\widehat{b} - \widehat{a}} \geq \frac{\widehat{b}}{H},$$

and similarly,

$$1 - \widehat{\delta}(W) = \frac{-\widehat{a}}{\widehat{b} - \widehat{a}} \geq -\frac{\widehat{a}}{H}.$$

When the estimated treatment-effect interval lies entirely on one side of zero, the inequalities follow directly from (3.3), since $H \geq 1$ and $\widehat{a}, \widehat{b} \in [-1, 1]$. They also hold trivially when $\widehat{a} = \widehat{b} = 0$.

Next, fix an arbitrary state $s \in \mathcal{S}_{h_0,h_1}$. Recall that

$$a = \ell_1 - u_0, \qquad b = u_1 - \ell_0,$$

and that $\mathbb{E}_s[\widehat{a}] = a$ and $\mathbb{E}_s[\widehat{b}] = b$. Taking expectations in (A.11) yields

$$\pi_{\widehat{\delta}}(s) \geq \frac{b}{H}, \qquad 1 - \pi_{\widehat{\delta}}(s) \geq -\frac{a}{H}.$$

Now, suppose that $\Delta(s) \geq 0$. Note that $\Delta(s) \leq b$, hence

$$R(\widehat{\delta}, s) = \Delta(s)(1 - \pi_{\widehat{\delta}}(s)) \leq b(1 - \pi_{\widehat{\delta}}(s)) \leq b\left(1 - \frac{b}{H}\right) \leq \frac{H}{4}, \tag{A.12}$$

where the last inequality follows from

$$b\left(1 - \frac{b}{H}\right) = \frac{H}{4} - \frac{(b - H/2)^2}{H}.$$

Conversely, when $\Delta(s) < 0$, $\Delta(s) \geq a$. Therefore,

$$R(\widehat{\delta}, s) = (-\Delta(s))\pi_{\widehat{\delta}}(s) \leq (-a)\pi_{\widehat{\delta}}(s) \leq (-a)\left(1 + \frac{a}{H}\right) \leq \frac{H}{4}. \tag{A.13}$$

Thus, for every $s \in \mathcal{S}_{h_0,h_1}$, $R(\widehat{\delta}, s) \leq H/4$, implying

$$\sup_{s \in \mathcal{S}_{h_0,h_1}} R(\widehat{\delta}, s) \leq \frac{H}{4}.$$

Since Theorem 3.4 establishes the same minimax lower bound, the result follows. □

A.5 Bernoulli coarsening and the implicit finite-sample minimax result

Proof of Lemma 3.10. Define $n = N_0 + N_1$ and denote the two minimax values by V_{can} and V_{Ber}.

First, $V_{\text{can}} \geq V_{\text{Ber}}$. Restricting an arbitrary canonical rule $\delta \in \mathcal{D}^{can}_{h_0,h_1}$ to the scaled Bernoulli support induces a Bernoulli rule $d_\delta(b) = \delta(((1 - h_t)b_{t,i})_{t,i})$. Note that for every $s \in \mathcal{S}^{ber}_{h_0,h_1}$ the assignment probabilities and regrets of d_δ and δ agree by construction. Hence,

$$\begin{aligned}\sup_{s \in \mathcal{S}^{can}_{h_0,h_1}} R(\delta, s) &\geq \sup_{s \in \mathcal{S}^{ber}_{h_0,h_1}} R(\delta, s) \\ &= \sup_{s \in \mathcal{S}^{ber}_{h_0,h_1}} R(d_\delta, s) \geq V_{\text{Ber}}.\end{aligned}$$

Because δ was arbitrary, we conclude $V_{\text{can}} \geq V_{\text{Ber}}$.

Second, $V_{\text{can}} \leq V_{\text{Ber}}$. Conditional on a canonical sample $X^n = ((X_{0,i})_{i=1}^{N_0}, (X_{1,i})_{i=1}^{N_1})$, construct auxiliary uniforms $U_{t,i} \sim \text{Uniform}[0, 1]$ that are independent of the canonical sample and of one another, and define

$$B_{t,i} = \mathbf{1}\left\{U_{t,i} \leq \frac{X_{t,i}}{1 - h_t}\right\}.$$

Note that conditional on observed $X_{t,i}$, each resulting $B_{t,i}$ is an independent Bernoulli variable with parameter $\frac{X_{t,i}}{1-h_t}$. Next, fix any canonical state s and let $p_t = \mathbb{E}_s[X_t]/(1-h_t)$. For any binary vector $b \in \{0,1\}^n$ by using independence we get

$$\mathbb{P}_s(B^n = b \mid X^n) = \prod_{t=0}^{1} \prod_{i=1}^{N_t} \left(\frac{X_{t,i}}{1-h_t}\right)^{b_{t,i}} \left(1 - \frac{X_{t,i}}{1-h_t}\right)^{1-b_{t,i}}.$$

Moreover,

$$\begin{aligned}\mathbb{P}_s(B^n = b) &= \prod_{t=0}^{1} \prod_{i=1}^{N_t} \mathbb{E}_s\left[\left(\frac{X_{t,i}}{1-h_t}\right)^{b_{t,i}} \left(1 - \frac{X_{t,i}}{1-h_t}\right)^{1-b_{t,i}}\right] \\ &= \prod_{t=0}^{1} \prod_{i=1}^{N_t} p_t^{b_{t,i}} (1-p_t)^{1-b_{t,i}}.\end{aligned}$$

where the first equality uses independence, and the second follows from $b_{t,i} \in \{0,1\}$ and $\mathbb{E}_s[X_{t,i}]/(1-h_t) = p_t$. Thus, we conclude that the joint distribution of the auxiliary sample B^n is

$$\mathcal{L}_s(B^n) = \text{Bernoulli}(p_0)^{\otimes N_0} \otimes \text{Bernoulli}(p_1)^{\otimes N_1}.$$

Then, we construct a Bernoulli state with the same treatment effect as s. Define a state $s^B \in \mathcal{S}^{ber}_{h_0,h_1}$ with reports $X^B_t = (1-h_t)B^B_t$, where $B^B_t \sim \text{Bernoulli}(p_t)$, and constant latent components $A^B_t = \mathbb{E}_s[A_t] = a_t$. Hence, $Y^B_{t,i} = (1-h_t)B^B_{t,i} + h_t a_t$. Therefore,

$$\mathbb{E}_{s^B}[Y^B_t] = (1-h_t)p_t + h_t a_t = \mathbb{E}_s[X_t + h_t A_t] = \mathbb{E}_s[Y_t],$$

which yields $\Delta(s^B) = \Delta(s)$, as desired. Thus, Bernoulli state s^B has the same treatment effect as s. Moreover, by construction, the normalized reporting sample B^n from s^B has the same distribution as auxilary sample B^n under s.

Next, fix any Bernoulli rule $d : \{0,1\}^n \to [0,1]$ and define the following treatment rule on $[0, 1-h_0]^{N_0} \times [0, 1-h_1]^{N_1}$:

$$\widetilde{d}(x^n) = \mathbb{E}[d(B^n) \mid X^n = x^n] = \sum_{b \in \{0,1\}^n} d(b) \prod_{t=0}^{1} \prod_{i=1}^{N_t} \left(\frac{x_{t,i}}{1-h_t}\right)^{b_{t,i}} \left(1 - \frac{x_{t,i}}{1-h_t}\right)^{1-b_{t,i}}. \tag{A.14}$$

Note that the weights are nonnegative and sum up to one. Therefore, $\widetilde{d} \in [0,1]$ and constitutes a valid canonical rule. Moreover, by iterated expectation

$$\pi_{\widetilde{d}}(s) = \mathbb{E}_s[\widetilde{d}(X^n)] = \mathbb{E}_s[d(B^n)].$$

Since the Bernoulli sample B^n under s^B has the same distribution as the auxiliary sample B^n under s, it also follows

$$\pi_{\widetilde{d}}(s) = \mathbb{E}_s[d(B^n)] = \mathbb{E}_{s^B}[d(B^n)] = \pi_d(s^B)$$

Finally, recall that regret depends only on treatment effect and treatment probability. Since under $\widetilde{d}$ and s those are equal to their counterparts under d and s^B, we conclude

$$R(\widetilde{d}, s) = R(d, s^B), \qquad \forall s \in \mathcal{S}^{can}_{h_0,h_1}.$$

Since we can construct s^B for every $s \in \mathcal{S}^{can}_{h_0,h_1}$, conclude for arbitrary rule $d \in \mathcal{D}^{ber}_{h_0,h_1}$ and induced canonical rule $\widetilde{d}$:

$$\sup_{s \in \mathcal{S}^{can}_{h_0,h_1}} R(\widetilde{d}, s) \leq \sup_{r \in \mathcal{S}^{ber}_{h_0,h_1}} R(d, r).$$

Since d was arbitrary, the latter yields $V_{\text{can}} \leq V_{\text{Ber}}$ and hence $V_{\text{can}} = V_{\text{Ber}}$, as desired.

For the second part of the lemma, let $d^* \in \mathcal{D}^{ber}_{h_0,h_1}$ be the minimax treatment rule in the Bernoulli problem. By construction, its induced canonical rule $\widetilde{d}^*$ as defined in (A.14) has no greater worst-case regret, and therefore is minimax in the canonical problem. Lemma 3.6 then transfers it to the original problem, which concludes the proof. □

Corollary A.2 (Bernoulli coarsening of count rules)**.** *Fix $0 \leq h_0, h_1 < 1$ and a Bernoulli matrix rule $\mathbf{d} = (d_{jk})$ as in (3.21). For an original sample W, define $\widetilde{x}_{t,i} = X_t(L_{t,i}, U_{t,i})/(1 - h_t)$ and $q_{t,j}(W)$ by (3.25). Then*

$$\delta_{\mathbf{d}}(W) = \sum_{j=0}^{N_0} \sum_{k=0}^{N_1} d_{jk} q_{0,j}(W) q_{1,k}(W)$$

is a measurable treatment rule taking values in $[0, 1]$. Moreover, for every original state $s \in \mathcal{S}_{h_0,h_1}$, there exists a Bernoulli state $s^B \in \mathcal{S}^{ber}_{h_0,h_1}$ such that

$$\Delta(s^B) = \Delta(s), \qquad \pi_{\delta_{\mathbf{d}}}(s) = \pi_{\mathbf{d}}(s^B), \qquad R(\delta_{\mathbf{d}}, s) = R(\mathbf{d}, s^B).$$

Proof. Fix an arbitrary original state $s \in \mathcal{S}_{h_0,h_1}$ and apply the canonical enlargement from Corollary A.1. Conditional on the original sample W, construct independent auxiliary Bernoulli variables

$$B_{t,i} \mid W \sim \text{Bernoulli}(\widetilde{x}_{t,i}), \qquad J^B_t = \sum_{i=1}^{N_t} B_{t,i}.$$

First, we show that $q_{t,j}(W)$ is the conditional probability of observing exactly j successes in the auxiliary sample from treatment arm t. For an arbitrary subset $I \subseteq \{1, \ldots, N_t\}$, conditional independence yields

$$\begin{aligned}\mathbb{P}_s(B_{t,i} = 1 \text{ for } i \in I,\, B_{t,i} = 0 \text{ for } i \notin I \mid W) \\ = \left(\prod_{i \in I} \widetilde{x}_{t,i}\right) \left(\prod_{i \notin I} (1 - \widetilde{x}_{t,i})\right).\end{aligned}$$

Note that the event $\{J^B_t = j\}$ is a disjoint union of the above events over all subsets I with exactly j elements. Hence,

$$\begin{aligned}\mathbb{P}_s(J^B_t = j \mid W) &= \sum_{\substack{I \subseteq \{1,\ldots,N_t\} \\ |I| = j}} \left(\prod_{i \in I} \widetilde{x}_{t,i}\right) \left(\prod_{i \notin I} (1 - \widetilde{x}_{t,i})\right) \\ &= q_{t,j}(W).\end{aligned}$$

Moreover, since the auxiliary draws are independent across treatment arms, it follows that

$$\mathbb{P}_s(J^B_0 = j, J^B_1 = k \mid W) = q_{0,j}(W) q_{1,k}(W).$$

Thus, by construction,

$$\delta_{\mathbf{d}}(W) = \mathbb{E}_s[d_{J^B_0, J^B_1} \mid W].$$

Note that $q_{t,j}(W) \geq 0$ and $\sum_{j=0}^{N_t} q_{t,j}(W) = 1$, since the events $\{J_t^B = j\}$ are mutually exclusive and exhaustive. Therefore, $\delta_{\mathbf{d}}(W)$ is a measurable convex combination of the entries of $\mathbf{d}$ and takes values in $[0,1]$.

Next, we show that the reconstructed rule has the same treatment probability as the Bernoulli matrix rule under a corresponding Bernoulli state. Define

$$p_t(s) = \frac{\mathbb{E}_s[X_t(L_t, U_t)]}{1 - h_t}, \qquad t = 0, 1.$$

Recall that the canonical transformation preserves i.i.d. sampling within each arm and independence between the two arm samples. Using this fact together with conditional independence of the auxiliary draws, for every $b \in \{0,1\}^{N_0+N_1}$ we obtain

$$\begin{aligned}\mathbb{P}_s(B^n = b) &= \mathbb{E}_s\left[\prod_{t=0}^{1}\prod_{i=1}^{N_t} \widetilde{x}_{t,i}^{b_{t,i}}(1-\widetilde{x}_{t,i})^{1-b_{t,i}}\right] \\ &= \prod_{t=0}^{1}\prod_{i=1}^{N_t} p_t(s)^{b_{t,i}}(1-p_t(s))^{1-b_{t,i}}.\end{aligned}$$

Hence, the joint distribution of the auxiliary sample is

$$\mathcal{L}_s(B^n) = \text{Bernoulli}(p_0(s))^{\otimes N_0} \otimes \text{Bernoulli}(p_1(s))^{\otimes N_1}.$$

By iterated expectations,

$$\begin{aligned}\pi_{\delta_{\mathbf{d}}}(s) &= \mathbb{E}_s[\delta_{\mathbf{d}}(W)] \\ &= \mathbb{E}_s[\mathbb{E}_s[d_{J_0^B, J_1^B} \mid W]] \\ &= \mathbb{E}_s[d_{J_0^B, J_1^B}] \\ &= D_{\mathbf{d}}(p_0(s), p_1(s)).\end{aligned}$$

Finally, we construct a Bernoulli state with the same treatment effect as s. By Corollary A.1, the canonical enlargement leaves the outcomes unchanged and gives the representation $Y_t = X_t + h_t A_t$, where $A_t \in [0,1]$. Let $a_t = \mathbb{E}_s[A_t]$. As in the proof of Lemma 3.10, define a Bernoulli state $s^B \in \mathcal{S}^{ber}_{h_0,h_1}$ with reports $X_t^B = (1-h_t)B_t^B$, where $B_t^B \sim \text{Bernoulli}(p_t(s))$, and constant latent components $A_t^B = a_t$. Then

$$\mathbb{E}_{s^B}[Y_t^B] = (1-h_t)p_t(s) + h_t a_t = \mathbb{E}_s[Y_t],$$

which yields $\Delta(s^B) = \Delta(s)$.

Moreover, under s^B, the Bernoulli count rule has treatment probability

$$\pi_{\mathbf{d}}(s^B) = D_{\mathbf{d}}(p_0(s), p_1(s)) = \pi_{\delta_{\mathbf{d}}}(s).$$

Thus, both treatment effect and treatment probability agree under the two states and their respective rules. By (1.5),

$$R(\delta_{\mathbf{d}}, s) = R(\mathbf{d}, s^B).$$

Since s was arbitrary, the result follows. □

Proof of Lemma 3.11. First, note that, for every $x^n \in \{0,1\}^{N_0} \times \{0,1\}^{N_1}$, $\overline{d}(x^n)$ averages over the values of a valid Bernoulli canonical treatment rules. Hence, it is measurable and belongs to $[0,1]$,

thus constituting a valid Bernoulli canonical treatment rule itself. Moreover, each permutation leaves the distribution of a Bernoulli canonical sample unchanged, hence

$$\mathbb{E}_s[\overline{d}(B^n)] = \frac{1}{|G|}\sum_{g\in G}\mathbb{E}_s[d(g\cdot B^n)] = \mathbb{E}_s[d(B^n)].$$

Since the treatment assignment probability under every Bernoulli canonical state is the same, d and $\overline{d}$ yield equal regret, as desired.

For the final part of the lemma, consider an arbitrary permutation $g_0 \in G$. For an arbitrary sample $x^n \in \{0,1\}^{N_0}\times\{0,1\}^{N_1}$, it follows that

$$\overline{d}(g_0\cdot x^n) = \frac{1}{|G|}\sum_{g\in G} d(g\cdot(g_0\cdot x^n)) = \frac{1}{|G|}\sum_{g\in G} d((g\circ g_0)\cdot x^n).$$

Because G is a group, for every $s\in G$ there exists a unique $g\in G$ such that $s = g\circ g_0$, namely $g = s\circ g_0^{-1}$. Hence, as g ranges over G, so does s, and

$$\overline{d}(g_0\cdot x^n) = \frac{1}{|G|}\sum_{g\in G} d((g\circ g_0)\cdot x^n) = \frac{1}{|G|}\sum_{s\in G} d(s\cdot x^n) = \overline{d}(x^n).$$

Thus, we conclude that $\overline{d}$ is invariant under sample ordering permutations and depends only on the success count J_t within each arm, as desired. □

Proof of Theorem 3.12. Note that Lemmas 3.6 and 3.10 equate the minimax regret of the original problem and the minimax regret of the Bernoulli canonical problem. Moreover, by Lemma 3.11, we can restrict decision maker to treatment rules of the matrix form (3.21), which depend on the sample only through the numbers of successes in the two treatment arms. Given these results, to prove Theorem 3.12, we proceed as follows. First, we characterize the resulting minimax optimization problem and establish the existence and attainability of the matrix minimax treatment rule. Second, we demonstrate how to transfer the resulting matrix treatment rule to the original problem.

Part 1: Existence and Characterization

Fix $p = (p_0, p_1) \in [0,1]^2$. Recall that in the Bernoulli-restricted experiment as defined in Lemma 3.10, the canonical reports have the form

$$X_{t,i} = (1-h_t)B_{t,i}, \qquad B_{t,i}\sim \text{Bernoulli}(p_t).$$

Let $J_t = \sum_{i=1}^{N_t} B_{t,i}$ denote the numbers of successes in arm t. Then

$$J_t \sim \text{Binomial}(N_t, p_t),$$

and

$$\mathbb{P}(J_0 = j, J_1 = k) = \beta_{N_0,j}(p_0)\beta_{N_1,k}(p_1).$$

by independence across arms. By (3.21), when $(J_0, J_1) = (j,k)$ the matrix rule $\mathbf{d}$ assigns treatment with probability d_{jk}. Hence, its unconditional assignment probability is

$$\pi_{\mathbf{d}}(p) = D_{\mathbf{d}}(p_0, p_1),$$

where $D_{\mathbf{d}}(p_0,p_1)$ is defined by (3.22).

Next, recall the canonical outcome representation (3.13)

$$Y_t = (1-h_t)B_t + h_t A_t, \qquad A_t \in [0,1].$$

Hence, $\mu_t = (1-h_t)p_t + h_t a_t$ and any treatment effect $\Delta \in \big[(1-h_1)p_1 - (1-h_0)p_0 - h_0,\ (1-h_1)p_1 - (1-h_0)p_0 + h_1\big]$ is feasible since $a_0, a_1 \in [0,1]$.

Fix p and define regret as a function of treatment effect Δ as

$$r_{D_{\mathbf{d}}(p)}(\Delta) = [\Delta]_+ (1 - D_{\mathbf{d}}(p)) + [-\Delta]_+ D_{\mathbf{d}}(p).$$

Given that the rule does not depend on latent a_t, nature maximizes the expression at the lower or upper bound of treatment effect as a function of p, which could be expressed as (3.23). Hence,

$$\sup_{\Delta \in [\underline{\Delta}(p), \overline{\Delta}(p)]} r_{D_{\mathbf{d}}(p)}(\Delta) = \max\{b_+(p)[1 - D_{\mathbf{d}}(p)], b_-(p) D_{\mathbf{d}}(p)\}.$$

However, nature could also choose any $p \in [0,1]^2$, which yields the overall worst-case regret of an arbitrary matrix rule

$$F(\mathbf{d}) = \max_{p \in [0,1]^2} \max\{b_+(p)[1 - D_{\mathbf{d}}(p)], b_-(p) D_{\mathbf{d}}(p)\}.$$

Taking the infimum over all matrices $\mathbf{d} \in [0,1]^{(N_0+1)\times(N_1+1)}$ yields (3.24) upon proving that it is actually attained.

To do so, write $m = (N_0+1)(N_1+1)$ and note that the matrix $\mathbf{d}$ entries could be represented as a vector ranging over the compact cube $K = [0,1]^m$. Hence, by Weierstrass' theorem it is sufficient to show that $F(\mathbf{d})$ is continuous on K.

To do so, first note that each $D_{\mathbf{d}}(p)$ as defined in (3.22) is a weighted average of the entries of $\mathbf{d}$ with weights that sum to one. Hence, two matrix treatment rules $\mathbf{d}, \mathbf{e} \in K$ satisfy

$$|D_{\mathbf{d}}(p) - D_{\mathbf{e}}(p)| \le \sum_{j=0}^{N_0} \sum_{k=0}^{N_1} |d_{jk} - e_{jk}| \beta_{N_0,j}(p_0) \beta_{N_1,k}(p_1) \le \|\mathbf{d} - \mathbf{e}\|_\infty, \tag{A.15}$$

where the first inequality is the triangle inequality and the second uses $\|\mathbf{d} - \mathbf{e}\|_\infty = \max_{j,k} |d_{jk} - e_{jk}|$.

Next, take any two treatment rules $\mathbf{d}, \mathbf{e} \in K$ and fix $p \in [0,1]^2$. Note that

$$\big|b_+(p)[1 - D_{\mathbf{d}}(p)] - b_+(p)[1 - D_{\mathbf{e}}(p)]\big| = b_+(p)|D_{\mathbf{d}}(p) - D_{\mathbf{e}}(p)| \le |D_{\mathbf{d}}(p) - D_{\mathbf{e}}(p)| \le \|\mathbf{d} - \mathbf{e}\|_\infty. \tag{A.16}$$

Similarly,

$$\big|b_-(p) D_{\mathbf{d}}(p) - b_-(p) D_{\mathbf{e}}(p)\big| = b_-(p)|D_{\mathbf{d}}(p) - D_{\mathbf{e}}(p)| \le |D_{\mathbf{d}}(p) - D_{\mathbf{e}}(p)| \le \|\mathbf{d} - \mathbf{e}\|_\infty. \tag{A.17}$$

Define the regret prior to maximization over p as

$$f(\mathbf{d}, p) = \max\{b_+(p)[1 - D_{\mathbf{d}}(p)], b_-(p) D_{\mathbf{d}}(p)\}.$$

By using the elementary inequality

$$|\max\{u_1, u_2\} - \max\{v_1, v_2\}| \le \max\{|u_1 - v_1|, |u_2 - v_2|\},$$

together with (A.16) and (A.17), we obtain

$$|f(\mathbf{d}, p) - f(\mathbf{e}, p)| \le \max\Big\{b_+(p)|D_{\mathbf{d}}(p) - D_{\mathbf{e}}(p)|, b_-(p)|D_{\mathbf{d}}(p) - D_{\mathbf{e}}(p)|\Big\} \le \|\mathbf{d} - \mathbf{e}\|_\infty.$$

Since this holds uniformly over all $p \in [0,1]^2$ and maximum over p is attained by continuity of $f(\mathbf{d}, p)$, it follows that

$$|F(\mathbf{d}) - F(\mathbf{e})| = \left| \max_{p \in [0,1]^2} f(\mathbf{d}, p) - \max_{p \in [0,1]^2} f(\mathbf{e}, p) \right| \le \max_{p \in [0,1]^2} |f(\mathbf{d}, p) - f(\mathbf{e}, p)| \le \|\mathbf{d} - \mathbf{e}\|_\infty.$$

Thus, F is Lipschitz continuous on K, which implies continuity. Therefore, by Weierstrass' theorem, the continuous function F attains its minimum on compact K. Consequently, the infimum is attained, as claimed in (3.24).

Part 2: Implementation in the original problem.

Let $\mathbf{d}^*$ be the minimax treatment rule in the Bernoulli canonical problem, and take an arbitrary original sample of interval reports W. By Corollary A.2, the coefficients $q_{t,j}(W)$ defined by (3.25) are the conditional probabilities of the auxiliary Bernoulli success counts, and (3.26) is a valid treatment rule for the original observations.

Next, fix any original state $s \in \mathcal{S}_{h_0,h_1}$. By Corollary A.1, the canonical enlargement preserves the true treatment effect and maps the original sample to a canonical sample with the corresponding distribution. Corollary A.2 then yields a Bernoulli state $s^B \in \mathcal{S}^{ber}_{h_0,h_1}$ with the same treatment effect and treatment assignment probability under $\mathbf{d}^*$ as under $\delta_{\mathbf{d}^*}$ and s. Hence,

$$R(\delta_{\mathbf{d}^*}, s) = R(\mathbf{d}^*, s^B) \leq F(\mathbf{d}^*) = V_{\text{Ber}} = V_{N_0,N_1}(h_0, h_1).$$

Since s was arbitrary and the minimax value is a lower bound on the worst-case regret of every feasible original rule,

$$\sup_{s \in \mathcal{S}_{h_0,h_1}} R(\delta_{\mathbf{d}^*}, s) = V_{N_0,N_1}(h_0, h_1).$$

The matrix optimum thus yields an implementable minimax rule for the original interval data, completing the argument. □