# Minimax Regret Treatment Choice with Limited Validity of Experiments or with Covariates

Jörg Stoye*

New York University

**- Preliminary -**

November 22, 2009

**Abstract**

This paper continues the investigation of minimax regret treatment choice initiated by Manski (2004). Consider a decision maker who must assign treatment to future subjects after observing outcomes experienced in a finite sample. A modified empirical success treatment rule is known to be minimax regret optimal in numerous variants of this decision problem. I investigate the sensitivity of these findings to perturbations of the decision environment in realistic directions. They are: (i) Treatment outcomes may be influenced by a covariate whose effect on outcome distributions is bounded in one of numerous probability metrics. This is interesting because introduction of a covariate with unrestricted effects leads to a pathological result. (ii) The sample population is a potentially selective (again in one of several metrics) subset of the sample population, thus samples generate potentially misleading signals even in the limit. This formalizes concerns about external validity via a "bounds" approach that turns the problem into one of partial identification. In both cases, small but positive perturbations leave the minimax regret decision rule unchanged (with caveats in one case). Thus, minimax regret analysis is not knife-edge dependent on ignoring certain aspects of realistic decision problems; to the contrary, it recommends to entirely disregard covariates whose effect is believed to be positive but small, as well as small enough amounts of missing data. All findings are finite sample results derived by game theoretic analysis.

*Paper prepared for the CeMMAP conference on Identification and Decisions in honor of Charles Manski's $60^{th}$ birthday. I thank audiences at that conference and at UCSD and Yale for comments. I also thank audiences at Harvard/MIT and Pittsburgh as well as conference audiences at the Cowles Foundation and at ISIPTA 07, Prague, for comments on older work superseded by this paper. Part of the paper was written while visiting the Cowles Foundation, whose hospitality is gratefully acknowledged. Address: Jörg Stoye, Department of Economics, New York University, 19 W. $4^{th}$ Street, New York, NY 10012, j.stoye@nyu.edu, http://homepages.nyu.edu/~js3909.

**Keywords:** Finite sample theory, statistical decision theory, minimax regret, treatment response, treatment choice.

# 1 Introduction

One recent focus of Charles Manski's research is to reapply statistical decision theory, and in particular the minimax regret criterion, to problems of treatment choice (Manski (2000, 2004, 2006, 2007a, 2007b, 2009); Brock and Manski (2008)). This has spawned a small but active literature[1] with numerous results, not all of which were positive. In particular, Stoye's (2009a) finding on covariates – to be elaborated below – appear discouraging.

In this paper, I provide two sets of new results that may be considered good news for minimax regret. They seem at first disparate but are driven by the same feature of minimax regret: In comparing the performance of decision rules across possible parameter values, the minimax regret criterion combines a concern for assigning the ex post best treatment, which is only indirectly important under maximin utility, with a concern for the stakes involved in doing so, which is ignored by minimax rules based on "tick" loss functions (e.g., hypothesis tests). In plain English, "getting it right" is intrinsically important, but concern for it depends on the stakes at play.

I reconsider a stylized model of sample-based treatment assignment previously investigated (with slight variations) by Canner (1970) and in a number of recent papers (Manski (2004), Manski and Tetenov (2007), Schlag (2006), Stoye (2009a)). Finite sample minimax regret treatment rules for this problem are known and, by and large, intuitively reasonable. It is also known, however, that an apparently modest modification of the decision problem, namely the introduction of an observable covariate, induces a pathological result (Stoye (2009a)). This finding seems to raise at least two questions: Can we redeem minimax regret when there are covariates? More generally, are the minimax regret recommendations previously discovered overly sensitive with respect to certain simplifications made in the stylized problem?

I investigate this question by analyzing two different modifications of the problem. First, I allow for the presence of a covariate, but bound its effect on potential outcome distributions in some distance metric (a menu of such metrics is offered). In a different extension, I take a cue from Manski (2007) and model limitations to external validity by allowing sampling populations to differ from treatment populations. Specifically, I provide two different partially identifying models of selection into the sampling population. All of these modifications have the following features in common: The perturbations of the decision problem can be scaled from zero (i.e. no perturbation) to very large, and a sufficiently large perturbation will induce a pathology in the form of a no-data rule. Yet both cases also lead to very strong "local stability" results: For any sample size $N$, there exists a positive (albeit of order $O(N^{-1/2})$) size of the perturbation such that the finite sample minimax regret treatment rule is

[1] Other than references cited below, see Brock (2006) and Stoye (2007a, 2007b). Minimax regret was also independently reconsidered by Bergemann and Schlag (2007, 2008), Eozenou et al. (2006), and Chamberlain (2000).

completely unchanged (in three out of four cases) or unchanged except for a numerical feature (in one case). Thus, while there are obviously many directions in which the treatment choice problem could be generalized, minimax regret is locally insensitive to generalization in some salient such directions. The result is especially welcome in the case of covariates, where existing findings seem to spell serious trouble. Indeed, covariates with small effect are a situation that minimax regret seems to handle better than tick loss functions.

All findings are finite sample results derived by analysis of a fictitious game between the decision maker and a malicious nature. Alternative approaches would be to use large-deviations inequalities to find finite sample bounds on regret (Manski (2004)) or to analyze local asymptotic experiments (Hirano and Porter (2009)).

The remainder of this paper is structured as follows. Section 2 sets the stage by setting up the decision problem and stating a quite general result that amalgamates previous work. It is the sensitivity of this result that will be examined. Section 3 contains such analysis with respect to the introduction of covariates, whereas section 4 is devoted to limited external validity. A brief numerical illustration is provided in section 5, and section 6 concludes. While intuitions for most results are given in the text, all technical arguments are collected in an appendix.

# 2 Setting the Stage

## 2.1 The Decision Problem

The decision problem is as in Manski (2004), and I use the same notation as in Stoye (2009a). A decision maker must assign one of two treatments $T \in \{0,1\}$ to members $j$ of a treatment population $J$. Each member of the treatment population has a response function $y^j(t) : \{0,1\} \to \{0,1\}$ that maps treatments onto outcomes. Substantively, what this really assumes is that a priori bounds on treatment outcomes exist, are known, and coincide across treatments. Restricting them to lie in $[0,1]$ is then a normalization. The additional restriction to binary outcomes is with less loss of generality than might appear because all treatment rules analyzed in this paper can be extended to general outcomes by a "binary randomization" technique due to Schlag (2006).[2] Specifically, one can define treatment rules for non-binary outcomes by first replacing every observed outcome with one realization of a binary, mean-preserving spread of itself and then operating rules defined for binary outcomes. This leaves intact minimax regret values and can, therefore, be used to find minimax regret efficiency bounds as well as decision rules that attain them.

[2] More precisely, the technique was first applied to statistical treatment choice by Schlag (2006); it had been independently discovered for related problems by Cucconi (1968), Gupta and Hande (1992), and Schlag (2003). See Stoye (2009a) for an elaboration in this paper's notation.

The population is a probability space $(J, \Sigma, P)$ and is "large" in the sense that $J$ is uncountable and $P(j) = 0$ for all $j$. The decision maker cannot distinguish between members of $J$, hence from her point of view, assigning treatment $t$ induces a random variable $Y_t$ (the *potential outcome*) with distribution $P(y^j(t))$. We will focus on the distribution $P(Y_0, Y_1)$ as unknown quantity. Specifically, a *state of the world* $s$ will be identified with $P(Y_0, Y_1)$, which is partially characterized by a couplet $(\mu_0, \mu_1) \equiv \mathbb{E}(Y_0, Y_1)$; this characterization will turn out to be sufficient. The set $\mathcal{S}$ collects feasible states of the world. I analyze both a situation of complete ignorance and the problem of testing an innovation, in which the behavior of treatment 0 is well understood. Formally, complete ignorance means that $\mathcal{S} = \Delta\{0,1\}^2$, the set of distributions over $\{0,1\}^2$; testing an innovation means that $\mathcal{S} = \left\{P(Y_0, Y_1) \in \Delta\{0,1\}^2 : \mathbb{E}Y_0 = \mu_0\right\}$, where $\mu_0$ is known. Notation will be extended to accommodate covariates as needed.

Assume that if $s$ were known, the decision maker would resolve the decision problem by maximizing expected outcomes, thus she would assign all subjects to $T = 1$ if $\mu_1 > \mu_0$, to $T = 0$ if $\mu_1 < \mu_0$, and she would be indifferent if $\mu_0 = \mu_1$. While this can capture risk aversion ($Y_t$ might be a utility), it does presumes a utilitarian social welfare function. The decision maker observes not $s$ but a signal of $s$, namely treatment outcomes experienced by a random sample of $N$ members of the treatment population. I will take $N$ to be known, although the generalization to $N$ being a random variable with known distribution is conceptually simple. The experiment generates a sample space $\Omega \equiv (\{0,1\} \times [0,1])^N$ with typical element $\omega = (t_n, y_n)_{n=1}^N$. The sampling distribution of $T$ depends on the *sample design*. The following designs will be considered:

- **Independent Randomization:** Both treatments are unknown, $N$ is odd or even, and $\{t_1, t_2, \ldots\}$ are i.i.d. realizations of Bernoulli variables with parameter 1/2 (i.e., independent tosses of fair coins).
- **Free Treatment Assignment:** Both treatments are unknown, $N$ is odd or even, and within-sample treatment assignment is a choice variable.
- **Testing an Innovation:** $\mu_0$ is known. Obviously, all sample subjects will be assigned to treatment 1.

Conditional on a realization $t_n$, $y_n$ is an independent realization of $Y_{t_n}$. The decision maker can specify a *statistical treatment rule* $\delta : \Omega \mapsto [0,1]$ that maps possible sample realizations $\omega$ onto treatment assignments $\delta(\omega) \in [0,1]$, where the value of $\delta$ is interpreted as probability of assigning treatment 1. Nonrandomized decision rules take values only in $\{0,1\}$, but randomization is allowed and will be used. The set of all such decision rules will be denoted by $\mathcal{D}$; in the case of free treatment assignment, the decision maker's action space is really $\Delta\{0,1\}^N \times \mathcal{D}$.[3]

The expected outcome generated by $\delta$ given $s$ is

$$u(\delta, s) \equiv \mu_0 \left(1 - \mathbb{E}\delta(\omega)\right) + \mu_1 \mathbb{E}\delta(\omega),$$

where expectations are taken with respect to the sampling distribution of $\omega$. Seen as a function of $s$, $u(\delta, s)$ is (the negative of) the *risk function* of treatment rule $\delta$. Absent a prior on $s$, attempts to optimize $u(\delta, s)$ induce a decision problem under ambiguity (Manski (2000)). The two most prominent resolutions of this problem are the Bayesian approach, i.e. to rank decision rules according to $\int u(\delta, s)d\pi$, where $\pi$ is a prior on $\mathcal{S}$ (see Chamberlain (2009) for an elucidation and Dehejia (2005) for an application to treatment choice) and the maximin utility approach, i.e. to rank decision rules according to $\min_{s\in\mathcal{S}} u(\delta, s)$. However, Manski (2004) initiated reconsideration of *minimax regret* in this context. To understand this criterion, first define the regret incurred by decision rule $\delta$ in state $s$,

$$R(\delta, s) \equiv \max_{d\in\mathcal{D}} u(d, s) - u(\delta, s),$$

the difference between the expected outcome induced by $\delta$ and the outcome that could have been achieved, had $s$ been known. Minimax regret is a maximin ranking with respect to regret loss, thus it recommends to pick

$$\delta^* \in \arg\min_{\delta\in\mathcal{D}} \max_{s\in\mathcal{S}} R(\delta, s)$$

if such a $\delta^*$ exists, as is the case in examples below. Minimax regret was originally introduced by Savages's (1951) reading of Wald (1950); see Stoye (2009c) for further references on history, axiomatizations, and applications. In order to avoid redundancy, I will jump directly to the core result that is this paper's starting point.

## 2.2 Existing Results

Exact (finite sample) solutions to the above treatment choice problems are available for both maximin utility and minimax regret. Indeed, one motivation for investigating minimax regret is a pathology of maximin utility. If both treatments are unknown, every decision rule achieves maximin utility because $\min_{s\in\mathcal{S}} u(\delta, s) = 0$ for all $\delta$, generated if $\mu_0 = \mu_1 = 0$. For testing an innovation, the no-data rule $\delta \equiv 0$ achieves maximin utility, generated by $\mu_1 = 0$. Both cases are instances of a degeneracy problem that was diagnosed already by Savage (1954) and, for the present problem or close variations of it, by Manski (2004), Schlag (2006), and Stoye (2009a).[4] In contrast, the minimax regret solutions to the decision problem appear quite natural. They will be presented now.

[3] This excludes sequential sample designs, i.e. $t_n$ cannot depend on lagged realizations of $y_n$. The according extension is left for future research.

[4] In the abstract, the same types of examples can be constructed for minimax regret (Parmigiani (1992)). They do not seem to occur naturally as one tries to model treatment choice problems, though.

Define

$$\delta_1^*(\omega) \equiv \begin{cases} 0, & I_N < 0 \\ 1/2, & I_N = 0 \\ 1, & I_N > 0 \end{cases},$$

where

$$\begin{aligned} I_N &\equiv N_1(\overline{y}_1 - 1/2) - N_0(\overline{y}_0 - 1/2) \\ &\propto [\#\,(\text{observed successes of treatment 1}) + \#\,(\text{observed failures of treatment 0})] \\ &\quad - [\#\,(\text{observed successes of treatment 0}) + \#\,(\text{observed failures of treatment 1})] \end{aligned}$$

with $N_t$ the number of sample subjects assigned to treatment $t$, $\overline{y}_t$ a sample mean that conditions on $T = t$, and the convention that $N_t(\overline{y}_t - 1/2) = 0$ if $N_t = 0$.

For the case of testing an innovation, define

$$\delta_2^*(\omega) \equiv \begin{cases} 0, & N\overline{y}_1 < n^* \\ \lambda^*, & N\overline{y}_1 = n^* \\ 1, & N\overline{y}_1 > n^* \end{cases},$$

where $n^* \in \{1, \ldots, N\}$ and $\lambda^* \in [0, 1]$ are characterized as follows:

$$\begin{aligned} &\max_{a \in [0,\mu_0]} (\mu_0 - a) \left[ \sum_{n > n^*} \binom{N}{n} a^n (1-a)^{N-n} + \lambda^* \binom{N}{n^*} a^{n^*} (1-a)^{N-n^*} \right] \qquad (1) \\ = &\max_{a \in [\mu_0,1]} (a - \mu_0) \left[ \sum_{n < n^*} \binom{N}{n} a^n (1-a)^{N-n} + (1-\lambda^*) \binom{N}{n^*} a^{n^*} (1-a)^{N-n^*} \right]. \end{aligned}$$

This rule is a threshold crossing rule that assigns treatment 1 if its success count exceeds a certain critical value, again with randomization on the threshold. Numerical evaluation reveals that the rule is quite close to being an empirical success rule, with a very slight bias toward believing that $\mu_1 = 1/2$. This renders it similar to rules derived from noninformative priors but quite different from rules based on hypothesis tests, which will be much more conservative.

The following result is found in Stoye (2009a).

**Proposition 1** ***(i)*** *Let both treatments be unknown and assume that the sample design is independent randomization. Then* $\delta_1^*$ *achieves minimax regret.*

***(ii)*** *Let both treatments be unknown and let sample design be a choice variable. Then minimax regret is achieved by independent randomization in conjunction with* $\delta_1^*$.

***(iii)*** *Let treatment* 0 *be known. Then minimax regret is achieved by* $\delta_2^*$.

**Corollary 1** *The decision problem from proposition 1(i)-(ii) has value*

$$\begin{aligned} R_1^*(N) &= \max_{a\in[1/2,1]} \left\{ (2a-1) \sum_{n<N/2} \binom{N'}{n} a^n (1-a)^{N'-n} \right\} \\ N' &= \max_{M\in\mathbb{N}} \{M \leq N : M \textit{ is odd}\}, \end{aligned}$$

*where* $R_1^*(0) = 1/2$.

*The decision problem from proposition 1(iii) has value*

$$R_2^*(N) = \max_{a\in[\mu_0,1]} \left\{ (a-\mu_0) \left[ \sum_{n<n^*} \binom{N}{n} a^n (1-a)^{N-n} + (1-\lambda^*) \binom{N}{n^*} a^{n^*} (1-a)^{N-n^*} \right] \right\}$$

*with* $(n^*, \lambda^*)$ *as in proposition 1(iii), and with* $R_2^*(0) = \mu_0(1-\mu_0)$.

An important technique to establish this kind of finite sample result is game theoretic analysis (Wald (1945)). Consider a fictitious zero-sum game in which the decision maker picks a decision rule $\delta \in \mathcal{D}$ (possibly at random; note, though, that the result of any randomization over decision rules can itself be expressed as element of $\mathcal{D}$) and Nature picks a state of the world $s$ (possibly at random, meaning that her strategy space is the set $\Delta\mathcal{S}$ of priors $\pi$ over $\mathcal{S}$). After both players moved, $s$ is drawn according to $\pi$, $\omega$ is drawn from $s$ according to the relevant sampling scheme, and $\delta$ is operated on $\omega$. The decision maker then pays $\max_{d\in\mathcal{D}} u(d,s) - u(\delta, s)$ to Nature. Any Nash equilibrium $(\delta^*, \pi^*)$ of this game characterizes a minimax regret treatment rule $\delta^*$, and Nature's strategy $\pi^*$ in such an equilibrium can be interpreted as the least favorable prior that is implicitly selected by the minimax regret criterion.

In the case of proposition 1, these least favorable priors are as follows. It turns out that $R(\delta, s)$ depends on $s$ only through $(\mu_0, \mu_1)$, thus identify states with couplets $(\mu_0, \mu_1)$. Then the least favorable prior for parts (i) and (ii) randomizes evenly over two symmetric states $(a, 1-a)$ and $(1-a, a)$. The least favorable prior in part (iii) randomizes (not in general evenly) over two states $(\mu_0, b)$ and $(\mu_0, c)$ with $b > \mu_0 > c$. Computation of $(a, b, c)$ is not required for the proofs but leads to the value functions provided in corollary 1.

# 3 Treatment Choice with Limited Internal or External Validity

Manski recently proposed to extend analysis of treatment choice by incorporating concerns that are not usually emphasized in econometric treatment evaluation, e.g. by incorporating deontological considerations into the loss function (Manski (2009)). I here pick up a different cue: In his recent monograph, Manski (2007) mentions external validity, i.e. the question whether conclusions that are valid within an experimental population can be extrapolated beyond the lab. In Manski's words, "an experiment is said to have external validity if the distribution of outcomes realized by a treatment group is the same as the distribution of outcomes that would be realized in an actual program" (p. 26). However, "participation in experiments ordinarily cannot be mandated in democracies. Hence experiments in practice usually draw subjects at random from a pool of persons who volunteer to participate. So one learns about treatment response within the population of volunteers rather than within the population of interest" (p. 138).

This section takes up the challenge of explicitly modelling such a concern. I will allow that the sampling population (that is, the universe from which samples are drawn) is a selective subset of the treatment population (that is, the population whose treatment responses are of ultimate interest). Obviously, interesting results are available only if the sampling population is at least somewhat informative about the treatment population,[5] and the relation between the two will accordingly be constrained. However, I avoid formulating an explicit model of selection that would restore identification. The approach will rather be to impose some assumptions that nonparametrically constrain the relation between the populations without pinning it down up to identified parameters. As a result, true treatment effects will be only partially identified, that is, only bounds on them are identifiable and therefore learnable even in the limit. The present analysis thereby connects finite sample analysis of treatment choice with another literature pioneered by Manski, namely partial identification.[6]

Specifically, consider the following two models.

**Example 1** ***Selection into Treatment as a Missing Data Problem***

*The sampling population is a subset of the treatment population. No assumption is made about how this subset is selected; however the subset has mass* $1-\varepsilon$*, where* $\varepsilon$ *is known. Thus, if* $Z \in \{0,1\}$ *indicates membership in the sampling population, then the distribution* $P(Y_0, Y_1)$ *of potential outcomes in the treatment population is characterized as*

$$P(Y_0, Y_1) = (1-\varepsilon)P(Y_0, Y_1|Z=1) + \varepsilon P(Y_0, Y_1|Z=0).$$

*This selection model can be thought of as a simple missing data problem. In particular, if both treatments are unknown,* $\Delta = \mathbb{E}(Y_1 - Y_0|Z=1)$ *is the average treatment effect in the sampling population, and* $\Delta^* = \mathbb{E}(Y_1 - Y_0)$ *is the average treatment effect of interest, then tight bounds on the latter are (Manski (1989)):*

$$(1-\varepsilon)\Delta - \varepsilon \leq \Delta^* \leq (1-\varepsilon)\Delta + \varepsilon.$$

[5] Perhaps this is not as obvious as one might have thought. Woorall (2007) relates the case of benoxaprofen, where a dramatic failure of external validity occurred because the medication was mainly used on elderly patients, but these were not at all sampled into the randomized clinical trials.

[6] See Manski (2003) for a survey. Minimax regret treatment choice has also been linked to partial identification analysis in Brock (2006), Stoye (2007a, 2007b, 2009b), Manski (2007a, 2007b, 2009), and Tetenov (2009).

*If $\mu_0$ is known, then*

$$(1-\varepsilon)\mathbb{E}(Y_1|Z=1)-\mu_0 \le \Delta^* \le (1-\varepsilon)\mathbb{E}(Y_1|Z=1)+\varepsilon-\mu_0.$$

*From a robust Bayesian perspective, this model can also be thought of as a contamination neighborhood model.*

**Example 2** ***Selection into Treatment as a Hidden Covariate Problem***

*Do not constrain the amount of missing data, but let the decision to join the sampling population (e.g., to volunteer) be modelled via a logit model with unobservable covariate $X \in [0,1]$. The effect of $X$ on treatment response is unconstrained, but the coefficient $\beta$ on $X$ in the selection model lies in $[0,b]$, where $b$ is known.*

*Under this selection model, the average treatment effect $\Delta^*$ is again not identified, but can be tightly bounded. If both treatments are unknown, one can verify that tight bounds in terms of $(\mathbb{E}(Y_0|Z=1), \mathbb{E}(Y_1|Z=1))$ are:*

$$\frac{\mathbb{E}(Y_1|Z=1)}{e^b+(1-e^b)\mathbb{E}(Y_1|Z=1)}-\frac{e^b\mathbb{E}(Y_0|Z=1)}{1+(e^b-1)\mathbb{E}(Y_0|Z=1)} \le \Delta^* \le \frac{e^b\mathbb{E}(Y_1|Z=1)}{1+(e^b-1)\mathbb{E}(Y_1|Z=1)}-\frac{\mathbb{E}(Y_0|Z=1)}{e^b+(1-e^b)\mathbb{E}(Y_0|Z=1)}.$$

*If $\mu_0$ is known, the bounds are*

$$\frac{\mathbb{E}(Y_1|Z=1)}{e^b+(1-e^b)\mathbb{E}(Y_1|Z=1)}-\mu_0 \le \Delta^* \le \frac{e^b\mathbb{E}(Y_1|Z=1)}{1+(e^b-1)\mathbb{E}(Y_1|Z=1)}-\mu_0.$$

*From a robust Bayesian perspective, this model can also be thought of as an odds-ratio neighborhood model.*

The first model formalizes selection into the sampling universe as a missing data problem. If only a small subset of the treatment population select themselves into the sampling population, this will naturally be insufficiently restrictive to allow for interesting conclusions. (In particular, it is easy to see that a no-data rule obtains if $\varepsilon = 1$.) In contrast, the hidden covariate model can allow for such conclusions even if the sampling population is a small subset of the treatment population. Of course, this is because this model is more stringent along other dimensions. While the effect of $X$ on treatment response is completely unconstrained, hence $X \in [0,1]$ is a vacuous assumption as far as distortion of treatment responses goes, the bound on $\beta$ combines with the upper limit on $X$ to constrain the effect of $X$ on selection probabilities. Specifically, variation in $X$ can change conditional odds ratios of entering the treatment population by a factor of at most $\exp(b)$.

I will take $\varepsilon$ respectively $b$ to be known to the researcher. If they can only be bounded, then it is w.l.o.g. to set them equal to their respective upper bound, because in both cases, identification decays monotonically as the parameter increases. It is conceivable that while not strictly known, $\varepsilon$ and/or $b$ are identifiable from data other than the ones under consideration. In other contexts, they are probably best thought of as a user-specified, “plausibility” or “sensitivity” parameter; such sensitivity parameters are by now quite common in both biostatistics and econometrics.

### 3.1 Two Unknown Treatments

Analysis of limited external validity will have to distinguish the case of two unknown treatments versus the case of testing an innovation. This section’s main result consists of a full characterization of finite sample minimax regret treatment rules for the former case. To do this for the missing data scenario, it is necessary to introduce a new decision rule. Let

$$\delta_3^*(\omega) = \begin{cases} 0, & \frac{1}{2}\left(1 + \frac{1-\varepsilon}{\varepsilon}\frac{I_N}{N}\right) < 0 \\ \frac{1}{2}\left(1 + \frac{1-\varepsilon}{\varepsilon}\frac{I_N}{N}\right), & 0 \leq \frac{1}{2}\left(1 + \frac{1-\varepsilon}{\varepsilon}\frac{I_N}{N}\right) \leq 1 \\ 1, & \frac{1}{2}\left(1 + \frac{1-\varepsilon}{\varepsilon}\frac{I_N}{N}\right) > 1 \end{cases}.$$

While not a threshold rule, $\delta_3^*$ is sensitive to the data in an intuitively reasonable manner because the probability of assigning treatment 1 increases in the success score $I_N$.[7] An interesting observation is that if $\varepsilon \geq 1/2$, i.e. the missing population has more probability mass than the observable one, then the truncation of $\frac{1}{2}\left(1 + \frac{1-\varepsilon}{\varepsilon}\frac{I_N}{N}\right)$ does not bind, and the decision rule is linear in $I_N$. In this case, $\delta_3^*$ is finite sample optimal, a finding that will be established below en passant but that would also follow by minimally extending a previous result of Manski (2007a). This section’s first main result is that for any $\varepsilon$ and $N$, it is finite sample minimax regret optimal to mix over just these two decision rules, where the mixture is degenerate both as $\varepsilon \geq 1/2$ and as $\varepsilon \to 0$.

**Proposition 2** *Consider the missing data setting of example 1. Define the decision rule*

$$\delta_{MD}^* = \alpha^*\delta_1^* + (1-\alpha^*)\delta_3^*,$$

*where*

$$\alpha^* = \begin{cases} 0, & \varepsilon \geq \frac{1}{2} \\ \dfrac{2^{N'-1}\frac{1-\varepsilon}{\varepsilon} - \sum_{z \geq \frac{N'(1-2\varepsilon)}{2-2\varepsilon}} \binom{N'}{z} \min\left\{\frac{1}{2} + \frac{1-\varepsilon}{\varepsilon}\left(\frac{z}{N'} - \frac{1}{2}\right), 1\right\}(2z-N')}{\sum_{z>N'/2}\binom{N'}{z}(2z-N') - \sum_{z \geq \frac{N'(1-2\varepsilon)}{2-2\varepsilon}} \binom{N'}{z} \min\left\{\frac{1}{2} + \frac{1-\varepsilon}{\varepsilon}\left(\frac{z}{N'} - \frac{1}{2}\right), 1\right\}(2z-N')}, & \frac{1}{2} < \varepsilon < \varepsilon_N^* \\ 1, & \varepsilon \leq \varepsilon_N^* \end{cases}$$

*with* $N' = \max\{M \leq N : M \text{ is odd}\}$ *as before and* $\varepsilon_N^* \equiv \left(1 + 2\sum_{z>N'/2}\binom{N'}{n}(2n-N')\right)^{-1}$.

*Then proposition 1(i)-(ii) applies with* $\delta_1^*$ *replaced by* $\delta_{MD}^*$. *In particular, proposition 1(i)-(ii) applies unchanged if* $\varepsilon \leq \varepsilon_N^*$.

[7] The rule can be made well-defined – and optimal – for $\varepsilon = 0$ by using the conventions $(1-\varepsilon)/\varepsilon = \infty$ and $0 \cdot (1-\varepsilon)/\varepsilon = 0$ for this case, but this detail is not substantively needed because $\delta_1^*$ is then optimal anyway.

The conceptually most interesting aspect of this result is probably the implied robustness result that resembles proposition 2: Small enough but strictly positive limitations to external validity (namely, $\varepsilon \leq \varepsilon_N^*$) will be completely ignored. The probability with which the simple threshold rule $\delta_1^*$ is played then decreases continuously as $\varepsilon$ rises above $\varepsilon_N^*$, and Manski's (2007a) result applies precisely as $\varepsilon$ reaches 1/2.

The result can be interpreted in terms of a switch between two equilibrium types in the underlying game. Nature has an incentive to make signals generated by the sampling population as uninformative as possible; in the extreme case, they might be misleading in the sense that the treatment which is better in the sampling population, and therefore likely to be more successful in the sample, is in fact the overall worse treatment. It turns out that this extreme scenario cannot occur in equilibrium. However, there are two different regions in parameter space: For $\varepsilon \leq \varepsilon_N^*$, the least favorable prior has the feature that with probability one, the optimal treatment is the one that is also optimal for the sample population, meaning that the decision maker's best response to the prior is completely unchanged. At $\varepsilon = \varepsilon_N^*$, the equilibrium switches to a "pooling equilibrium" in which both treatments are equally good for the sample population, thus the data are pure noise.

An interesting feature of pooling equilibria is that if the data are uninformative, then the value of the game cannot depend on how many data there are. This is here reflected in the fact that the maximal regret incurred by the minimax regret strategy equals $\varepsilon/2$ whenever $\varepsilon \geq \varepsilon_N^*$. Thus, if one holds $\varepsilon < 1/2$ fixed and increases $N$, finite sample minimax regret will cease to decrease as soon as $\varepsilon_N^* \geq \varepsilon$. Beyond this point, the value of additional data to a minimax regret decision maker drops to zero; one could say that the identification problem completely dominates estimation issues.[8] This observation is related to an independent finding by Tetenov (2009) with regard to both the flavor of the result and the underlying mechanism.

I now extend the finding to the hidden covariate setting. For this purpose, define

$$\delta_4^*(\omega) = \begin{cases} 0, & \frac{1}{2} + \frac{I_N}{N}\frac{4\exp(b)}{\exp(2b)-1} < 0 \\ \frac{1}{2} + \frac{I_N}{N}\frac{4\exp(b)}{\exp(2b)-1}, & 0 \leq \frac{1}{2} + \frac{I_N}{N}\frac{4\exp(b)}{\exp(2b)-1} \leq 1 \\ 1, & \frac{1}{2} + \frac{I_N}{N}\frac{4\exp(b)}{\exp(2b)-1} > 1 \end{cases}$$

for $b > 0$. This rule plays a similar role here as $\delta_3^*$ above. It is relatively easy to show that it achieves minimax regret if $b \geq \log\left(2 + \sqrt{5}\right)$; this is again the threshold at which the truncation of $\delta_4^*(\omega)$ never binds. The work is in showing that $\delta_1^*$ is optimal for $b$ small enough and that exact minimax regret can be achieved for intermediate $b$ by mixing the rules. The algebraic detail is somewhat elaborate, but should of course be ignored by a casual reader.

[8] As an aside, this imediately implies that the minimax regret treatment rule identified here is not unique. Once a pooling equilibrium is reached, the DM has many minimax regret equivalent decision rules at her disposal because she can choose to discard data and just use $\delta_{MD}^*$ for the accordingly adjusted sample size $M$ (as long as one has $\varepsilon \geq \varepsilon_M^*$).

| N | 1 | 3 | 5 | 10 | 20 | 50 | 100 | 200 | 500 | 1000 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\boldsymbol{\varepsilon}_{\mathbf{N}}^{*}$ | 0.5000 | 0.4000 | 0.3478 | 0.2889 | 0.2210 | 0.1512 | 0.1116 | 0.0815 | 0.0531 | 0.0381 |
| $\exp(\mathbf{b}_N^*)$ | 4.23 | 3.00 | 2.53 | 2.10 | 1.71 | 1.42 | 1.28 | 1.19 | 1.12 | 1.08 |

Table 1: Numerical illustration of propositions 4 and 5.

**Proposition 3** *Consider the hidden covariate setting of example 2. Define the decision rule*

$$\delta_{HC}^* = \alpha^* \delta_1^* + (1 - \alpha^*) \delta_4^*,$$

*where*

$$\alpha^* = \begin{cases} 0, & b \geq \log\left(2+\sqrt{5}\right) \\ \frac{\frac{4\exp(b)}{\exp(2b)-1} - \sum_{z \geq N'(1-\exp(2b)/2)} \binom{N'}{z} \min\left\{\frac{1}{2} + \frac{4\exp(b)}{\exp(2b)-1}\left(\frac{z}{N'} - \frac{1}{2}\right), 1\right\}\left(2z - N'\right)}{\sum_{z > N'/2} \binom{N'}{z}(2z-N') - \sum_{z \geq N'(1-\exp(2b)/2)} \binom{N'}{z} \min\left\{\frac{1}{2} + \frac{4\exp(b)}{\exp(2b)-1}\left(\frac{z}{N'} - \frac{1}{2}\right), 1\right\}(2z-N')}, & b_N^*(N) < b < \log\left(2+\sqrt{5}\right) \\ 1, & b \leq b_N^* \end{cases}$$

*with* $N' = \max\{M \leq N : M \textit{ is odd}\}$ *as before and*

$$b_N^* = \log\left(\left(\sum_{n<N^*/2} \binom{N'}{n}(2n - N')\right)^{-1} + \left(1 + \left(\sum_{n<N^*/2} \binom{N'}{n}(2n - N')\right)^{-2}\right)^{1/2}\right).$$

*Then proposition 1(i)-(ii) applies with* $\delta_1^*$ *replaced by* $\delta_{HC}^*$*. In particular, proposition 1(i)-(ii) applies unchanged if* $b \leq b_N^*$*.*

The threshold distortions below which proposition 1 applies unchanged are displayed in table 2. As an example, if $N = 10$, minimax regret prescribes to ignore contaminations of up to 29%, i.e. scenarios in which the signal generated by the sample population could be quite misleading indeed. The corresponding number for a sample size of $N = 100$ still exceeds 10%.

## 4 Limited Internal Validity

**(This section brand new, rough around the edges, and probably full of typos)**

I now turn attention to limitations of internal validity, that is, causal inference may be inappropriate even within the sampling population because not all sample subjects comply with assigned treatment, and noncompliance may be selective. I present two models in which one can observe treatment assigned but not treatment received. The exact same results emerge if one can observe treatment received but not treatment assigned; this is because only the wedge between the population analog of the naive treatment estimator, $\Delta = \mathbb{E}(Y_1 - Y_0)$, and the true ATE will matter. I expect a qualitatively similar result to apply if both assigned and received treatment are observable, but closed-form results then appear infeasible for reasons that will be explained.

Thus, assume that sampling and treatment population coincide, but that a probability mass $\varepsilon$ of subjects are noncompliers, that is, they may not receive the assigned treatment. I will presume in the following that assigned treatment is observed but received treatment is not. For example, this scenario applies to field experiments in which one cannot verify whether assigned treatment was in fact taken. It also applies if the "treatment" variable is a covariate whose effect on outcomes is of interest but which is observed subject to classification error. Finally, the same results reported below apply if one observes received but not assigned treatment, an example being social experiments in which treatment was assigned at random, treatment received is recorded, but there is a worry that some sample subjects switched treatment groups without this being recorded in the data (e.g., in the Tennessee STAR experiment).

The scenario again generates a situation of partial identification. If nothing is assumed about the behavior of noncompliers, then worst-case bounds on the true ATE $\Delta^*$ in terms of the observable ATE $\Delta$ are

$$\Delta - 2\varepsilon \leq \Delta^* \leq \Delta + 2\varepsilon.$$

They are achieved by all noncompliers being "anti-compliers" and having a treatment effect of $-1$ respectively $+1$.[9] This behavioral assumption is frequently implausible. I therefore also consider a model in which noncompliers are either "always-takers" or "never-takers," that is they simply ignore their treatment assignment. The assumption that there are no "anti-compliers" is a crucial identification assumption in the literature on Latent Average Treatment Effects (LATE, as in Angrist, Imbnes, and Rubin (1996)), where it is called monotonicity. What's more, it is essentially equivalent to imposing a threshold crossing model for selection into treatment, with treatment assignment working as an instrument (Vytlacil (2006)). Under this assumption, bounds improve to

$$\Delta - \varepsilon \leq \Delta^* \leq \Delta + \varepsilon,$$

achieved if the treatment effect for all noncompliers is $-1$ respectively $+1$, and irrespective of the fraction of always-takers within the set of noncompliers.

Results similar to propositions 4 and 5 can be generated for these scenarios.

**Proposition 4** *Consider the noncompliance scenario with unrestricted behavior of noncompliers. Define the decision rule*

$$\delta_5^*(\omega) = \begin{cases} 0, & \frac{1}{2} + \frac{I_N}{2\varepsilon N} < 0 \\ \frac{1}{2} + \frac{I_N}{2\varepsilon N}, & 0 \leq \frac{1}{2} + \frac{I_N}{2\varepsilon N} \leq 1 \\ 1, & \frac{1}{2} + \frac{I_N}{2\varepsilon N} > 1 \end{cases}$$

[9] This scenario has observable implications: It implies that both $\mu_1$ and $\mu_0$ lie in $[\varepsilon, 1-\varepsilon]$, and therefore might be eventually learned to not apply, which does not affect minimax regret analysis however. Note that if the implication fails, then the bounds displayed do not lie within $[-1, 1]$ and should, of course, be accordingly truncated.

*and the decision rule*

$$\delta^*_{NC1} = \alpha^* \delta^*_1 + (1 - \alpha^*)\delta^*_5,$$

*where*

$$\alpha^* = \max\left\{0, \frac{\frac{1}{2\varepsilon} - \sum_{z \geq N'(1-\varepsilon)} \binom{N'}{z} \min\left\{\frac{1}{2} + \frac{1}{2\varepsilon}\left(\frac{z}{N'} - \frac{1}{2}\right), 1\right\}(2z - N')}{\sum_{z > N'/2} \binom{N'}{z}(2z - N') - \sum_{z \geq N'(1-\varepsilon)} \binom{N'}{z} \min\left\{\frac{1}{2} + \frac{1}{2\varepsilon}\left(\frac{z}{N'} - \frac{1}{2}\right), 1\right\}(2z - N')}\right\},$$

*with* $N' = \max\{M \leq N : M \text{ is odd}\}$.

*Then proposition 1(i)-(ii) applies with* $\delta^*_1$ *replaced by* $\delta^*_{NC1}$. *In particular, proposition 1(i)-(ii) applies unchanged if* $\varepsilon \leq \left(2\sum_{z > N'/2}\binom{N'}{n}(2n - N')\right)^{-1}$.

**Proposition 5** *Consider the noncompliance scenario and impose monotonicity. Define the decision rule*

$$\delta^*_6(\omega) = \begin{cases} 0, & \frac{1}{2} + \frac{I_N}{4\varepsilon N} < 0 \\ \frac{1}{2} + \frac{I_N}{4\varepsilon N}, & 0 \leq \frac{1}{2} + \frac{I_N}{4\varepsilon N} \leq 1 \\ 1, & \frac{1}{2} + \frac{I_N}{4\varepsilon N} > 1 \end{cases}$$

*and the decision rule*

$$\delta^*_{NC2} = \alpha^* \delta^*_1 + (1 - \alpha^*)\delta^*_6,$$

*where*

$$\alpha^* = \begin{cases} 0, & \varepsilon \geq 1/2 \\ \frac{\frac{1}{4\varepsilon} - \sum_{z \geq N'(1-2\varepsilon)} \binom{N'}{z} \min\left\{\frac{1}{2} + \frac{1}{4\varepsilon}\left(\frac{z}{N'} - \frac{1}{2}\right), 1\right\}(2z - N')}{\sum_{z > N'/2} \binom{N'}{z}(2z - N') - \sum_{z \geq N'(1-2\varepsilon)} \binom{N'}{z} \min\left\{\frac{1}{2} + \frac{1}{4\varepsilon}\left(\frac{z}{N'} - \frac{1}{2}\right), 1\right\}(2z - N')}, & \varepsilon^*_N < \varepsilon < 1/2 \\ 1, & \varepsilon \leq \varepsilon^*_N \end{cases},$$

*with* $N' = \max\{M \leq N : M \text{ is odd}\}$ *and* $\varepsilon^*_N = \left(4\sum_{z > N'/2}\binom{N'}{n}(2n - N')\right)^{-1}$.

*Then proposition 1(i)-(ii) applies with* $\delta^*_1$ *replaced by* $\delta^*_{NC1}$. *In particular, proposition 1(i)-(ii) applies unchanged if* $\varepsilon \leq \varepsilon^*_N$.

Technically, these two results follow very much along the same lines as propostion 4. In particular, it must be admitted that closed-form analysis is feasible because of the simple nature of the worst-case bounds on $\Delta^*$ given $\Delta$ (these bounds are, of course, achieved with probability 1 under the least favorable prior). As a contrast, consider the situation where both assigned and received treatment are observable. Bounds on $\Delta^*$ are then known, but generally are characterized as the solution to a linear programming problem involving observable features of the joint distribution of assigned treatments, received treatments, and outcomes (Balke and Pearl (2000)). The bounds are well-behaved as functions of $\Delta$, so it is expected that minimax regret treatment rules based on them will have qualitatively similar features to the above; in fact, insensitivity to a small amount of noncompliance can be shown. A full closed-form analysis does not appear feasible, however.

## 4.1 The Case of Testing an Innovation

Parallels between the symmetric testing case and the case of testing an innovation are not as tight in this section as in the preceding one. For one thing, I am not able to produce a complete equilibrium analysis and will only present a robustness finding along the lines of section 3. Even this robustness finding is, however, weaker than propositions 4 and 5. For limitations to external validity that are small enough, the minimax regret decision rule looks similar to the one from proposition 1 and indeed approaches it as $\varepsilon \to 0$ respectively $b \to 0$. But this time around, the approximation is only the weaker one that one would have hoped for as a minimum, that is, small perturbations of the environment lead to small changes of the rule. Specific results are as follows.

**Proposition 6** *Consider the missing data setting of example 1. There exists $\varepsilon_2 > 0$ s.t. if $\varepsilon \leq \varepsilon_2$, proposition 1(iii) continues to apply, with the modification that condition (1) turns into*

$$\begin{aligned}
& \max_{a\in[0,1]} \left\{((1-\varepsilon)a + \varepsilon - \mu_0)\left(\Pr(N\overline{y}_1 < n^*) + (1-\lambda^*)\Pr(N\overline{y}_1 = n^*)\right)\right\} \\
= & \max_{a\in[0,1]} \left\{(\mu_0 - (1-\varepsilon)a)\left(\Pr(N\overline{y}_1 > n^*) + \lambda^*\Pr(N\overline{y}_1 = n^*)\right)\right\},
\end{aligned}$$

*where the probabilities refer to $N\overline{y}_1$ being binomially distributed with parameters $(N, a)$.*

**Proposition 7** *Consider the hidden covariate setting of example 2. There exists $b_2 > 0$ s.t. if $b \leq b_2$, proposition 1(iii) continues to apply, with the modification that condition (1) turns into*

$$\begin{aligned}
& \max_{a\in[0,1]} \left\{\frac{a}{\exp(b) + (1-\exp(b))a}\left(\Pr(N\overline{y}_1 < n^*) + (1-\lambda^*)\Pr(N\overline{y}_1 = n^*)\right)\right\} \\
= & \max_{a\in[0,1]} \left\{\frac{\exp(b)a}{1 + (\exp(b)-1)a}\left(\Pr(N\overline{y}_1 > n^*) + \lambda^*\Pr(N\overline{y}_1 = n^*)\right)\right\},
\end{aligned}$$

*where the probabilities refer to $N\overline{y}_1$ being binomially distributed with parameters $(N, a)$.*

The basic intuition behind these results resembles the "robustness" part of propositions 4 and 5: For small enough disturbances, the least favorable prior still renders the data informative, and the optimal decision rule will be a threshold rule. However, it turns out that the very strong robustness result for the case of two unknown treatments relied not only on this idea, but also on the decision problem's symmetry. This symmetry is broken when one treatment is known, and in order to keep minimax regret equalized between (worst-case) type I and type II error scenarios, the decision parameters $(n^*, \lambda^*)$ must be (smoothly) adjusted.

Here is a brief assessment of how other decision rules would cope with these types of problems. Maximin utility would still be caught in the triviality trap. A Bayesian could, of course, put a prior on contaminations. More interestingly, consider a robust Bayesian who accepts this section's framework as description of her sets of priors. Propositions 4 through 7 would apply, so in the case of two unknown treatments, the $\Gamma$-minimax regret decision is completely unaffected by the contamination's presence as long as $\varepsilon$ or $b$ are small enough. Tick loss functions will encounter the exact same problem as in the preceding section. There are states of the world under which the sign of the average treatment effect (hence the identity of the better treatment) varies between sampling and treatment population, in which case sample signals are actually misleading; minimaxing of tick loss functions will guard against this by prescribing a no-data rule. Finally, to use hypothesis tests in this problem, one would have to implicitly impute the missing data when formulating the hypothesis. For testing an innovation, a natural approach would be to recommend the innovation if the null hypothesis $H_0 : \mu_{11} \leq \mu_{01}/(1-\varepsilon)$ is rejected. This point null is least favorable within a compound null that allows for all possible values of $(\mu_{00}, \mu_{10})$. In effect, one then assumes that the missing data are least favorable for the innovation. Commonsensically, this may well be reasonable if one is specifically concerned about false positives, but then, if there are such concerns, one might rather want to write them into the loss functions.

# 5 Treatment Choice with Covariates

## 5.1 A Problem with Minimax Regret

Most real-world economic decision and prediction problems involve covariates. The according modification of the above problem is, therefore, of obvious interest. The crucial part of this modification is that a covariate $X \in \mathcal{X}$ exists and is observable for both sample and treatment subjects, thus the decision maker can attempt to (implicitly) estimate the mean regression $(\mu_{0x}, \mu_{1x}) \equiv \mathbb{E}(Y_{0x}, Y_{1x})$ and can provide accordingly conditioned treatment recommendations.

To begin, assume that $\mathcal{X}$ is finite with $K$ elements and that the distribution of $X$ is known. Potential outcomes are, then, random variables $Y_{tx}$ that depend on treatment as well as covariate. A state of the world $s$ is a distribution $P\left((Y_{0x}, Y_{1x})_{x\in\mathcal{X}}\right)$ with marginals $s_x \equiv P(Y_{0x}, Y_{1x})$. Complete specification of the problem also requires a state space; for the moment, let this be the perhaps most natural, and largest one, namely $\mathcal{S} = \Delta\left\{0,1\right\}^{2K}$. Define also $\mathcal{S}_x \equiv \{s_x : s \in \mathcal{S}\}$. A sample $\omega$ collects realizations $(t_n, x_n, y_n)$, where the distributions of both $T$ and $X$ depend on the sample design and $y_n$ is an independent realization of $Y_{t_n x_n}$.[10] A statistical treatment rule maps samples $\omega$ into vectors of treatment assignment probabilities $\delta(\omega) \in [0,1]^K$, whose components $\delta_x(\omega)$ are identified with probabilities of assigning treatment 1 to subjects with covariate value $x$. A treatment rule's risk function is $u(\delta, s) \equiv \sum_{x\in\mathcal{X}} \Pr(X = x)\left(\mu_{0x}\left(1 - \mathbb{E}\delta_x(\omega)\right) + \mu_{1x}\mathbb{E}\delta_x(\omega)\right)$, where $(\mu_{0x}, \mu_{1x}) \equiv \mathbb{E}(Y_{0x}, Y_{1x})$. Regret is $R(\delta, s) \equiv \max_{d\in\mathcal{D}} u(d, s) - u(\delta, s)$ as before.

[10] To keep the argument simple, I will again exclude sequential sample designs, thus $(T, X)$ cannot depend on lagged realizations of $Y_{t_n x_n}$.

An obvious question is whether in this extended problem, minimax regret is achieved by pooling observations across covariates, by conditioning on covariates, or by something in between. The intuitive trade-off is between the resolution of a decision rule and its sensitivity to sampling variation. Using large deviations bounds, Manski (2004) discovered that a lower bound on regret incurred by pooling exceeds an upper bound incurred by conditioning on covariates for rather small sample sizes. The tentative conclusion was that prevailing practice may err in the direction of too much pooling.

Stoye (2009a) re-analyzed this problem in terms of finite sample regret and found that Manski's result merely approximates a much stronger, and pathological one.[11] Minimax regret recommends to condition treatment choice on all available covariates, even if this leads to empty sample cells. This conclusion extends to cases of many sample cells and small samples, where the resulting decision rule is essentially a no-data rule. Indeed, if a covariate takes infinitely many values, then a no-data rule achieves minimax regret. This result reverses the thrust of previous findings, raising more questions about minimax regret than about prevailing practice. It motivates this section's analysis.

Here is an intuition for why the problem obtains. Both Manski (2004) and Stoye (2009a) use the state space $\mathcal{S} = \Delta[0,1]^{2K}$ specified above. This state space allows for Nature to choose priors under which $s_x$ and $s_{x'}$ are independent random variables. Observations of treatment outcomes for covariate $x$ are then uninformative about potential outcomes for covariate $x'$, thus treatment rules which best respond to such priors separate inference across covariates. But the additive separability by covariate of $R(\delta, s)$ can be used to show that priors of this sort, in turn, best respond to these decision rules. An alternative, non-game theoretic intuition is as follows: Ceteris paribus, minimax regret selects for parameter values that make it hard to learn. In the presence of covariates, cross-covariate learning is hardest if cross-covariate signals are vacuous. Indeed, if a covariate takes infinitely many values, then uniform learning about the better treatment is impossible, and any minimax-type rule that values learning is expected to encounter trouble. For example, maximin utility or minimax loss with respect to the "tick" loss function (i.e., a trade-off between size and power) will have the exact same problem here.

However, uniform learning is impossible because some perhaps highly implausible priors are allowed for. Specifically, $(\mu_{0x}, \mu_{1x})$ and $(\mu_{0x'}, \mu_{1x'})$ are equally likely to be very similar (indeed identical under conditions on $P(X)$) and very different. This is substantively implausible. Consider a medical trial in which the race of sample subjects or, even worse, their date of birth is recorded. Clearly, there exists prior information to the effect that these covariates will matter only to a limited degree. In more technical terms, the least favorable prior is demanding of the state space, which must essentially be a

[11] A precise statement requires additional notation not otherwise used in this paper, so I will try to get away with a paraphrase.

Cartesian product of covariate-wise state spaces.[12] The problem may, therefore, lie with (technically) a too permissive state space that (substantively) reflects underspecification of prior information. One might hope that it can be alleviated by introducing plausible prior constraints. The next subsection introduces some such constraints and shows that they can have dramatic effects.

### 5.2 Limiting the Effect of Covariates

This section does not make any assumptions about $\mathcal{X}$, thus the setting is technically more general than the one for which pathological results were obtained. In particular, $\mathcal{X}$ could be continuous, so that the below result could be rephrased in terms of nonparametric minimax regret mean regression. To ensure well-definedness of state space, let $\mathcal{X}$ be endowed with some algebra $\Sigma$; a state is then a $\Sigma$-measurable function that maps covariate values $x \in \mathcal{X}$ into distributions $P(Y_{0x}, Y_{1x}) \in \Delta\{0,1\}^2$. Finally, for this section's result to hold as stated in the case of testing an innovation, I need to impose that in the case where $\mu_{0x}$ is known, it is constant. Some remarks on the more general case will be provided.

The innovation is to restrict the effect that $X$ can have on the marginal distribution $P(Y_{tx})$. This will be done through bounding $||P(Y_{tx}), P(Y_{tx'})||$, where $||\cdot||$ can stand for one of several metrics. I begin by analyzing the effect of the following assumption:

**Assumption 3**

$$|\mu_{tx} - \mu_{tx'}| \leq \kappa$$

*for all* $x, x' \in \mathcal{X}$.

Findings will first be stated for this assumption and then be extended to a number of other probability metrics. Whichever metric is used, some restrictions of this type will be exceedingly plausible in many cases. Again, medical trials constitute a nice example. Researchers would probably be willing to bound the effect of race on outcomes and to quite severely bound the effect of birthdays. At the same time, it will rarely be honest to bound a covariate's effect at precisely zero. Even birthdays might have a very small effect on reaction to medication (because they proxy for season of birth), so that restricting their effect to be exactly zero is likely an approximation. This is why proposition 2 below may be of some interest. In words, it states that for every $N$, there exists $\kappa$ that should be thought of as "small but positive" s.t. covariates whose cumulative effect can be bounded by $\kappa$ should be completely ignored. A precise statement is somewhat more long-winded and can be differentiated depending on how much of the sample design is exogenous.

**Proposition 8** *Let there be a covariate $X$. Let $\overline{\delta}_1^*$ and $\overline{\delta}_2^*$ be the decision rules that mimic the preceding ones and ignore the existence of $X$, thus $\overline{\delta}_{t,x}^*(\omega) = \delta_t^*(\omega)$ for all $(x, \omega)$ and $t = 1, 2$. Then:*

[12] More formally, one must have that for any $(s_x)_{x\in\mathcal{X}} \in \times_{x\in\mathcal{X}}\mathcal{S}_x$, there exists a state $s \in \mathcal{S}$ with marginals $(s_x)_{x\in\mathcal{X}}$.

***(i)*** *Let both treatments be unknown. If the sample design is any of matched pairs, unconstrained randomization, or constrained randomization, the sample is a simple random sample with respect to $X$, and assumption 1 holds with $\kappa \leq 2R_1^*(N)$, then $\overline{\delta}_1^*$ achieves minimax regret.*

***(ii)*** *Let both treatments be unknown. If the sample design is any of matched pairs, unconstrained randomization, or constrained randomization, the sample design with respect to $X$ is a choice variable, and assumption 1 holds with $\kappa \leq 2R_1^*(N)$, then simple random sampling in conjunction with $\overline{\delta}_1^*$ achieves minimax regret.*

***(iii)*** *Let both treatments be unknown. If the sample design with respect to both $T$ and $X$ is a choice variable, and assumption 1 holds with $\kappa \leq 2R_1^*(N)$, minimax regret is achieved by simple random sampling in conjunction with any of the above sample designs and $\overline{\delta}_1^*$.*

***(iv)*** *Let $\mu_{0x}$ be known and constant and let the sample be a simple random sample with respect to $X$. If assumption 1 holds with $\kappa \leq 4R_2^*(N)$, then $\overline{\delta}_2^*$ achieves minimax regret.*

***(v)*** *Let $\mu_{0x}$ be known and constant and let sample design with respect to $X$ be a choice variable. If assumption 1 holds with $\kappa \leq 4R_2^*(N)$, then minimax regret is achieved by simple random sampling in conjunction with $\overline{\delta}_2^*$.*

*In all cases, if $X$ is continuous, then the bound is tight in the sense that if $\kappa$ exceeds the relevant threshold, the above decision rules do not in general achieve minimax regret. Finally:*

***(vi)*** *No-data rules strictly fail to achieve minimax regret if assumption1 holds with any $\kappa < 1$.*

Again, the proposition essentially states that if the potential effect of $X$ is small enough, then its presence should be completely ignored. An intuition for this result goes as follows: If the decision maker ignores $X$, then Nature has two types of responses. By playing states in which nothing is lost by ignoring $X$, she can mimic the least favorable prior from proposition 1 and thus enforce the regret values from corollary 1. By playing states in which the correct treatment is different for different $x$, Nature can ensure that the above decision rules assign suboptimal treatments with rather high probabilities; but the bound on a covariate's effect can be exploited to bound regret by $\kappa/2$ uniformly over states of this type and even by $\kappa/4$ if $\mu_{0x}$ is known and constant (the bound of $\kappa/2$ would go through without constancy). It follows that if $\kappa$ is small enough, Nature is better off mimicking the least favorable prior from proposition 1.

Most of this argument not require $\mu_{0x}$ constant nor $X$ continuous, and as a result, the general finding that the covariate should be ignored for some positive but sufficiently small $\kappa$ does not depend on these additional restrictions. The effect of these restrictions is rather the following. If $X$ is continuous, then Nature can achieve the regret bounds, hence they are tight in the sense stated. In the case of testing an innovation, constancy of $\mu_{0x}$ enables the sharpening of the bound to $\kappa/4$ and also ensures that Nature can mimic the least favorable prior underlying proposition 1 by making $\mu_{1x}$ constant. If constancy of

$\mu_{0x}$ fails, then statements (iv) and (v) would fail as stated, but it is clear from the rpoof that they would still apply with $4R_2^*(N)$ replaced by some (smaller) threshold value $\kappa_N > 0$.

Finally, it is worth noting that while a full equilibrium analysis for arbitrary $P(X)$ and $\kappa$ seems elusive, the amelioration of the "triviality trap" discovered in Stoye (2009a) is substantially stronger than indicated by parts (i)-(v). Minimal expansion of the proof of (i)-(v) shows that for any $\kappa < 1$, decision rules $\delta_1^*$ respectively $\delta_2^*$, applied without any regard for $X$, incur maximal regret of strictly less than 1/2 (if $\mu_0$ is unknown) respectively $\mu_0(1-\mu_0)$ (if $\mu_0$ is known), whereas any no-data rule incurs at least these values. Thus, it is actually the no-data result that is in some sense sensitivie: Imposing assumption 1 with any nonvacuous choice of $\kappa$ eliminates it.

I now extend the proposition to other notions of distance between probability distribution. Fix any probability distributions $P$ and $Q$ on the Borel sets in $\mathbb{R}$. Define the Total Variation distance $||\cdot||_{TV}$ by $||P,Q||_{TV} = \max_{E\in\Sigma}|P(E)-Q(E)|$, the log odds ratio distance by $||P,Q||_{LOR} = \max_{E\in\Sigma}|\log(P(E)/Q(E))|$, and the Kullback-Leibler divergence (a.k.a. relative entropy) by $D_{KL}(P||Q) = \int \log(dP/dQ)dP$ (assuming existence of these quantities).[13] Then the following obtains:

**Proposition 9** ***(i)*** *Assume that* $||P(Y_{tx}),P(Y_{tx'})||_{TV} \leq \kappa$ *for all* $x,x' \in \mathcal{X}$*, where* $\kappa > 0$ *is a user specified parameter. Then the conclusion of proposition 2 obtains.*

***(ii)*** *Assume that* $||P(Y_{tx}),P(Y_{tx'})||_{LOR} \leq \gamma$ *for all* $x,x' \in \mathcal{X}$*, where* $\gamma > 0$ *is a user specified parameter. Then the conclusion of proposition 2 obtains if*

$$\gamma \leq 2\log\frac{1+2R_1^*(N)}{1-2R_1^*(N)} \text{ or } \gamma \leq 2\log\frac{1+4R_2^*(N)}{1-4R_2^*(N)},$$

*depending on whether one or both treatments are unknown, and with the caveat that the threshold may not be best possible if* $\mu_0$ *is known.*

***(iii)*** *Assume that* $D_{KL}(P(Y_{tx})||P(Y_{tx'})) \leq \delta$ *for all* $x,x' \in \mathcal{X}$*, where* $\rho > 0$ *is a user specified parameter. Then the conclusion of proposition 2 obtains if*

$$\rho \leq 2R_1^*(N)\log\frac{1+R_1^*(N)}{1-R_1^*(N)} \text{ or } \rho \leq 4R_2^*(N)\log\frac{1+2R_2^*(N)}{1-2R_2^*(N)},$$

*depending on whether one or both treatments are unknown, and with the caveat that the threshold may not be best possible if* $\mu_0$ *is known.*

The proposition can be established by showing that bounds in these different metrics induce bounds on $|\mu_{tx}-\mu_{tx'}|$ and then invoking proposition 2 (modulo a change in variables for (ii) and (iii)). The

[13] If $P$ and $Q$ are Bernoulli with parameters $\mu_P$ and $\mu_Q$, then these simplify to $||P,Q||_{TV} = |\mu_P - \mu_Q|$ (and this then also coincides with the Kolmogorov-Smirnov and other distances), $||P,Q||_{LOR} = |\log((\mu_P(1-\mu_Q))/(\mu_Q(1-\mu_P)))|$, and $D_{KL}(P||Q) = \mu_P\log(\mu_P/\mu_Q)+(1-\mu_P)\log((1-\mu_P)/(1-\mu_Q))$. The proof of proposition 3 uses these simplifications, but the result does not materially depend on binary outcomes. Also, bounding the log odds ratio can alternatively be justified by imposing a logit model of treatment outcomes and limiting the coefficient on $X$ in the model, as will be shown in detail in example 2 below.

proof strategy might seem inefficient because bounds on Kullback-Leibler divergences and log odds ratios are strictly stronger than the implied bounds on differences in means, thus there might be slack in the threshold values given above. But this is not so: The least favorable priors which render proposition 2 tight also maximize $|\mu_{tx} - \mu_{tx'}|$ subject to the respective bound on $D_{KL}$ or $||\cdot||_{LOR}$. Thus, the thresholds given in proposition 3 are tight in the case of two unknown treatments, and in the case of testing an innovation, cannot be improved without using knowledge of a particular $\mu_{0x}$. One could, therefore, think of bounding $|\mu_{tx} - \mu_{tx'}|$ as an efficient approach in the sense of getting the desired effect through the weakest among a set of restrictions. Noting that $|\mu_{tx} - \mu_{tx'}| = 1$ implies infinite log odds ratio distance as well as Kullback-Leibler divergence, one can eliminate ni-data rules by imposing any finite bound on either quantity.

The result is of interest not least because the existence of a covariate with unrestricted domain causes problems for some other decision criteria. For example, the mean regression of interest $(\mu_{0x}, \mu_{1x})$ may have a sufficiently rich domain that a noninformative prior appears elusive; thus imposing a prior means to impose a substantial amount of structure even beyond possible bounds on the effects of covariates.[14] Having said that, there is a Bayesian interpretation to this section's results. Consider a robust (multi-prior) Bayesian who is willing to allow any prior s.t. $|\mu_{tx} - \mu_{tx'}| \leq \kappa$ holds with probability 1.[15] Proposition 2 applies to this Bayesian's $\Gamma$-minimax regret decision; thus, a robust Bayesian might want to completely ignore the covariate as well.

Minimax-type rules based on tick loss functions will encounter serious problems in the present setting. Bounding the covariate's effect does not affect the "no-data" conclusion from these loss functions at all, unless the bound is set at zero. The problem is that this section's assumptions do not imply that uniform learning about the correct treatment assignment becomes feasible; they merely ensure that under those parameter values where it remains infeasible, the penalty for not learning is small. Minimax regret, because it combines concern for picking the right treatment with concern for the damage caused by not doing so, is able to exploit this feature.[16]

This section's result can be illustrated numerically. Table 1 displays the threshold values from proposition 7 for different $N$, both for the case of two unknown treatments and for testing an innovation, in which case the thresholds depend on $\mu_0$. The values of $\mu_0$ selected are the ones from tables in Manski

[14] A uniform prior on the set of possible $\mu_{0x} : \mathcal{X} \to [0,1]$ would certainly have to imply a uniform distribution over the indicator functions of events in $\mathcal{X}$, hence over the the power of $\mathcal{X}$ – but this distribution does not even exist for $\mathcal{X} = [0,1]$.

[15] Note this is written as a restriction on parameter values. It is not sufficient to impose a similar condition for prior expectations.

[16] These problems of tick loss functions are known, and one frequent resolution is to record positive loss only if the difference between expected outomes exceeds some threshold. This author is skeptical about such solutions: They introduce an additional tuning parameter and frankly seem to be crude approximations of regret.

| N | 1 | 3 | 5 | 10 | 20 | 50 | 100 | 200 | 500 | 1000 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\boldsymbol{\mu_0}$ **unknown** | 0.2500 | 0.1760 | 0.1411 | 0.1086 | 0.0764 | 0.0482 | 0.0340 | 0.0240 | 0.0158 | 0.0108 |
| $\boldsymbol{\mu_0}$**= .5** | 0.2500 | 0.1760 | 0.1411 | 0.1086 | 0.0764 | 0.0482 | 0.0340 | 0.0240 | 0.0158 | 0.0108 |
| $\boldsymbol{\mu_0}$**= .25** | 0.3599 | 0.1577 | 0.1379 | 0.0929 | 0.0664 | 0.0416 | 0.0295 | 0.0208 | 0.0132 | 0.0093 |
| $\boldsymbol{\mu_0}$**= .05** | 0.1632 | 0.1261 | 0.0991 | 0.0572 | 0.0343 | 0.0217 | 0.0149 | 0.0105 | 0.0066 | 0.0047 |

Table 2: Numerical illustration of propositions 2 and 3.

(2004). Thresholds are the same for unknown $\mu_0$ and for $\mu_0 = .5$ because in thisw case, it can be shown that $R_1^*(N) = 2R_2^*(N)$.

As a practical illustration, consider an example from Stoye (2009a): There is a binary covariate, $X \in \{m, f\}$ say, the decision maker can freely choose the sample design, and both treatments are unknown. Then if the effect of $X$ is unrestricted, the decision maker will want to separately sample males and females and to conduct entirely separate inference (according to proposition 1) within either group, no matter how small the overall sample size. In contrast, if $N = 10$ and the decision maker is willing to bound $\left|\mu_{tm} - \mu_{tf}\right|$ by 10% – which would appear reasonable in many contexts –, the finite sample minimax regret recommendation is to completely ignore gender. Specifically, minimax regret would be achieved by taking a simple random sample of the population, not even recording $X$, assigning within-sample treatment by constrained randomization, and applying $\delta_1^*$. A bound on $\left|\mu_{tm} - \mu_{tf}\right|$ of 1% – still something one would be willing to impose for many obscure covariates – would justify ignoring the covariate at a rather substantive sample size of $N = 1000$.

# 6 Conclusion

Manski (2004) placed minimax regret treatment choice on the agenda of a small but active literature in econometrics. This literature was quite successful in characterizing finite sample minimax regret experimental designs and treatment rules in stylized problems, but recently also contained some findings that would have to worry proponents of minimax regret. In particular, no-data recommendations, which are a core argument against maximin utility in this context, affect minimax regret as soon as covariates are introduced.

This paper extended previous findings in a direction that may be more charitable to minimax regret. While the introduction of covariates with arbitrarily large effects as well as of unrestricted disagreement between response functions in sample and treatment population causes massive problems for minimax regret (and, indeed, for any other decision rule), minimax regret is locally insensitive to the introduction of these features. To repeat, the interesting finding is not that the effect of these features on minimax regret treatment rules approaches zero as their potential magnitude approaches zero; such a result should be expected from any reasonable treatment rule. It is not obvious, however, that in three of four cases, said effect becomes exactly zero for some positive (and in some cases, arguably quite large) magnitude of the effect. Furthermore, these findings are finite sample results, so they cannot be artifacts of approximations of any kind.

Regarding the methodological significance of these findings, I emphasize that they are driven by a specific feature of minimax regret not shared by many other decision criteria. Bayesian and maximin utility criteria attempt to maximize some increasing functional of the risk function $u(\delta, s)$, thus they intrinsically care about "good" versus "bad" states but not about how much efficiency loss – meaning underperformance relative to what could have been achieved ex post – a decision rule causes in a given state. Hypothesis testing is based on tick loss functions that have the opposite intuition: They are only sensitive to whether the ex post correct decision was made and not to the extent of utility loss caused. Minimax regret seems combines aspects of both, reflecting a consideration for ex post optimality but also for the stakes at play. The local robustness features of minimax regret decision rules can be directly linked to this feature, that is, it does indeed drive the proofs. This link between the intuition and the technical result is some cause for optimism that similar robustness features could be uncovered with respect to other modifications of the problem. It is hoped that along with axiomatic discussions and empirical applications, findings like these advance our understanding of the trade-offs involved in choosing among treatment rules and experimental designs for real-world, sample based decision problems.

# A Proofs

**Least Favorable Priors Underlying Proposition 1** This proposition is not new, but an explanation that in particular exhibits the least favorable priors is helpful. It is easy to see that regret depends on states of the world $s$ only through $(\mu_0, \mu_1)$, so states will be identified with these parameters.

**(i)** Assume that $N$ is odd. Then a Nash equilibrium of the fictitious game is given by $(\delta_1^*, \pi_1^*)$, where $\pi_1^*$ randomizes evenly over $\{(a^*, 1-a^*), (1-a^*, a^*)\}$ for some $a^* > 1/2$ to be characterized. To verify this, note that to be a best response to $\delta_1^*$, a state $(\mu_0, \mu_1)$ must solve

$$\max_{(\mu_0,\mu_1)\in[0,1]^2} \max \left\{ \begin{array}{c} (\mu_1 - \mu_0) B\left(\frac{N-1}{2}; N, \frac{1+\mu_1-\mu_0}{2}\right), \\ (\mu_0 - \mu_1)\left(1 - B\left(\frac{N-1}{2}; N, \frac{1+\mu_1-\mu_0}{2}\right)\right) \end{array} \right\},$$

where $B(\cdot; N, p)$ denotes the binomial c.d.f. with indicated parameters. The expression uses that $\delta_1^*$ can alternatively be written as $\delta_1^*(\omega) = 1\{z > 0\}$, where $z \equiv (1 + I_N/N)/2$ is distributed as a binomial variable with parameters $(N, (1 + \mu_1 - \mu_0/2))$. The objective is symmetric in $(\mu_0, \mu_1)$, thus $(a^*, b^*)$

maximizes it iff $(b^*, a^*)$ does; furthermore, $(\mu_0, \mu_1)$ enters the objective only through $\Delta \equiv \mu_1 - \mu_0$. It follows that one best response to $\delta_1^*$ is $\pi_1^*$ with $a^* = (1 - \Delta^*)/2$, where $\Delta^*$ solves

$$\max_{\Delta \in [0,1]} \phi(\Delta), \phi(\Delta) = \Delta B\left(\frac{N-1}{2}; N, \frac{1+\Delta}{2}\right).$$

This establishes equilibrium. The value of the game is $R_1^*(N) = \phi(\Delta^*)$, which upon algebraic evaluation leads to the corollary.

Now let $N$ be even. $I_N$ is necessarily even in this case, so if $I_N \neq 0$, $\delta_1^*(\omega)$ is invariant to dropping the last observation from the sample. If $I_N = 0$, the 50/50 tie-breaking can equivalently be achieved by dropping the last observation from the sample. Thus if the planner plays $\delta_1^*$, Nature's best response problem is as if $\delta_1^*$ were played on a sample of size $(N-1)$, and $\pi_1^*$ (computed for $N-1$, of course) is a best response. On the other hand, $\delta_1^*$ is a best response to any prior with the symmetry properties of $\pi_1^*$ for any sample size. Thus an equilibrium has been found, and $R_1^*(N) = R_1^*(N-1)$ in this case.

As a side note, Schlag (2006) recommends constrained randomization (the first treatment assignment is by coin toss and alternating thereafter) which (as well as matched pairs for $N$ even) achieves minimax regret as well and (as evident from algebra in Schlag (2006)) weakly dominates independent randomization in the benchmark problem. This dominance will not apply in the below modifications, however, because independent randomization is preferred in states of the world where the signal generated by observable populations is misleading.

**(ii)** Let sample design be a choice variable. Then $\delta_1^*$ combined with *any* sample design is a best response to $\pi_1^*$, but $\pi_1^*$ is a best response to $\delta_1^*$ in combination with the above (and possibly other) sample designs, thus a Nash equlibrium has been found.

**(iii)** The least favorable prior $\pi_2^*$ randomizes over $\{(\mu_0, b^*), (\mu_0, c^*)\}$, where $(b^*, c^*)$ are characterized by

$$\begin{aligned} b^* &= \arg\max_{b \in [\mu_0, 1]} \left\{ (b - \mu_0) \left[ \sum_{n < n^*} \binom{N}{n} b^n (1-b)^{N-n} + (1 - \lambda^*) \binom{N}{n^*} b^{n^*} (1-b)^{N-n^*} \right] \right\} \\ c^* &= \arg\max_{c \in [0, \mu_0]} \left\{ (\mu_0 - c) \left[ \sum_{n > n^*} \binom{N}{n} c^n (1-c)^{N-n} + \lambda^* \binom{N}{n^*} c^{n^*} (1-c)^{N-n^*} \right] \right\}. \end{aligned}$$

For this to be a Nash equilibrium, it is needed that both maximization problems have the same value, hence $(n^*, \lambda^*)$ must fulfil condition (1). Randomization need not be even; weights are rather determined by the requirement that $\delta_2^*$ be Bayes against $\pi_2^*$. The corollary follows by evaluating the above maximization problems.

**Proposition 4** The proof will be conducted explicitly for unconstrained randomization with $N$ odd; the extension to the other scenarios is as before.

The Nash equilibrium of the fictitious game is as follows. It is easy to see that regret only depends on the marginal distributions of $(Y_{00}, Y_{01}, Y_{10}, Y_{11})$ and thus on their expectations $(\mu_{00}, \mu_{01}, \mu_{10}, \mu_{11})$, hence identify states $s$ with vectors $(\mu_{00}, \mu_{01}, \mu_{10}, \mu_{11}) \in [0,1]^4$. Then a Nash equilibrium of the fictitious game is given by $(\delta^*_{MD}, \pi^*_{MD})$, where $\pi^*_{MD}$ randomizes evenly over two states of the form $(1, a, 0, 1-a)$ and $(0, 1-a, 1, a)$, where $a > 1/2$ if $\varepsilon < \varepsilon^*_N$ and $a = 1/2$ if $\varepsilon \geq \varepsilon^*_N$.

**Step 1: Simplifying Natures best response problem.** Nature's best-response problem can be written as

$$\max_{(\mu_{00},\mu_{01},\mu_{10},\mu_{11})\in[0,1]^4} \max \left\{ \begin{array}{c} ((1-\varepsilon)(\mu_{11}-\mu_{01}) + \varepsilon(\mu_{10}-\mu_{00}))(1-\alpha^*\mathbb{E}\delta_1(\omega) - (1-\alpha^*)\mathbb{E}\delta_3(\omega)), \\ ((1-\varepsilon)(\mu_{01}-\mu_{11}) + \varepsilon(\mu_{00}-\mu_{10}))(\alpha^*\mathbb{E}\delta_1(\omega) + (1-\alpha^*)\mathbb{E}\delta_3(\omega)), \end{array} \right\}, \tag{2}$$

where $\delta_1$ and $\delta_3$ are the decision rules $\delta^*_1$ and $\delta^*_3$ defined in the text. (Dropping the asterisks anticipates an upcoming reduction in DM's action space.) This can be substantially simplified. First, recall from the proof of proposition 2 that $z \equiv (I_N + N)/2$ is distributed binomial with parameters $(N, (\mu_{11} + 1 - \mu_{01})/2)$. It follows that both $\mathbb{E}\delta_1(\omega)$ and $\mathbb{E}\delta_3(\omega)$ depend on the state of the world only through $\Delta \equiv \mu_{11} - \mu_{01}$. Thus, write $f_1(\Delta) \equiv \mathbb{E}\delta_1(\omega)$, $f_3(\Delta) \equiv \mathbb{E}\delta_3(\omega)$. Note furthermore symmetry of the two functions compared in the max-operator and the linear way in which $(\mu_{00}, \mu_{10})$ enter the objective. These observations imply that a vector $(a, b, c, d)$ solves (2) iff either $(a, b, c, d) = (0, \pi^*, 1, \rho^*)$ or $(a, b, c, d) = (1, \rho^*, 0, \pi^*)$, where $(\pi^*, \rho^*)$ solve

$$\max_{(\pi,\rho)\in[0,1]^2} ((1-\varepsilon)(\rho-\pi) + \varepsilon)(1 - \alpha^* f_1(\rho-\pi) - (1-\alpha^*) f_3(\rho-\pi)).$$

Noting that the objective depends on $(\pi, \rho)$ only through $\Delta$, one can further simplify (2) to

$$\max_{\Delta\in[-1,1]} \phi(\Delta;\alpha^*,\varepsilon), \phi(\Delta;\alpha^*,\varepsilon) = ((1-\varepsilon)\Delta + \varepsilon)(1 - \alpha^* f_1(\Delta) - (1-\alpha^*) f_3(\Delta)), \tag{3}$$

the problem that will be analyzed henceforth. I finally note that $\phi(\Delta; \alpha^*, \varepsilon) < 0$ for $\Delta < -\varepsilon/(1-\varepsilon)$, so attention will be restricted to $\Delta \in [-\varepsilon/(1-\varepsilon), 1]$ henceforth.

**Step 2: Introducing a simplified game.** For this and the next few steps, consider a simplified game in which the decision maker's strategy set consists of randomizations over $\{\delta_1, \delta_2\}$. Identify feasible strategies for DM with probabilities $\alpha$ of playing $\delta_1$. Nature, on the other hand, is restricted to choosing $\Delta$ (possibly at random); every choice of $\Delta$ is identified with even randomization over $(1, (1+\Delta)/2, 0, (1-\Delta)/2), (0, (1-\Delta)/2, 1, (1+\Delta)(/2))$. This game has a Nash equilibrium (Glicksberg (1952)). Nature's best-response problem in this game is (3), and the previous step's analysis showed that any best response for Nature in this game is also a best response for her in the original one. On the other hand, it is easy to see the following: If Nature's strategy is supported on $(0, 1]$, then $\alpha^* = 1$

is a unique best response for DM in both the simplified and the original game. If Nature plays $\Delta^* = 0$, then the sample data are noise and any choice of $\alpha$ is a best response in both the simplified and the original game; this equilibrium will henceforth be called "pooling." If Nature's strategy is supported on $[-1, 0)$, then $\alpha^* = 0$ is a unique best response in the simplified (albeit not the original) game. The latter case (and other, more intricate ones) will not in fact obtain, so that equilibria of the simplified game correspond to equilibria of the original one.

**Road map for steps 3 to 6.** The next steps involve some algebra; here is a brief explanation of what's going on. There are two types of equilibria, both of which involve pure strategies by Nature in the simplified (albeit not the original) game. If $\varepsilon \leq \varepsilon_N^*$, then Nature plays $\Delta^* \geq 0$ and DM plays $\alpha^* = 1$; else, Nature plays $\Delta^* = 0$ and DM plays some $\alpha^*$ to be characterized. After dealing with the special case of $\varepsilon \geq 1/2$, this finding will emerge in three steps. Conjecture that $\alpha^* = 1$, then $\phi$ can be shown to be quasiconcave in $\Delta$, thus Nature has a unique best response $\Delta^*$. For equilibrium to obtain, one must have $\Delta^* \geq 0$, which (in view of quasiconcavity) requires $\phi'(0; 1, \varepsilon) \geq 0$, which turns out to obtain iff $\varepsilon \leq \varepsilon_N^*$ (step 4). Next, one can show that if there exists an equilibrium with $\Delta^* = 0$ for some $\varepsilon$, then there also exists such an equilibrium for every $\varepsilon' \geq \varepsilon$. In words, as the distortion $\varepsilon$ grows, the equilibrium will potentially switch from informative to pooling but never back. Because $\Delta^* = 0$ for $\varepsilon = \varepsilon_N^*$, equilibrium will be pooling for $\varepsilon > \varepsilon_N^*$ (step 5). Closed-form expressions for $\alpha^*$ can now be found by evaluating the necessary condition $\phi'(0; \alpha^*, \varepsilon) = 0$ (step 6).

**Step 3: Equilibrium when $\varepsilon \geq 1/2$.** Let $\varepsilon \geq 1/2$, then the projection of $\delta_3^*$ onto $[0, 1]$ never binds and $\delta_3^*$ is linear with expectation $f_3(\Delta) = \frac{1}{2} + \frac{1-\varepsilon}{2\varepsilon}\Delta$. The equilibrium is characterized by $\alpha^* = 0$ in this case. To verify this, simplify (3) to

$$\max_{\Delta \in [-1,1]} ((1-\varepsilon)\Delta + \varepsilon) \frac{1}{2} \left(1 - \frac{1-\varepsilon}{\varepsilon}\Delta\right) = \max_{\Delta \in [-1,1]} \left\{ \frac{\varepsilon}{2} - \frac{(1-\varepsilon)^2}{2\varepsilon}\Delta^2 \right\},$$

thus a unique best response is $\Delta^* = 0$.

For future use, observe also that

$$\frac{\partial^2 \phi(\Delta; 0, \varepsilon)}{\partial \Delta \partial \varepsilon} = 2\Delta \left( 2\frac{1-\varepsilon}{\varepsilon} + \frac{(1-\varepsilon)^2}{\varepsilon^2} \right),$$

which is strictly positive for $\varepsilon < 1/2$. It follows that for $\varepsilon < 1/2$, any best response to $\alpha^* = 0$ is characterized by $\Delta^* > 0$, which in turn leads to $\alpha^* = 1$ as best response, breaking the conjectured equilibrium. Thus, if $\varepsilon < 1/2$, there exists no equilibrium with $\alpha^* = 0$. For the remainder of this proof, presume $\varepsilon \leq 1/2$.

**Step 4: Equilibrium when $\alpha^* = 1$.** Next, conjecture that $\alpha^* = 1$, simplifying (3) to

$$\max_{\Delta\in[-1,1]} \phi(\Delta;\alpha^*,\varepsilon), \phi(\Delta;\alpha^*,\varepsilon) = ((1-\varepsilon)\Delta+\varepsilon)(1-f_1(\Delta)).$$

For equilibrium to obtain, this has to be solved by some $\Delta^* \geq 0$. To verify conditions for this, note first that $\phi$ is quasiconcave. To see this, write

$$f_1(\Delta) = 1 - B\left(\frac{N-1}{2}; N, \frac{1+\Delta}{2}\right) = 1 - \underbrace{\frac{N+1}{2}\left[\begin{array}{c} N \\ \frac{N-1}{2}\end{array}\right]}_{\rho}\int_0^{\frac{1-\Delta}{2}} \left(\frac{t}{2}\right)^{\frac{N-1}{2}}\left(\frac{1-t}{2}\right)^{\frac{N-1}{2}} dt$$

and note that

$$\begin{aligned}
\phi'(\Delta;1,\varepsilon) &= (1-\varepsilon)(1-f_1(\Delta)) - ((1-\varepsilon)\Delta+\varepsilon)\, f_1'(\Delta) \\
\phi''(\Delta;1,\varepsilon) &= -2(1-\varepsilon)f_1'(\Delta) - ((1-\varepsilon)\Delta+\varepsilon)\, f_1''(\Delta) \\
f_1'(\Delta) &= \rho 2^{-N}(1-\Delta^2)^{\frac{N-1}{2}} \\
f_1''(\Delta) &= -\rho 2^{-N}(N-1)\Delta(1-\Delta^2)^{\frac{N-3}{2}}.
\end{aligned}$$

This immediately reveals that $\phi''(\Delta;1,\varepsilon) < 0$ whenever $\Delta \in [\varepsilon/(\varepsilon-1), 0]$. Furthermore,

$$\phi''(\Delta;1,\varepsilon) = \rho 2^{-N}(1-\Delta^2)^{\frac{N-3}{2}}\left[\underbrace{(N+1)(1-\varepsilon)\Delta^2 + \varepsilon(N-1)\Delta - 2(1-\varepsilon)}_{\psi(\Delta)}\right].$$

The sign of $\phi''$ equals the sign of $\psi$ on $[0,1]$ except that $\phi''(1;1,\varepsilon) = 0$. Noting that $\psi(0) < 0$, $\psi(1) > 0$, and $\psi$ has at most two sign changes on the real line because it is quadratic in $\Delta$, it follows that $\phi$ is first concave and then convex on $[\varepsilon/(\varepsilon-1), 1]$. Also noting that $\phi(\varepsilon/(\varepsilon-1);1,\varepsilon) = \phi(1;1,\varepsilon) = 0$, one finds that $\phi$ is quasiconcave on $[\varepsilon/(\varepsilon-1), 1]$ with an interior maximum characterized by a first-order condition; furthermore, this maximum is characterized by $\Delta^* \geq 0$ iff $\phi'(0;1,\varepsilon) \geq 0$. It remains to write

$$\begin{aligned}
\phi'(0;1,\varepsilon) &= \frac{1-\varepsilon}{2} - \varepsilon f_1'(0) \\
\Longrightarrow \phi'(0;1,\varepsilon) &\geq 0 \text{ iff } \varepsilon \leq \varepsilon_N^* = \frac{1}{1+2f_1'(0)}.
\end{aligned}$$

The latter condition is, therefore, necessary and sufficient for this equilibrium to obtain. The closed-form expression for $\varepsilon_N^*$ provided in the proposition follows upon computing

$$\begin{aligned}
f_1(\Delta) &= \sum_{z>N/2}\binom{N}{z}2^{-N}(1+\Delta)^z(1-\Delta)^{N-z} \\
\Longrightarrow f_1'(0) &= \sum_{z>N/2}\binom{N}{z}2^{-N}(2z-N)
\end{aligned}$$

**Step 5: Analysis of revealing equilibria in general.** Call an equilibrium revealing if Nature's strategy is *not* degenerate at $\Delta^* = 0$. Suppose by contradiction that Nature's strategy is supported on $[-1, 0)$, then DM's best response is to set $\alpha^* = 0$; but this equilibrium was ruled out in step 3. It follows that in any revealing equilibrium, (3) is solved by at least one $\Delta^* > 0$. The value of the equilibrium can therefore be bounded above by

$$\max_{\Delta \in [0,1]} \min_{\alpha \in [0,1]} \phi(\Delta; \alpha, \varepsilon).$$

On this restricted domain,

$$\frac{\partial \phi(\Delta; \alpha^*, \varepsilon)}{\partial \varepsilon} = (1-\Delta)(1 - \alpha^* f_1(\Delta) - (1-\alpha^*) f_3(\Delta)) \leq 1/2$$

for any $(\alpha^*, \varepsilon)$, hence $\frac{\partial}{\partial \varepsilon} \max_{\Delta \in [0,1]} \min_{\alpha \in [0,1]} \phi(\Delta; \alpha, \varepsilon) \leq 1/2$ by an envelope theorem. In contrast, a pooling equilibrium (if it exists) has value $\phi(0; \alpha^*, \varepsilon) = \varepsilon/2$ (independent of $\alpha$) with derivative 1/2. It follows that if there exists a pooling equilibrium for some $\varepsilon$, then there also exists a pooling equilibrium for any $\varepsilon' > \varepsilon$. But inspection of algebra in step 4 reveals that $\Delta^* = 0$, thus the equilibrium is pooling, at $\varepsilon = \varepsilon_N^*$. Hence, there exists a pooling equilibrium for every $\varepsilon \geq \varepsilon_N^*$.

**Step 6: Characterization of the pooling equilibrium.** A necessary condition for a pooling equilibrium is that $\max_{\Delta \in [-1,1]} \phi(\Delta; \alpha^*, \varepsilon) = \varepsilon/2$. As $\phi(0; \alpha^*, \varepsilon) = \varepsilon/2$, this requires a first-order condition to hold at $\Delta^* = 0$, thus

$$\phi'(0; \alpha^*, \varepsilon) = -\varepsilon \left(\alpha^* f_1'(0) + (1-\alpha^*) f_2'(0)\right) + \frac{1-\varepsilon}{2} = 0 \Longrightarrow \alpha^* = \frac{\frac{1-\varepsilon}{2\varepsilon} - f_2'(0)}{f_1'(0) - f_2'(0)}.$$

The closed-form expressions for $\alpha^*$ provided in the proposition follow upon substituting for $f_1'(0)$ from above and using

$$\begin{aligned} f_2(\Delta) &= \sum_{z: \frac{1}{2} + \frac{1-\varepsilon}{\varepsilon}\left(\frac{z}{N} - \frac{1}{2}\right) \geq 0} \binom{N}{z} 2^{-N} (1+\Delta)^z (1-\Delta)^{N-z} \max\left\{\frac{1}{2} + \frac{1-\varepsilon}{\varepsilon}\left(\frac{z}{N} - \frac{1}{2}\right), 1\right\} \\ \Longrightarrow f_2'(0) &= \sum_{z \geq \frac{N(1-2\varepsilon)}{2-2\varepsilon}} \binom{N}{z} 2^{-N} \max\left\{\frac{1}{2} + \frac{1-\varepsilon}{\varepsilon}\left(\frac{z}{N} - \frac{1}{2}\right), 1\right\} (2z - N). \end{aligned}$$

This always leads to a positive expression because $f_2'(0) > f_1'(0)$ and $\frac{1-\varepsilon}{2\varepsilon} \geq f_1'(0)$ can be verified. $\alpha^*$ will equal 0 whenever $\frac{1-\varepsilon}{2\varepsilon} = f_1'(0)$, that is whenever the truncation of $\delta_2$ fails to bind; this is the case iff $\varepsilon \geq 1/2$, confirming algebra from step 3. $\alpha^*$ will equal 1 iff $f_1'(0) = \frac{1-\varepsilon}{2\varepsilon}$, which occurs iff $\varepsilon = \varepsilon_N^*$. $\alpha^* > 1$ if $\varepsilon < \varepsilon_N^*$, in which case the pooling equilibrium does not exist, confirming algebra from step 4.

**Proposition 5** The proof plan is similar to the one of proposition 4. To minimize redundancy, I follow the preceding proof's steps and only point out necessary adjustments. I also omit algebraic steps that are conceptually similar, if different in detail, to ones in the preceding proof. To describe priors, let $(\mu_{01}, \mu_{11})$ denote expected outcomes in the sampling population and $(\mu_0, \mu_1)$ expected outcomes in the treatment population. The least favorable prior then randomizes evenly over two states, one of which has $\mu_{11} = 1 - \mu_{01} = a$ and $\mu_1 = 1 - \mu_0 = \frac{\exp(b)\mu_{11}}{1+(\exp(b)-1)\mu_{11}}$, the other one is generated from this by exchanging the treatment labels. Similar to before, a pooling equilibrium (here meaning that $a = 1/2$) obtains if $b \geq b_N^*$ and a revealing equilibrium (here meaning that $a > 1/2$) obtains if $b < b_N^*$.

**Step 1: Simplifying Nature's best response problem.** Nature's best response problem is

$$\max_{(\mu_{01},\mu_{11},\mu_0,\mu_1)\in[0,1]^4} \max \left\{ \begin{array}{c} (\mu_1 - \mu_0)\left(1 - \alpha^* \mathbb{E}\delta_1(\omega) - (1-\alpha^*)\mathbb{E}\delta_4(\omega)\right), \\ (\mu_0 - \mu_1)\left(\alpha^* \mathbb{E}\delta_1(\omega) + (1-\alpha^*)\mathbb{E}\delta_4(\omega)\right) \end{array} \right\}$$

$$\text{s.t.} \qquad \frac{\mu_{01}}{\exp(b) + (1-\exp(b))\mu_{01}} \leq \mu_0 \leq \frac{\exp(b)\mu_{01}}{1+(\exp(b)-1)\mu_{01}}$$

$$\frac{\mu_{11}}{\exp(b) + (1-\exp(b))\mu_{11}} \leq \mu_1 \leq \frac{\exp(b)\mu_{11}}{1+(\exp(b)-1)\mu_{11}}.$$

Inspection of the way in which $(\mu_0, \mu_1)$ enter the objective reveals immediate simplification to

$$\max_{(\mu_{01},\mu_{11})\in[0,1]^2} \max \left\{ \begin{array}{c} \left(\frac{\exp(b)\mu_{11}}{1+(\exp(b)-1)\mu_{11}} - \frac{\mu_{01}}{\exp(b)+(1-\exp(b))\mu_{01}}\right)\left(1 - \alpha^* \mathbb{E}\delta_1(\omega) - (1-\alpha^*)\mathbb{E}\delta_4(\omega)\right), \\ \left(\frac{\exp(b)\mu_{01}}{1+(\exp(b)-1)\mu_{01}} - \frac{\mu_{11}}{\exp(b)+(1-\exp(b))\mu_{11}}\right)\left(\alpha^* \mathbb{E}\delta_1(\omega) + (1-\alpha^*)\mathbb{E}\delta_4(\omega)\right) \end{array} \right\}.$$

As before, $\mathbb{E}\delta_1(\omega)$ and $\mathbb{E}\delta_4(\omega)$ depend on $(\mu_{01}, \mu_{11})$ only through $\Delta = \mu_{11} - \mu_{01}$. This is not true for the other part of the objective. However, evaluation of first and second derivatives reveals that the problem

$$\max_{\substack{\mu_{01},\mu_{11}\in[0,1]^2: \\ \mu_{11}=\mu_{01}+\Delta}} \left\{ \frac{\exp(b)\mu_{11}}{1+(\exp(b)-1)\mu_{11}} - \frac{\mu_{01}}{\exp(b)+(1-\exp(b))\mu_{10}} \right\}$$

is solved by $\mu_{01} = (1-\Delta)/2 = 1 - \mu_{11}$, and similarly for $\left(\frac{\exp(b)\mu_{10}}{1+(\exp(b)-1)\mu_{01}} - \frac{\mu_{11}}{\exp(b)+(1-\exp(b))\mu_{11}}\right)$. It follows that a best response for Nature must have $\mu_{01} = 1 - \mu_{11}$, in which case the maximization problem can be further simplified to

$$\max_{\Delta\in[-1,1]} \max \left\{ \begin{array}{c} \underbrace{\frac{\exp(b)(1+\Delta)-(1-\Delta)}{\exp(b)(1+\Delta)+(1-\Delta)}\left(1 - \alpha^* f_1(\Delta) - (1-\alpha^*)f_4(\Delta)\right),}_{\psi_1(\Delta)} \\ \underbrace{\frac{1+\Delta-\exp(b)(1-\Delta)}{1+\Delta+\exp(b)(1-\Delta)}\left(\alpha^* f_1(\Delta) + (1-\alpha^*)f_4(\Delta)\right)}_{\psi_2(\Delta)} \end{array} \right\},$$

where $f_1(\Delta) = \mathbb{E}\delta_1(\omega)$ as before and $f_4(\Delta) = \mathbb{E}\delta_4(\omega)$. The arguments of the inner max-operator are symmetric: $\psi_1(\Delta) = -\psi_2(-\Delta)$. It follows that any maximizer of $\psi_1(\Delta)$ is also a maximizer of $\max\{\psi_1(\Delta), \psi_2(\Delta)\}$. For the purpose of constructing a simplified game, I therefore drop $\psi_2(\Delta)$ from the best response problem, leading to

$$\max_{\Delta\in[-1,1]} \phi(\Delta;\alpha,b), \phi(\Delta;\alpha,b) = \underbrace{\frac{\exp(b)(1+\Delta)-(1-\Delta)}{\exp(b)(1+\Delta)+(1-\Delta)}}_{\rho(\Delta)}\left(1 - \alpha^* f_1(\Delta) - (1-\alpha^*)f_4(\Delta)\right), \qquad (4)$$

the problem that will be analyzed henceforth.

**Step 2: Introducing a simplified game.** As before.

**Step 3: Equilibrium when $b \geq \ln\left(2+\sqrt{5}\right)$.** The claim is that $\alpha^* = 0$ in this case. To verify it, note that $f_4(\Delta) = \frac{1}{2} + \frac{\Delta}{\exp(b)-1}$ in this case, thus (4) simplifies to

$$\max_{\Delta \in [-1,1]} \rho(\Delta)\left(\frac{1}{2} - \frac{\Delta}{\exp(b)-1}\right).$$

By evaluating derivatives, this can be verified to be solved by $\Delta^* = 0$ as required.

**Step 4: Equilibrium when $\alpha^* = 1$.** Assume for the moment that $\phi$ is quasiconcave, then the equilibrium obtains whenever $\phi'(0) \geq 0$. Algebraic evaluation reveals that this condition is equivalent to $\exp(b) \leq 1/f_1'(0) + \sqrt{1 + (1/f_1'(0))^2}$.

To see quasiconcavity, observe first that $\rho$ is increasing and concave, whereas algebra from the preceding proof revealed that $(1 - f_1(\Delta))$ is decreasing and concave, over $\Delta \in [-1, 1]$. These facts together imply that $\phi$ is concave on $[-1, 0]$. It is also easily verified that $\phi(0; \alpha, b) = \frac{\exp(b)-1}{2(\exp(b)+1)} > 0$ and that $\phi(1; \alpha, b) = 0$. It follows that $\phi(\Delta; \alpha, b)$ has a critical point $\widetilde{\Delta}$ on $[0, 1]$. Let $k = \rho(\widetilde{\Delta})$ and $l = \rho'(\widetilde{\Delta})$, then $\widetilde{\rho}(\Delta) = \rho(\widetilde{\Delta}) + \rho'(\widetilde{\Delta})(\Delta - \widetilde{\Delta})$ is the tangent to $\rho$ at $\widetilde{\Delta}$. As $\rho$ is concave with positive intercept, $\widetilde{\rho}$ is a positive affine function of $\Delta$, thus $\phi(\Delta) \leq \widetilde{\rho}(\Delta)(1 - f_1(\Delta))$. But with $\widetilde{\rho}$ being positive affine, algebra from step 4 of the preceding proof can be entirely mimicked to show that $\widetilde{\rho} \cdot (1 - f_1)$ as a function of $\Delta$ is quasiconcave on $[0, 1]$, is positive and increasing at 0, and equals 0 at 1. It follows that $\widetilde{\rho} \cdot (1 - f_1)$ has at most one maximum on $[0, 1]$, and any such maximum obtains iff a first-order condition holds. But $\frac{d}{d\Delta}\widetilde{\rho}(\widetilde{\Delta})(1 - f_1(\widetilde{\Delta})) = \widetilde{\rho}(\widetilde{\Delta})(-f_1'(\widetilde{\Delta})) + \widetilde{\rho}'(\widetilde{\Delta})(1 - f_1(\widetilde{\Delta})) = \rho(\widetilde{\Delta})(-f_1'(\widetilde{\Delta})) + \rho'(\widetilde{\Delta})(1 - f_1(\widetilde{\Delta})) = 0$, so the maximum occurs at $\widetilde{\Delta}$. Since also $\widetilde{\rho}(\widetilde{\Delta})(1 - f_1(\widetilde{\Delta})) = \rho(\widetilde{\Delta})(1 - f_1(\widetilde{\Delta}))$, it follows that $\phi(0; \alpha, b)$ is uniquely maximized at $\widetilde{\Delta}$.[17]

**Step 5: Analysis of revealing equilibria in general.** As before.

**Step 6: Characterization of the pooling equilibrium.** Algebra similar to before reveals that $\phi'(0; \alpha^*, b) = 0$ requires

$$\alpha^* = \frac{\frac{4\exp(b)}{\exp(2b)-1} - f_4'(0)}{f_1'(0) - f_4'(0)}.$$

$f_1'(0)$ was computed before, and evaluation of $f_4'(0)$ is similar to the one of $f_3'(0)$.

[17] This footnote is for the referee's consumption. The intuition behind the linearization trick is as follows. Concavity of $\rho$ means that examination of second derivatives as in the preceding proof becomes intractable. But substantively, $\rho$ being strictly concave rather than linear makes Nature's objective "more concave" and therefore should not be a problem. This is reflected in the fact that linearization of $\rho$ around the FOC delivers the result.

**Proposition 6** As before, it is easy to see that regret depends on states of the world only through $(\mu_{10}, \mu_{11}) = (\mathbb{E}(Y_1|Z=0), \mathbb{E}(Y_1|Z=1))$, thus identify states $s$ with vectors $(\mu_{01}, \mu_{11}) \in [0,1]^2$. For $\varepsilon$ small enough, an equilibrium of the fictitious game is given by $(\delta_2^*, \widetilde{\pi}_2^*)$, where $\widetilde{\pi}_2^*$ randomizes over some $\{(a^*, 1), (b^*, 0)\}$ s.t. $(1-\varepsilon)a^* + \varepsilon \geq \mu_0 \geq (1-\varepsilon)b^*$ and $a^* \geq b^*$. The randomization is not in general even; the equilibrium value for $p \equiv \widetilde{\pi}_2^*(a^*, 1)$ will be characterized below.

To verify that (the modification of) $\delta_2^*$ is a best response, note that a Bayesian with prior $\widetilde{\pi}_2^*$ will weakly prefer treatment 1 whenever

$$\mu_0 \leq \mathbb{E}(Y_1|\omega) = \frac{p\left((1-\varepsilon)a^* + \varepsilon\right)\Pr(\omega|s=(a^*,1)) + (1-p)(1-\varepsilon)b^*\Pr(\omega|s=(b^*,0))}{p\Pr(\omega|s=(a^*,1)) + (1-p)\Pr(\omega|s=(b^*,0))},$$

and will weakly prefer treatment 0 if the opposite inequality holds. The r.h. expression increases in $\overline{y}_1$, hence the decision maker will prefer treatment 0 whenever $\overline{y}_1$ is below some threshold value and will prefer treatment 1 whenever $\overline{y}_1$ is above it. To support $\delta_1^*$, $p$ must be chosen s.t. this threshold value is $n^*/N$. For any $\overline{y}_1$, the r.h. expression increases from $(1-\varepsilon)b^*$ to $((1-\varepsilon)a^* + \varepsilon)$ as $p$ decreases from 1 to 0, so $p$ can always be appropriately.

Next, $(n^*, \lambda^*)$ must be chosen s.t. both of $\{(a^*, 1), (b^*, 0)\}$ are best responses. This requires that

$$\begin{aligned} & \max_{a\in[0,1]} \left\{((1-\varepsilon)a + \varepsilon - \mu_0)\left(\Pr(N\overline{y}_1 < n^*|\mu_{11}=a) + (1-\lambda^*)\Pr(N\overline{y}_1 = n^*|\mu_{11}=a)\right)\right\} \\ = & \max_{b\in[0,1]} \left\{(\mu_0 - (1-\varepsilon)b)\left(\Pr(N\overline{y}_1 > n^*|\mu_{11}=b) + \lambda^*\Pr(N\overline{y}_1 = n^*|\mu_{11}=b)\right)\right\} \end{aligned} \tag{5}$$

and furthermore that

$$\begin{aligned} & \max\left\{\arg\max_{a\in[0,1]} \left\{((1-\varepsilon)a + \varepsilon - \mu_0)\left(\Pr(N\overline{y}_1 < n^*|\mu_{11}=a) + (1-\lambda^*)\Pr(N\overline{y}_1 = n^*|\mu_{11}=a)\right)\right\}\right\} \\ \geq & \min\left\{\arg\max_{b\in[0,1]} \left\{(\mu_0 - (1-\varepsilon)b)\left(\Pr(N\overline{y}_1 > n^*|\mu_{11}=b) + \lambda^*\Pr(N\overline{y}_1 = n^*|\mu_{11}=b)\right)\right\}\right\}, \end{aligned} \tag{6}$$

so that $(a^*, b^*)$ can be picked s.t. $a^* \geq b^*$.

First, $(n^*, \lambda^*)$ can always be picked s.t. (5) obtains. To see this, note that the conservativeness of $\delta_2^*$ is $(n^*, \lambda^*)$ can be smoothly indexed by $\alpha \equiv n + 1 - \lambda$; $\alpha = 0$ $[N+1]$ corresponds to the rule that always [never] picks treatment 1. The l.h. objective pointwise continuously decreases in $\alpha$, hence so does its maximum; the r.h. maximum continuously increases in $\alpha$. If $\alpha = 0$, then the l.h. maximum is $(1-\mu_0)$ and the r.h. maximum is 0; if $\alpha = 1$, then the l.h. maximum is 0 and the r.h. maximum is $\mu_0$. It follows that equality obtains at an intermediate value of $\alpha$.

Second, condition (6) applies for some positive (but small enough) $\varepsilon$. To see that, note that if $\varepsilon = 0$, the equilibrium is the one from proposition 1(iii); in particular, $a^* > \mu_0 > b^*$ obtains. Let $(n_0^*, \lambda_0^*)$ denote the values of $(n^*, \lambda^*)$ in that equilibrium and let $(a_0^*, b_0^*)$ be according values of $(a^*, b^*)$, then one has

$$\begin{aligned}
&\max_{a\in[0,1]} \{((1-\varepsilon)a+\varepsilon-\mu_0)\left(\Pr(N\overline{y}_1 < n^*|\mu_{11}=a)+(1-\lambda^*)\Pr(N\overline{y}_1 = n^*|\mu_{11}=a)\right)\} \\
\geq\ & ((1-\varepsilon)a_0^*+\varepsilon-\mu_0)\left(\Pr(N\overline{y}_1 < n_0^*|\mu_{11}=a_0^*)+(1-\lambda_0^*)\Pr(N\overline{y}_1 = n_0^*|\mu_{11}=a_0^*)\right) \\
=\ & \frac{((1-\varepsilon)a_0^*+\varepsilon-\mu_0)}{a_0^*-\mu_0} R_2^*(N) > R_2^*(N),
\end{aligned}$$

but one also has the crude bound

$$\begin{aligned}
&\max_{a\in[0,\mu_0]} \{((1-\varepsilon)a+\varepsilon-\mu_0)\left(\Pr(N\overline{y}_1 < n^*|\mu_{11}=a)+(1-\lambda^*)\Pr(N\overline{y}_1 = n^*|\mu_{11}=a)\right)\} \\
&\qquad\qquad\qquad \leq ((1-\varepsilon)\mu_0+\varepsilon-\mu_0) = \varepsilon(1-\mu_0).
\end{aligned}$$

For $\varepsilon$ small enough, one has $\varepsilon(1-\mu_0) \leq R_2^*(N)$ and therefore

$$\max\left\{\arg\max_{a\in[0,1]} \{((1-\varepsilon)a+\varepsilon-\mu_0)\left(\Pr(N\overline{y}_1 < n^*|\mu_{11}=a)+(1-\lambda^*)\Pr(N\overline{y}_1 = n^*|\mu_{11}=a)\right)\}\right\} \geq \mu_0$$

(elementwise if the arg max is a set). Similarly, one finds that for $\varepsilon$ small enough,

$$\min\left\{\arg\max_{b\in[0,1]} \{(\mu_0-(1-\varepsilon)b)\left(\Pr(N\overline{y}_1 > n^*|\mu_{11}=b)+\lambda^*\Pr(N\overline{y}_1 = n^*|\mu_{11}=b)\right)\}\right\} \leq \mu_0.$$

**Proposition 7** The proof plan is the same as for proposition 6, with some different algebraic detail that much resembles the proof of proposition 5.

**Proposition 2**

**(i)** I establish the claim for unconstrained randomization with $N$ odd, the extension to other cases is as above. The claim is that if $R_1^*(N) \geq 2\kappa$, then an equilibrium of the fictitious game is given by $(\overline{\delta}^*, \overline{\pi}_1^*)$, where $\overline{\pi}_1^*$ is a prior that has $s_x = s_{x'}$ a.s. for all $x, x'$ and that otherwise mimics $\pi_1^*$. In words, the equilibrium mimics the equilibrium from proposition 1, with both players ignoring the existence of $X$.

As the distribution of $X$ is known, it is without loss of generality to let it be uniform on $[0,1]$. Noting that regret depends on $s_x$ only through the marginal distributions of $(Y_{0x}, Y_{1x})$, identify states $s$ with $\Sigma$-measurable functions that map covariate values $x \in X$ onto $(\mu_{0x}, \mu_{1x}) = \mathbb{E}(Y_{0x}, Y_{1x}) \in [0,1]^2$. Also define the unconditional expectations $\mu_t = \int \mu_{tx} dx$.[18] Then Nature's best-response problem

[18] The expectations in this sentence are features of $s$; to avoid confusion between parameter values and sampling expectations, I will otherwise reserve the symbol $\mathbb{E}$ for the latter and write features of $s$ as integrals.

against decision rule $\overline{\delta}^*$ is given by

$$\begin{aligned}
& \max_{s\in\mathcal{S}} R(\overline{\delta}^*, s) \\
= & \max_{s\in\mathcal{S}} \int \Big(\max\{\mu_{0x},\mu_{1x}\} - \mu_{0x}(1-\mathbb{E}\overline{\delta}^*_x(\omega)) - \mu_{1x}\mathbb{E}\overline{\delta}^*_x(\omega)\Big)\, dx \\
= & \max_{s\in\mathcal{S}} \int \Big(\mathbb{E}\overline{\delta}^*(\omega)\cdot(\mu_{0x}-\mu_{1x})\cdot 1\{\mu_{0x}>\mu_{1x}\} + (1-\mathbb{E}\overline{\delta}^*(\omega)\cdot(\mu_{1x}-\mu_{0x})\cdot 1\{\mu_{1x}\geq\mu_{0x}\}\Big)\, dx \\
= & \max_{s\in\mathcal{S}} \left\{\mathbb{E}\overline{\delta}^*(\omega)\int(\mu_{0x}-\mu_{1x})\cdot 1\{\mu_{0x}>\mu_{1x}\}dx + (1-\mathbb{E}\overline{\delta}^*(\omega))\int(\mu_{1x}-\mu_{0x})\cdot 1\{\mu_{1x}\geq\mu_{0x}\}dx\right\} \\
= & \max_{s\in\mathcal{S}} \left\{\mathbb{E}\overline{\delta}^*(\omega)\, pa + (1-\mathbb{E}\overline{\delta}^*(\omega))(1-p)b\right\}, \qquad (7)
\end{aligned}$$

where the first step substitutes in for the definition of $R(\overline{\delta}^*, s)$, the next two steps rearrange terms and use that $\overline{\delta}^*_x(\omega) = \overline{\delta}^*(\omega)$ for all $x$, and the final step defines

$$\begin{aligned}
a &= \frac{\int(\mu_{0x}-\mu_{1x})1\{\mu_{0x}>\mu_{1x}\}dx}{\int 1\{\mu_{0x}>\mu_{1x}\}dx} \\
b &= \frac{\int(\mu_{1x}-\mu_{0x})1\{\mu_{0x}\leq\mu_{1x}\}dx}{\int 1\{\mu_{0x}\leq\mu_{1x}\}dx} \\
p &= \int 1\{\mu_{0x}>\mu_{1x}\}dx.
\end{aligned}$$

If there exists a solution to this problem s.t. $p \in \{0,1\}$, then there exists a solution s.t. $(\mu_{0x}, \mu_{1x})$ is constant. To see this, note that if $p = 1$, then $R(\overline{\delta}^*, s)$ depends on $s$ only through $\mathbb{E}\overline{\delta}^*(\omega)$ and $a$; but $\mathbb{E}\overline{\delta}^*(\omega)$ depends on $s$ only through $(\mu_0, \mu_1)$, and $a$ simplifies to $(\mu_0 - \mu_1)$ if $p = 1$. The argument for $p = 0$ is similar. In particular, if any solution to the problem is s.t. $p \in \{0,1\}$, then $\overline{\pi}_1^*$ is a best response and the value of the best-response problem is $R_1^*(N)$.

The remainder of this proof establishes that if (7) has an interior (in the sense of $p \in (0,1)$) solution, then its value can be bounded above by $\kappa/2$. If $\kappa \leq 2R_1^*(N)$, then it follows that $\overline{\pi}_1^*$ is a best response to $\overline{\delta}^*$. That $\overline{\delta}^*$ is a best response to $\overline{\pi}_1^*$ is known, so that a Nash equilibrium has been established. I will finally show that the bound of $\kappa/2$ on the value of interior solutions cannot be improved upon without further assumptions.

If $p \in (0,1)$, then there exist covariate values $x, x' \in \mathcal{X}$ s.t. $(\mu_{1x} - \mu_{0x})(\mu_{1x'} - \mu_{0x'}) < 0$, i.e. ex post optimal treatment is different for these two values. One can then write

$$\mu_{1x} - \mu_{0x} + \mu_{0x'} - \mu_{1x'} = \underbrace{\mu_{1x} - \mu_{1x'}}_{\leq\kappa} + \underbrace{\mu_{0x'} - \mu_{0x}}_{\leq\kappa} \leq 2\kappa.$$

This holds for any $x, x'$ with said property, thus

$$2\kappa \geq \sup_{x\in\mathcal{X}}\{\mu_{0x} - \mu_{1x}\} + \sup_{x\in\mathcal{X}}\{\mu_{1x} - \mu_{0x}\} \geq a + b.$$

The next step is to replace $\mathbb{E}\overline{\delta}^*(\omega)$ with a tractable expression. Observe that $\mathbb{E}\overline{\delta}^*(\omega) \geq 1/2$ iff $\mu_1 \geq \mu_0$ iff $b(1-p) \geq pa$. Also, the argument in (7) increases in $\mathbb{E}\overline{\delta}^*(\omega)$ iff $\mu_1 \leq \mu_0$ iff $b(1-p) \leq pa$.

It follows that the argument increases in $\mathbb{E}\overline{\delta}^*(\omega)$ iff $\mathbb{E}\overline{\delta}^*(\omega) \leq 1/2$. Replacing $\mathbb{E}\overline{\delta}^*(\omega)$ with a quantity that is known to be between $\mathbb{E}\overline{\delta}^*(\omega)$ and $1/2$ will, therefore, increase the value of problem (7).

Such a quantity is given by $(1+b(1-p)-pa)/2$. To see this, observe that under unconstrained randomization and with $N$ odd, $\delta^*$ can be rewritten as $\delta^*(\omega) = 1\{(N+I_N)/2 > 1/2\}$, where $(N+I_N)/2 = \#\{\text{observed successes of treatment 1}\} + \{\#\text{observed failures of treatment 0}\}$ is distributed binomially with parameters $(N, (\mu_1 + 1 - \mu_0)/2) = (N, (1+b(1-p)-pa)/2)$. Denoting by $B(\cdot; N, \pi)$ the binomial c.d.f. with indicated parameters and using the regularized Beta distribution representation of this c.d.f., one thus has

$$\begin{aligned}
& \mathbb{E}\overline{\delta}^*(\omega) \\
= \; & 1 - B((N-1)/2; N, (1+b(1-p)-pa)/2) \\
= \; & 1 - \frac{N+1}{2}\binom{N}{\frac{N-1}{2}} \int_0^{(1+b(1-p)-pa)/2} L^{(N-1)/2}(1-L)^{(N-1)/2} dL.
\end{aligned}$$

Direct evaluation of derivatives reveals that $\mathbb{E}\overline{\delta}^*(\omega)$ as a function of $(1+b(1-p)-pa)/2$ is convex if $(1+b(1-p)-pa)/2 \leq 1/2 \Leftrightarrow \mu_1 \leq \mu_0$ and concave otherwise. Also, it is easy to verify that $\mathbb{E}\overline{\delta}^*(\omega) = (1+b(1-p)-pa)/2$ for $(1+b(1-p)-pa)/2 \in \{0, 1/2, 1\}$. It follows that on $[0,1]$, the graph of $\mathbb{E}\delta^*(\omega)$ is first below the one of $(1+b(1-p)-pa)/2$, equals it at $(1/2, 1/2)$, and is above it thereafter. This implies the claim.

One therefore has

$$\begin{aligned}
& \sup_{s \in \mathcal{S}: p \in (0,1)} \left\{ \mathbb{E}\overline{\delta}^*(\omega) \cdot ap + \left(1 - \mathbb{E}\overline{\delta}^*(\omega)\right) \cdot b(1-p) \right\} \\
\leq \; & \sup_{s \in \mathcal{S}: p \in (0,1)} \left\{ \frac{1+b(1-p)-pa}{2} ap + \frac{1-b(1-p)+pa}{2} b(1-p) \right\} \\
= \; & \sup_{s \in \mathcal{S}: p \in (0,1)} \left\{ \frac{1}{2}\left(pa + (1-p)b - (pa-(1-p)b)^2\right) \right\}.
\end{aligned}$$

All in all, $\sup_{s \in \mathcal{S}: p \in (0,1)} R(\overline{\delta}^*, s)$ can be bounded from above by

$$\max_{\substack{(a,b,p) \in [0,2\kappa] \times [0,2\kappa] \times [0,1]: \\ a+b \leq 2\kappa}} \frac{1}{2}\left(pa + (1-p)b - (pa-(1-p)b)^2\right), \tag{8}$$

and this supremum can be attained on $\{s \in \mathcal{S} : p \in (0,1)\}$ only if the maximum in in (8) is attained by some $p \in (0,1)$. It will now be shown that (8) is bounded from above by $\max\{R_1^*(N), \kappa/2\}$, where furthermore the bound equals $\kappa/2$ whenever it is achieved by a $p \in (0,1)$.[19]

[19] This footnote is for the referee's consumption. The remainder of the proof mechanically if tediously verifies the very last statement. The money was in two preceding steps: First, use the limited selectivity assumption to conclude $a + b \leq 2\kappa$. This quite transparently leads to a bound on regret of $\kappa$ – after all, DM could just determine treatment by coin-flips independent of the data and get it right 1/2 of the time. The additional improvement in the bound, which also

Consider first optimization with respect to $b$, keeping $(a,p)$ fixed:

$$\max_{b\in[0,2\kappa-a]}\frac{1}{2}\left(pa+(1-p)b-(pa-(1-p)b)^2\right). \tag{9}$$

Evaluation of derivatives reveals that the objective is concave, thus the solution is either a corner solution or an interior one characterized by first-order condition

$$(1-p)(1+2(pa-(1-p)b)\overset{!}{=}0\Longrightarrow b=\frac{ap+1/2}{1-p},$$

which (for future reference) is consistent with $b\leq 2\kappa-a$ iff

$$\frac{ap+1/2}{1-p}\leq 2\kappa-a\Longleftrightarrow a\leq 2\kappa-2\kappa p-\frac{1}{2}.$$

If $(a^*,b^*,p^*)$ solves (8), then $b^*$ must solve (9) with parameters $(a^*,p^*)$; thus

$$b^*\in\left\{0,\left(a^*p^*+1/2\right)/\left(1-p^*\right),2\kappa-a^*\right\},$$

where the middle value is feasible only if $a^*\leq 2\kappa-2\kappa p^*-1/2$. These cases will be analyzed separately.

First, assume $b^*=0$, then (8) simplifies to

$$\max_{(a,p)\in[0,2\kappa]\times[0,1]}\frac{pa-(pa)^2}{2}. \tag{10}$$

It is then w.l.o.g. to set $p^*=1$, the remaining free variable being $a=\int\left(\mu_{1x}-\mu_{0x}\right)dx$ (the simplification uses $p^*=1$). This is a boundary solution, and its value can be achieved by priors that set $\mu_{0x}$ and $\mu_{1x}$ constant; in particular, its value is bounded by $R_1^*(N)$.

Next, assume $b^*=\left(a^*p^*+1/2\right)/(1-p^*)$. Then problem (8) is unaffected by adding the constraint that $b=\left(ap+1/2\right)/(1-p)$. Note, however, from above that this case is feasible only if $a^*\leq 2\kappa-2\kappa p^*-1/2$. Substituting for these observations leads to

$$\max_{\substack{(a,p)\in[0,2\kappa]\times[0,1]:\\ a+b\leq s\kappa}}\frac{1}{2}\left(pa+\left(ap+\frac{1}{2}\right)-\left(pa-\left(ap+\frac{1}{2}\right)\right)^2\right)=\max_{\substack{(a,p)\in[0,2\kappa]\times[0,1]:\\ a\leq 2\kappa-2\kappa p-1/2}}\left\{pa+\frac{1}{8}\right\}.$$

This objective increases in $a$ for every $p$, so the constraint will bind and can be substituted for, leading to

$$\max_{p\in[0,1]}\left\{p\left(2\kappa-2\kappa p-\frac{1}{2}\right)+\frac{1}{8}\right\}.$$

leads to a tight bound, comes because the DM can use the signal more smartly – if there is a signal, that is. Recalling that $pa-(1-p)b=\mu_1-\mu_0$, this "more smartly" shows up in the "$-(pa-(1-p)b)^2$" term in (8), which punishes Nature for making $\mu_1$ and $\mu_0$ different and the better treatment thereby detectable. Without this term, the bound would indeed be the aforementioned bound of $\kappa$, achieved by $(a,b,p)=(2\kappa,0,1)$. The linearization of $\mathbb{E}\overline{\delta}^*$ actually has an intuition: An exact calculation for large finite $N$ seeming infeasible, I just replaced $\mathbb{E}\overline{\delta}^*$ with what it would be if $N=1$ (and for $N=1$, all subsequent calculations are exact). This should be expected to shrink $\mathbb{E}\overline{\delta}^*$ toward 1/2 because increasing $N$ surely improves detection of the better treatment.

The objective is strictly concave in $p$. Algebraic evaluation yields $p^* = \max\{1/2 - 1/8\kappa, 0\}$, where $p^* > 0$ iff $\kappa > 1/4$. Thus for $\kappa < 1/4$, we have $p^* = 0$, i.e. a corner solution (that furthermore coincides (10)). If $\kappa \geq 1/4$, then the constrained solution can be evaluated. It is characterized by $p^* = 1/2 - 1/8\kappa$, which can be substituted into the constraint to find $a^* = \kappa - 1/4$, leading to

$$\max_{\substack{(a,p)\in[0,2\kappa]\times[0,1]:\\ a\leq 2\kappa-2\kappa p-1/2}} \left\{pa + \frac{1}{8}\right\} = \left(\frac{1}{2} - \frac{1}{8\kappa}\right)\left(\kappa - \frac{1}{4}\right) + \frac{1}{8} = \frac{\kappa}{2} + \frac{1}{32\kappa} - \frac{1}{8} \leq \frac{\kappa}{2},$$

where the last step used that $\kappa \geq 1/4$.

Finally, assume $b^* = 2\kappa - a^*$, then substituting in simplifies (8) to

$$\max_{(a,p)\in[0,2\kappa]\times[0,1]} = \frac{1}{2}\left(pa + (1-p)(2\kappa - a) - (pa - (1-p)(2\kappa - a))^2\right).$$

Assuming any boundary solution, i.e. $a^* \in \{0, 2\kappa\}$ or $p \in \{0,1\}$, would again collapse the problem into (10). It remains to evaluate a possible interior maximum. Evaluation of first order conditions

$$2p^* - 1 + 4\kappa - 4\kappa p^* - 2a^* \overset{!}{=} 0$$
$$2a^* - 2\kappa + 8\kappa^2 - 8\kappa^2 p^* - 4\kappa a^* \overset{!}{=} 0$$

leads to the unique candidate solution

$$(a^*, p^*) = (\kappa, 1/2).$$

This is the only interior candidate solution to (8), and it has value $\kappa/2$.

Finally, if $X$ is continuous, then the bound can be achieved, thus the proposition fails for $\kappa > 2R_1^*(N)$. If $X$ is continuous, it is w.l.o.g. to assume that it is continuously distributed on $[0,1]$. Mimicking Stoye (2009a, proposition 4), construct a sequence of priors $\pi_i, i = 1, 2, \ldots$ as follows. Define the partition $W_i \equiv \{[0, 1/i], (1/i, 2/i], \ldots ((i-1)/i, 1]\}$ of the unit interval. Let $(w_i^j)_{j=1}^{2^i}$ collect the subsets of $W_i$ in arbitrary order. Define the collection of distributions $\left(s_i^j\right)_{j=1}^{2^i}$ by identifying $s_i^j$ with the degenerate distribution concentrated at

$$(\mu_{0x}, \mu_{1x})_{x\in\mathcal{X}} = \begin{pmatrix} (1+\kappa/2)\,1\left\{x \in w_i^j\right\} + (1-\kappa/2)\,1\left\{x \in w_i^j\right\}, \\ (1-\kappa/2)\,1\left\{x \in w_i^j\right\} + (1+\kappa/2)\,1\left\{x \in w_i^j\right\} \end{pmatrix}_{x\in\mathcal{X}}.$$

Let $\pi_i$ be the uniform distribution over states $\left(s_i^j\right)_{j=1}^{2^i}$, i.e. $\pi_i$ assigns probability $2^{-i}$ to every $s_i^j$. Notice the following features of $\pi_i$: (i) The prior expectation of $(Y_{0x}, Y_{1x})$ equals $(1/2, 1/2)$. (ii) With slight abuse of notation, let $w_i(x)$ be the element of $W_i$ that contains $x$. Then $s_x$ and $s_{x'}$ are independent whenever $w_i(x) \neq w_i(x')$. Now minimal adaptation of algebra from Stoye (2009a, proposition 4) shows that $\lim_{i\to\infty} \min_{\delta\in D} \int R(\delta, s) d\pi_i = \kappa/2$, thus $\lim_{i\to\infty} \min_{\delta\in D} \max_{s\in\mathcal{S}} R(\delta, s) \geq \kappa/2$. (This is the construction that established the "covariate problem" to begin with, just adapted to insure $|\mu_{tx} - \mu_{tx'}| \leq \kappa$.)

**(ii)** Consider the extended game in which the decision maker may also choose the within-sample treatment assignment rule. If Nature plays $\overline{\pi}_1^*$, then it follows from the preceding paragraph's observations that $\overline{\delta}_1^*$ in conjunction with any within-sample treatment assignment rule is a best response. If DM chooses any of the treatment assignment rules under consideration, then Nature's best response problem is just the one from part (i), hence $\overline{\pi}_1^*$ is a best response, establishing equilibrium.

**(iii)** Consider the extended game in which the decision maker may also choose a sample design with respect to $X$. This game has an equilibrium in which the decision maker picks the simple random sample and equilibrium strategies otherwise equal the ones from part (i). If the decision maker's strategy is as conjectured, then Nature's best response problem is just the one from part (i), hence $\overline{\pi}_1^*$ is a best response. But since $X$ is irrelevant given $\overline{\pi}_1^*$, one best response to $\overline{\pi}_1^*$ is decision rule $\overline{\delta}_1^*$ in conjunction with simple random sampling. This establishes the equilibrium.

**(iv)** The argument is similar to the one in (i), and I will merely point out the differences.

If $p \in (0,1)$, then there exist covariate values $x, x' \in \mathcal{X}$ s.t. $(\mu_{1x} - \mu_{0x})(\mu_{1x'} - \mu_{0x'}) < 0$, i.e. ex post optimal treatment is different for these two values. One can then write

$$\mu_{1x} - \mu_{0x} + \mu_{0x'} - \mu_{1x'} = \underbrace{\mu_{1x} - \mu_{1x'}}_{\leq \kappa} + \underbrace{\mu_{0x'} - \mu_{0x}}_{=0} \leq \kappa.$$

This holds for any $x, x'$ with said property, thus

$$\kappa \geq \sup_{x \in \mathcal{X}} \{\mu_{0x} - \mu_{1x}\} + \sup_{x \in \mathcal{X}} \{\mu_{1x} - \mu_{0x}\} \geq a + b.$$

This improves the bound from the proof of (i) by a factor of 2, which is reflected in result.

Two modifications are needed to argue that the substitution of $(1 + b(1-p) - pa)/2$ for $\mathbb{E}\delta^*(\omega)$ shrinks the latter toward $1/2$. First, $\delta^*$ is not the decision rule discovered in proposition 1, it is rather identified with the cutoff rule

$$\widetilde{\delta}_x(\omega) \equiv \begin{cases} 0, & N\overline{y}_1 < \widetilde{n} \\ \widetilde{\lambda}, & N\overline{y}_1 = \widetilde{n} \\ 1, & N\overline{y}_1 > \widetilde{n} \end{cases}.$$

where $(\widetilde{n}, \widetilde{\lambda})$ are s.t. $\mathbb{E}\widetilde{\delta}(\omega) = 1/2$ iff $\mu_1 = \mu_0$. By bounding maximal regret incurred by this decision rule, one a forteriori bounds minimax regret.

One can then bound

$$\sup_{s \in \mathcal{S}: p \in (0,1)} \left\{\mathbb{E}\widetilde{\delta}(\omega) \cdot pa + \left(1 - \mathbb{E}\widetilde{\delta}(\omega)\right) \cdot b(1-p)\right\},$$

where $(a, b, p)$ are as in part (i).

Let $\mu_0 \in (0,1)$. (The decision problem is trivial if $\mu_0 \in \{0,1\}$.) Then $\mathbb{E}\widetilde{\delta}(\omega)$ as a function of $\mu_1$ passes through the following points: $\{(0,0),(\mu_0,1/2),(1,1)\}$. Furthermore, write

$$\begin{aligned}
\mathbb{E}\widetilde{\delta}(\omega) &= 1-\left(1-\widetilde{\lambda}\right)B(\widetilde{n}-1;N,\mu_1)-\widetilde{\lambda}B(\widetilde{n};N,\mu_1)\\
&= 1-\left(1-\widetilde{\lambda}\right)(N-\widetilde{n}+1)\binom{N}{\widetilde{n}-1}\int_0^{\mu_1} t^{N-\widetilde{n}}(1-t)^{\widetilde{n}+1}dt-\widetilde{\lambda}\,(N-\widetilde{n})\binom{N}{\widetilde{n}}\int_0^{\mu_1} t^{N-\widetilde{n}-1}(1-t)^{\widetilde{n}}dt.
\end{aligned}$$

Some algebra yields

$$\frac{\partial^2\mathbb{E}\widetilde{\delta}(\omega)}{\partial\mu_1^2}=\binom{N}{\widetilde{n}}(1-\mu_1)^{N-\widetilde{n}-2}\mu_1^{\widetilde{n}-2}\cdot$$

$$\left[\underbrace{\left(1-\widetilde{\lambda}\right)\widetilde{n}[(\widetilde{n}-1)(1-\mu_1)^2-(N-\widetilde{n})\mu_1(1-\mu_1)]+\widetilde{\lambda}(N-\widetilde{n})[\widetilde{n}\mu_1(1-\mu_1)-(N-\widetilde{n}-1)\mu_1^2]}_{\psi}\right]$$

This expression is zero at $\{0,1\}$ and its sign otherwise equals the sign of $\psi$. $\psi$ is strictly positive at $\mu_1=0$ and strictly negative at $\mu_1=1$, thus it has an odd number of sign changes on $[0,1]$. But it is also quadratic in $\mu_1$ and therefore has at most two sign changes on the real line. It follows that $\mathbb{E}\widetilde{\delta}(\omega)$ is first convex and then concave on $[0,1]$. Assuming w.l.o.g. that $\mu_0\geq 1/2$, compare this to

$$\frac{\mu_1}{2\mu_0}=\frac{1}{2}\frac{\mu_0+(\mu_1-\mu_0)}{\mu_0}=\frac{1}{2}\left(1+\frac{(1-p)b-pa}{\mu_0}\right),$$

where $(a,b,p)$ are as in part (i). This linear function passes through $\{(0,0),(\mu_0,1/2),(1,1/(2\mu_0))\}$; note that $1/(2\mu_0)\leq 1$. It is now seen that replacing $\mathbb{E}\delta^*(\omega)$ with $(1/2(1+((1-p)b-pa)/\mu_0))$ in the maximization problem will bound its value from above.

It follows that

$$\begin{aligned}
&\sup_{s\in\mathcal{S}:p\in(0,1)}\left\{\mathbb{E}\widetilde{\delta}(\omega)\cdot pa+\left(1-\mathbb{E}\widetilde{\delta}(\omega)\right)\cdot b(1-p)\right\}\\
\leq\ &\sup_{s\in\mathcal{S}:p\in(0,1)}\left\{\frac{1}{2}\left(pa\left(1+\frac{(1-p)b-pa}{\mu_0}\right)+(1-p)b\left(1-\frac{(1-p)b-pa}{\mu_0}\right)\right)\right\}\\
=\ &\sup_{s\in\mathcal{S}:p\in(0,1)}\left\{\frac{1}{2}\left(pa+(1-p)b-\frac{(pa-(1-p)b)^2}{\mu_0}\right)\right\}\\
\leq\ &\sup_{s\in\mathcal{S}:p\in(0,1)}\left\{\frac{1}{2}\left(pa+(1-p)b-(pa-(1-p)b)^2\right)\right\}.
\end{aligned}$$

From here on, the argument proceeds as before, and the value of an interior solution to the problem can be bounded at $\kappa/4$ (the improvement relative to $\kappa/2$ stems from the tightened constraint on $a+b$).

To see that Nature can achieve the bound, let prior $\pi_i$ be as in (i) except that

$$(\mu_{1x})_{x\in\mathcal{X}}=\left((\mu_0+\kappa/2)\,1\left\{x\in w_i^j\right\}+(\mu_0-\kappa/2)\,1\left\{x\in w_i^j\right\}\right)_{x\in\mathcal{X}},$$

then $\lim_{i\to\infty}\min_{\delta\in D}\int R(\delta,s)d\pi_i = \kappa/4$, thus $\lim_{i\to\infty}\min_{\delta\in D}\max_{s\in\mathcal{S}} R(\delta,s) \geq \kappa/4$. An additional subtlety is that this construction is only feasible if $\{\mu_0-\kappa/2, \mu_0+\kappa/2\} \subseteq [0,1]$, which obtains iff $\kappa \leq \min\{2\mu_0, 2-2\mu_0\}$. However, the construction is only needed if $\kappa \leq R_2^*(N)/4$. $R_2^*(N)$ decreases in $N$ and equals $\mu_0(1-\mu_0)$ for $N=0$, thus one can assume $\kappa \leq \mu_0(1-\mu_0)$, and it is easy to verify $\mu_0(1-\mu_0) \leq \min\{2\mu_0, 2-2\mu_0\}$.

**(v)** Follows like part (iii).

**(vi)** Is about to be added.

**Proposition 3**

**(i)** With binary outcomes, $||P(Y_{tx}), P(Y_{tx'})||_{TV} = |\mu_{tx}-\mu_{tx'}|$, thus the claim restates proposition 2.

**(ii)** With binary outcomes, $||P(Y_{tx}), P(Y_{tx'})||_{LOR} \leq \gamma$ iff $\log\left((\mu_{tx}(1-\mu_{tx'}))/(\mu_{tx'}(1-\mu_{tx}))\right) \leq \gamma$. I will show that this, in turn, implies $|\mu_{tx}-\mu_{tx'}| \leq (\exp(\gamma/2)-1)\,/\,(\exp(\gamma/2)+1)$, which in turn implies the claim through proposition 2. While the implication

$$||P(Y_{tx}), P(Y_{tx'})||_{LOR} \leq \gamma \Rightarrow |\mu_{tx}-\mu_{tx'}| \leq (\exp(\gamma/2)-1)\,/\,(\exp(\gamma/2)+1)$$

is one-sided, it will be seen that the result is tight.

Thus, consider the problem

$$\max_{(\mu,\rho)\in[0,1]^2} |\mu-\rho| \quad \text{s.t. } \log\frac{\mu(1-\rho)}{\rho(1-\mu)} \leq \gamma.$$

Noting the problem's symmetry, replace $|\mu-\rho|$ with $(\mu-\rho)$ in the objective for tractability. Letting $\lambda$ denote the Lagrange multiplier on the constraint, relevant partial derivatives are

$$\begin{aligned}\nabla_\mu(x) &= 1-\lambda\left(\frac{1}{x}+\frac{1}{1-x}\right)\\ \nabla_\rho(x) &= -1+\lambda\left(\frac{1}{x}+\frac{1}{1-x}\right).\end{aligned}$$

Clearly $\nabla_\mu(x) = \nabla_\mu(1-x) = -\nabla_\rho(x) = -\nabla_\rho(1-x)$. As $\mu=\rho$ corresponds to minimizing the objective, a solution $(\mu^*,\rho^*)$ must have the feature that $\mu^*=1-\rho^*$. Using this constraint to eliminate $\rho$ and reparameterizing the problem by setting $\mu=(1+\Delta)/2 \Leftrightarrow \Delta=|\mu-\rho|$, one has the new problem

$$\max_{\Delta\in[0,1]} \Delta \quad \text{s.t. } \gamma \geq \log\frac{\left(\frac{1+\Delta}{2}\right)^2}{\left(\frac{1-\Delta}{2}\right)^2} = 2\log\frac{1+\Delta}{1-\Delta}.$$

The solution to this is characterized by the constraint binding, establishing the claim.

To see tightness of the bound for the case of two unknown treatments, note that a least favorable prior in proposition 2(i) has the feature that $\mu_{0x} = 1 - \mu_{1x}$, thus $(\mu_{0x}, \mu_{1x})$ from that prior solve the above problem for some $\gamma$. This last argument is not easily extended to the case of known $\mu_0$ however.

**(iii)** With binary outcomes, $D_{KL}(P(Y_{tx})||P(Y_{tx'})) \leq \gamma$ iff $\mu_{tx}\log(\mu_{tx}/\mu_{tx'}) + (1-\mu_{tx})\log((1-\mu_{tx})/(1-\mu_{tx'}))$. The argument is then similar to part (ii), except it is based on analyzing the maximization problem

$$\max_{(\mu,\rho)\in[0,1]^2} |\mu - \rho| \quad \text{s.t. } \mu \log \frac{\mu}{\rho} + (1-\mu)\log\frac{1-\mu}{1-\rho} \leq \gamma.$$

This problem has the same symmetry properties as the one in (ii), so a similar reparameterization leads to

$$\max_{\Delta\in[0,1]} \Delta \quad \text{s.t. } \gamma \geq (1+\Delta/2)\log\frac{1+\Delta/2}{1-\Delta/2} + (1-\Delta/2)\log\frac{1-\Delta/2}{1+\Delta/2} = \Delta\log\frac{1+\Delta/2}{1-\Delta/2}$$

and thus to the expression given in the proposition. Other remarks are as in (ii).

# References


[1] Bahadur, R.R. and L.J. Savage, 1956. The Nonexistence of Certain Statistical Procedures in Nonparametric Problems. *Annals of Mathematical Statistics*, 27, 1115-1122.

[2] Bergemann, D. and K.H. Schlag, 2007. Robust Monopoly Pricing. Cowles Foundation Discussion Paper 1527R.

[3] —, 2008. Pricing Without Priors. *Journal of the European Economic Association (Papers and Proceedings)*, 6, 560-569.

[4] Brock, W.A., 2006. Profiling Problems with Partially Identified Structure. *Economic Journal,* 116, F427-F440.

[5] Brock, W.A., and C.F. Manski, 2008. Competitive Lending with Partial Knowledge of Loan Repayment. Discussion paper, Northwestern University.

[6] Canner, P.L., 1970. Selecting One of Two Treatments when the Responses are Dichotomous. *Journal of the American Statistical Association*, 65, 293-306.

[7] Chamberlain, G., 2000. Econometrics and Decision Theory. *Journal of Econometrics*, 95, 255-283.

[8] —, 2009. Bayesian Aspects of Treatment Choice. Discussion paper, Harvard University.

[9] Cucconi, O, 1968. Contributi all'Analisi Sequenziale nel Controllo di Accettazione per Variabili. *Atti dell' Associazione Italiana per il Controllo della Qualità*, 6, 171-186.

[10] Dehejia, R., 2005. Program Evaluation as a Decision Problem. *Journal of Econometrics*, 125, 141-173.

[11] Eozenou, P., J. Rivas, and K.H. Schlag, 2006. Minimax Regret in Practice – Four Examples on Treatment Choice. Discussion paper, European University Institute.

[12] Glicksberg, I.L., 1952. A Further Generalization of the Kakutani Fixed Point Theorem, with Application to Nash Equilibrium Points. *Proceedings of the American Mathematical Society*, 3, 170-174.

[13] Gupta, S.S., and S.N. Hande, 1992. On Some Nonparametric Selection Procedures. In A.K.Md.E.Saleh (ed.), *Nonparametric Statistics and Related Topics*. Elsevier.

[14] Hirano, K. and J.R. Porter, 2009. Asymptotics for Statistical Treatment Rules. *Econometrica*, 77, 1683-1701.

[15] Manski, C.F., 1989. Anatomy of the Selection Problem. *Journal of Human Resources*, 24, 343-360.

[16] —, 2000. Identification Problems and Decisions under Ambiguity: Empirical Analysis of Treatment Response and Normative Analysis of Treatment Choice. *Journal of Econometrics*, 95, 415-442.

[17] —, 2003. *Partial Identification of Probability Distributions*. Springer Verlag.

[18] —, 2004. Statistical Treatment Rules for Heterogeneous Populations. *Econometrica*, 72: 1221-1246.

[19] —, 2006. Search Profiling with Partial Knowledge of Deterrence. *Economic Journal*, 116, F385-F401.

[20] —, 2007a. Minimax-Regret Treatment Choice with Missing Outcome Data. *Journal of Econometrics*, 139, 105-115.

[21] —, 2007b. *Identification for Prediction and Decision*. Harvard University Press.

[22] —, 2009. Diversified Treatment under Ambiguity. *International Economic Review*, forthcoming.

[23] Manski, C.F. and A. Tetenov, 2007. Admissible Treatment Rules for a Risk-averse Planner with Experimental Data on an Innovation. *Journal of Statistical Planning and Inference* 137: 1998-2010.

[24] Parmigiani, G., 1992. Minimax, Information, and Ultrapessimism. *Theory and Decision*, 33, 241-252.

[25] Savage, L.J., 1951. The Theory of Statistical Decisions. *Journal of the American Statistical Association*, 46, 55-67.

[26] —, 1954. *The Foundations of Statistics*. Wiley.

[27] Schlag, K.H., 2003. How to Minimize Maximum Regret under Repeated Decision-Making. Discussion paper, European University Institute.

[28] —, 2006. Eleven. Discussion paper, European University Institute.

[29] Stoye, J., 2007a. Minimax Regret Treatment Choice with Incomplete Data and Many Treatments. *Econometric Theory*, 23, 190-199.

[30] —, 2007b. Minimax Regret Treatment Choice with Finite Samples and Missing Outcome Data. In G. de Cooman, J. Veinarová, and M. Zaffalon (eds.), *Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications*.

[31] —, 2009a. Minimax Regret Treatment Choice with Finite Samples. *Journal of Econometrics* 151: 70-81.

[32] —, 2009b. Minimax Regret Treatment Choice with Missing Data: An Application to Young Offenders. *Journal of Statistical Theory and Practice*, 3, 239-254.

[33] —, 2009c. Minimax Regret. In S.N. Durlauf and L.E. Blume (eds.), *The New Palgrave Dictionary of Economics*.

[34] Tetenov, A., 2009. Measuring Precision of Statistical Inference on Partially Identified Parameters. Discussion paper, Collegio Carlo Alberto.

[35] Wald, A., 1945. Statistical Decision Functions Which Minimize the Maximum Risk. *Annals of Mathematics*, 46, 265-280.

[36] —, 1950. *Statistical Decision Functions*. Wiley.

[37] Worrall, J., 2007. Evidence in Medicine and Evidence-Based Medicine. *Philosophical Compass* 2: 981-1022.