Minimax Regret Treatment Choice with Finite Samples and Missing Outcome Data

Jörg Stoye
New York University
j.stoye@nyu.edu

Abstract

This paper uses the minimax regret criterion to analyze choice between two treatments when one has observed a finite sample that is plagued by missing data. The analysis is entirely in terms of exact finite sample regret, as opposed to asymptotic approximations or finite sample bounds. It thus extends Manski (2007), who largely abstracts from finite sample problems, as well as Stoye (2006a), who provides finite sample results but abstracts from missing data. Core findings are: (i) Minimax regret is achieved by randomizing over two rules that were identified in the aforecited papers. (ii) For every sample size, there exists a sufficiently small (but positive) proportion of missing data such that if less data are missing, the missing data problem is ignored altogether and Stoye's (2006a) results apply. (iii) For every positive fraction of missing data, the value of additional observations drops to zero at a finite sample size. I also provide the decision problem's value function and briefly touch on optimal sample design as well as unknown propensity scores.

Keywords. Minimax regret, missing data, statistical decision theory, partial identification, treatment evaluation.

1 Introduction

Consider a planner who has to decide whether to assign a binary treatment – e.g., a medical treatment or a labor market intervention – to members of some target population. She can base her choice on observations of outcomes experienced by a sample of subjects, some of whom received the treatment and some of whom served as control group. The signal generated by these observations has two limitations: First, it is generated by a finite sample. Second, it is assumed that some data are missing, that is, a subset of the target population is not represented in the sample, and members of this subset may react to treatment differently from the observable subjects. Thus, the choice scenario simultaneously generates finite sample problems (a standard issue in statistics and econometrics) and problems of incomplete identification (a less standard issue; see Manski 2003 for a survey).

I analyze this situation using minimax regret with respect to expected outcomes as optimality criterion. Importantly, the analysis is entirely in terms of exact finite sample regret, as opposed to asymptotic approximations (as in Hirano and Porter 2005) or bounds on finite sample quantities (as in Manski 2004). It thus extends, and connects, Manski (2007), who analyzes a somewhat more general case but largely abstracts from finite sample problems, and Stoye (2006a), who provides finite sample results but abstracts from missing data.

In fact, both of the aforementioned papers analyze special cases of the present scenario, and their results are linked in a specific way here. Stoye (2006a), by ignoring missing data, characterizes finite sample minimax regret rules for the boundary case where the proportion of missing data vanishes. Manski (2007) provides a finite sample minimax regret rule if at least half the data are missing. The respective solutions are quite different from each other. The perhaps surprising upshot of this paper is that minimax regret can generally be achieved by randomizing over them, where the mixture is degenerate on the previous two findings' domains but creates a smooth transition in between. For the case where the proportion of observable data p is known a priori, two intriguing aspects of the result are the following:

- For every sample size N, there exists a critical p_N^* such that if $p \geq p_N^*$, then the presence of missing data is ignored altogether and the treatment rules from Stoye (2006a) apply. Those decision rules therefore have a certain degree of robustness to missing data.
- The minimax regret value of the decision problem

exhibits nonstandard asymptotic behavior: For every p, the limiting value of regret is exactly achieved beyond some finite N. Thus the value of additional observations drops to zero at some finite sample size.

If p is not known a priori, minimax regret is achieved by presuming that p equals the lowest value that the decision maker considers possible. In particular, if p cannot be bounded away from zero a priori, then minimax regret is achieved by a "no-data rule."

The remainder of this paper is structured as follows. I first set up the decision problem, introduce notation, and provide a brief motivation for minimax regret. The heart of this paper is section 2.2, which provides relevant results from the aforecited papers and then their joint generalization. In section 2.3, I show how to compute the decision problem's value function, section 2.4 briefly discusses optimal sample design, and section 2.5 considers unknown p. Section 3 concludes, and the appendix collects all proofs.

2 Analysis of the Treatment Choice Problem

2.1 Setup and Notation

There is a binary treatment, $T \in \{0,1\}$, that must be assigned, possibly at random, to members of a target population. Two classic examples are clinical trials, where the target population would be all people who suffer from a certain condition, $T = 1$ would denote a medical innovation, and $T = 0$ would be the status quo treatment, and job training for the unemployed, where $T = 1$ would denote training and $T = 0$ no training. To model treatment effects, I use the standard "potential outcomes" notation (Rubin 1974): For every member of the target population, the random variable $Y_1 \in [0,1]$ denotes the outcome that she would experience if assigned to treatment, whereas $Y_0 \in [0,1]$ is the outcome she would experience if assigned to the control group.[1] Of course, only one of the two random variables will be actualized; the other realization remains counterfactual.

The decision maker observed outcomes experienced in a size N simple random sample from a sample population. Members of the sample were assigned treatment according to some design that will initially be taken as given; the question of optimal sample design is considered later. The sample generates an imperfect signal for two reasons: First, it is finite, and random variation in observed outcomes will be fully considered. Second, only a subset of the target population is observable. I model this by presuming that the sample population is a subset of relative probability mass p of the target population. Importantly, it is assumed that while (Y_0, Y_1) is distributed identically across the sample population, its distribution in the unobservable part of the target population can be different. A leading example is if a study was performed on volunteers who might not be fully representative of the target population. As a consequence, the distribution of (Y_0, Y_1) would only be partially revealed even by an infinitely large sample. This is why the problem is inherently a decision problem under ambiguity, and very similar in structure to interval probability problems as well as robust Bayesian inference. See, in particular, Manski (2002, 2005). Previous analyses of the same problem either largely abstracted from the finite sample problem (Manski 2007) or from the ambiguity caused by missing data (Stoye 2006a).

To model the problem, let the random variable $Z \in \{0,1\}$ indicate whether a member of the target population is in the sample population ($Z = 1$) or not ($Z = 0$). Define the random variables $Y_{tz} \equiv (Y_t | Z = z)$ and write $p \equiv \Pr(Z = 1)$, the proportion of observable subjects in the population. I initially assume that p is known. Then a *state of nature* s can be identified with a true distribution of $(Y_{01}, Y_{00}, Y_{11}, Y_{10})$. Assume that a priori bounds on Y_0 and Y_1 are finite, coincide, and that there are no restrictions on their joint distribution, then it is without further loss of generality to set the state space $\mathcal{S}$ equal to $\Delta([0,1]^4)$, the set of distributions over $[0,1]^4$. Most of the discussion will actually restrict outcomes to be binary, i.e. set $\mathcal{S} = \Delta\left(\{0,1\}^4\right)$, more on which below. It is worth noting that $(Y_{01}, Y_{00}, Y_{11}, Y_{10})$ are not restricted to be independent. I will use the following notational conventions: If Y_i is a random variable, then μ_i denotes its expectation and $\overline{y}_i$ a sample mean.

The sample is a simple random sample from the sample population. For any sample point, one treatment is assigned according to the sample design and the according outcome is observed, thus the decision maker sees realizations (t, y_{t1}). Let $\mathcal{S}_N = (\{0,1\} \times [0,1])^N$, with typical element s_N, denote the sample space induced by a sample of size N, i.e. the collection of possible sample realizations. The decision maker has to choose a treatment rule $\delta : \mathcal{S}_N \to [0,1]$ that maps possible sample outcomes into probabilities of assigning treatment 1. In particular, she is allowed to randomize. The set of decision rules δ is labelled $\mathcal{D}$.

Any combination of state and decision rule induces an

[1] The restriction to $[0,1]$ is w.l.o.g. if, and only if, some bounds on outcomes are known a priori.

expected outcome

$$
\begin{aligned}
u(\delta, s) &\equiv \mu_1 \mathbb{E}\delta(s_N) + \mu_0 \left(1 - \mathbb{E}\delta(s_N)\right) \\
&= \left(p\mu_{11} + (1-p)\mu_{10}\right) \mathbb{E}\delta(s_N) \\
&\quad + \left(p\mu_{01} + (1-p)\mu_{00}\right)\left(1 - \mathbb{E}\delta(s_N)\right)
\end{aligned}
$$

Here, $\mathbb{E}\delta(s_N)$ is evaluated given s; although suppressed in the notation, it will also depend on the sample design. Strictly speaking, u is already a risk function with respect to an underlying loss function $L(y_t) = -y_t$. I will take for granted that if s were known, treatments rules would be evaluated according to u. With unknown s, the efficacy of δ will be measured in terms of *minimax regret* relative to u, thus a treatment rule δ^* is optimal if

$$
\begin{aligned}
\delta^* &\in \arg\min_{\delta \in \mathcal{D}} \left\{ \max_{s \in \mathcal{S}} R(\delta, s) \right\}, \\
R(\delta, s) &\equiv \max_{\delta' \in \mathcal{D}} \left\{ u(\delta', s) \right\} - u(\delta, s).
\end{aligned}
$$

The minimax regret criterion minimizes worst-case performance relative to the ex-post optimal expected outcome or, equivalently, relative to the performance of an infeasible "oracle" treatment rule that utilizes full knowledge of $\mathbb{E}(Y_{01}, Y_{00}, Y_{11}, Y_{10})$. Minimax regret was originally suggested by Savage (1951). In the present formulation – which is not the only possible one – it was recently reconsidered in statistics and related fields (Droge 1998, 2006; Eldar et al. 2003; Hirano and Porter 2005; Manski 2004, 2005, 2007; Schlag 2003, 2006; Stoye 2006a, 2007).[2] A motivation for it is that it avoids the imposition of priors and optimizes against states of the world in which the decision maker's action has a large effect. This sets it apart from its main competitors: The Bayesian decision rule requires specification of a subjective prior over states; maximin utility also avoids priors but optimizes against states in which outcomes are very bad, irrespective of whether they are affected by actions. For a historical overview and further heuristic as well as axiomatic discussion, see Stoye (2006b).

Of course, there are many possible sample designs. I will focus on those considered in Stoye (2006a); they may serve as stylized models of real-world sampling schemes and will turn out to be minimax regret optimal. By *stratified assignment*, I henceforth mean that N is even and that exactly half of the sample is allocated to treatment 1. By *randomized assignment*, I mean that sample points are assigned to treatments by independent tosses of a fair coin.

Some comments on this setup are in order.

- In this paper, p cannot depend on t, and N is not a random variable. The story behind this setting is that missing data occur before treatments are assigned, an example being selection of subjects into experimental pools. Manski (2007) considers the more general case where attrition from an experimental pool can be selective by, and potentially in reaction to, treatment assignment. Unfortunately, finite sample analysis of this case is extremely involved, because sample composition becomes a random variable whose exact distribution must be taken into account. Although one specific such case is analyzed below, a general treatment is left to future research.

- The below results presume binary outcomes, i.e. $Y_0, Y_1 \in \{0, 1\}$. For lemma 2, it will be pointed out that this is not necessary. For the other cases, minimax regret treatment rules for $Y_0, Y_1 \in [0, 1]$ can – under regularity conditions on the state space – be generated by a technique due to Schlag (2003, 2006). The trick, which will be called *binary randomization*, is to replace every sample realization y_i by the outcome of one independent toss of a coin with parameter y_i and then apply the below treatment rules to the resulting, binary samples.

- Covariates are not introduced into this paper's notation, but the results immediately extend to the case of finite-valued covariates by means of proposition 3 in Stoye (2006a). Specifically, let there be a covariate X and let the sample be stratified by covariate, then minimax regret is achieved by applying the below treatment rules separately across covariates. For treatment assignment conditional on $X = x$, the treatment rule therefore utilizes only the subsample with covariate value x. The surprising aspect of this is that there is no inference across covariates. See Stoye (2006a) for an in-depth discussion.

- The decision problem can clearly be interpreted as an imprecise probability problem. The present specification respresents a very special case, however, because complete ignorance about true probabilities is presumed. Prior information can be introduced by restricting the state space $\mathcal{S}$, as is done in Stoye (2006a). This poses no conceptual difficulties but may, of course, change computations.

The remainder of the paper is concerned with finding δ^* for different decision scenarios. The proofs exploit the fact that δ^* can be represented as the decision maker's equilibrium strategy in a fictitious zero-sum

[2] Minimax regret is also closely related to the competititve ratio; indeed, it could as well be called *competititve difference.*

game against Nature. This allows one to infer existence of a minimax regret treatment rule from known game theoretic results (Glicksberg 1952). Other than that, it just restates that the minimax regret decision rule can be characterized as a saddle point, but the game theoretic interpretation facilitates the import of heuristics and solution strategies developed by economists.

2.2 Treatment Rules

The first step is to analyze the aforementioned boundary cases, namely $p = 1$ and $p \leq 1/2$.

Lemma 1 *If $p = 1$, minimax regret is achieved by*

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

where

$$\begin{aligned} I_N &\equiv \#\,(\textit{observed successes of treatment } 1) \\ &\quad +\#\,(\textit{observed failures of treatment } 0) \\ &\quad -\#\,(\textit{observed failures of treatment } 1) \\ &\quad -\#\,(\textit{observed successes of treatment } 0) \\ &\propto N_1(\overline{y}_{11} - 1/2) - N_0(\overline{y}_{01} - 1/2), \end{aligned}$$

where N_t is the number of sample subjects assigned to treatment t. For a stratified sample design, this is equivalent to

$$\delta_1^* \equiv \begin{cases} 0, & \overline{y}_{11} < \overline{y}_{01} \\ 1/2, & \overline{y}_{11} = \overline{y}_{01} \\ 1, & \overline{y}_{11} > \overline{y}_{01} \end{cases}.$$

Lemma 2 *If $p \leq 1/2$, minimax regret is achieved by*

$$\delta_2^* \equiv \frac{1}{2} + \frac{p}{2(1-p)}\frac{I_N}{N}.$$

This applies for either stratified or randomized sample design; in the former case, it can be rewritten as

$$\delta_2^* \equiv \frac{1}{2} + \frac{p\,(\overline{y}_{11} - \overline{y}_{01})}{2(1-p)} = \frac{(p\overline{y}_{11} + 1 - p) - p\overline{y}_{01}}{2(1-p)}$$

as in Manski's (in press) proposition 2.

Lemma 1 is from Stoye (2006a, proposition 1). Lemma 2 follows from Manski (2007, proposition 2) for stratified samples but not for randomized ones. The latter design allows for empty sample cells, a case that Manski has to exclude. The generalization presented here is new.

Also, while lemma 2 is here stated for $Y_0, Y_1 \in \{0, 1\}$, inspection of the proof reveals that δ_2^* can be extended to $Y_0, Y_1 \in [0, 1]$ by using the above definition of I_N in terms of $(N_0, N_1, \overline{y}_{01}, \overline{y}_{11})$. This is not the rule that would emerge from applying the binary randomization technique and then operating δ_2^* on the resulting, binary sample; in particular, it randomizes with probability 0 if (Y_{01}, Y_{11}) has a continuous distribution. This illustrates that minimax regret treatment rules need not be unique.

The next lemma and definition set the stage for the general problem, i.e. $p \in [0, 1]$. From Manski (2007, see also Stoye 2007), we know a minimax regret decision rule for the limiting case where the expectations (μ_{01}, μ_{11}) of (Y_{01}, Y_{11}) are known.

Lemma 3 *Let (μ_{01}, μ_{11}) be known, then minimax regret is achieved by*

$$\begin{aligned} \delta_3^* &\equiv \begin{cases} 0, & \delta < 0 \\ \delta, & 0 \leq \delta \leq 1 \\ 1, & 1 < \delta \end{cases} \\ \delta &\equiv \frac{1}{2} + \frac{p}{2(1-p)}(\mu_{11} - \mu_{01}). \end{aligned}$$

This rule is essentially the population analog of δ_2^*; it just adds a truncation to insure that $\delta_3^* \in [0, 1]$. As final preliminary step, I define its sample analog:

Definition 1 *The sample analog of δ_3^* is*

$$\begin{aligned} \delta_4^* &\equiv \begin{cases} 0, & \delta < 0 \\ \delta, & 0 \leq \delta \leq 1 \\ 1, & 1 < \delta \end{cases} \\ \delta &\equiv \frac{1}{2} + \frac{p}{2(1-p)}\frac{I_N}{N}. \end{aligned}$$

To accommodate both sample designs, δ_4^* is based on I_N/N rather than $(\overline{y}_{11} - \overline{y}_{01})$. Of course, these expressions coincide under the stratified design.

I am now ready to state this paper's main result.

Theorem 4 *Consider any fixed $N < \infty$ as well as $p \in (0, 1]$. Then minimax regret is achieved by the following randomization over δ_1^* and δ_4^*:*

$$\delta^* \equiv \begin{cases} \delta_1^* & \textit{with probability } \alpha^* \\ \delta_4^* & \textit{with probability } (1-\alpha^*) \end{cases},$$

where

$$\begin{aligned}
\alpha^* &\equiv \min\left\{\frac{\frac{p}{2(1-p)}-A}{B-A},1\right\}\\
A &\equiv 2^{-N^*}\sum_{n\geq N^*(1-\frac{1}{2p})}\left(\begin{array}{c}N^*\\n\end{array}\right)(2n-N^*)\\
&\quad\times\min\left\{\frac{1}{2}+\frac{p}{2(1-p)}\frac{2n-N^*}{N^*},1\right\}\\
B &\equiv 2^{-N^*}\sum_{n>N^*/2}\left(\begin{array}{c}N^*\\n\end{array}\right)(2n-N)
\end{aligned}$$

and

$$N^*=\left\{\begin{array}{ll}N, & N \textit{ is odd}\\ N-1, & N \textit{ is even}\end{array}\right. .$$

In particular, α^* equals 1, and the decision rule therefore collapses to δ_1^*, iff $p\geq p_N^*\equiv\frac{2B}{2B+1}$. This threshold value converges to 1 as $N\to\infty$. On the other hand, α^* equals 0, and the decision rule therefore collapses to δ_4^*, iff $p\leq 1/2$.

If $p=0$, then $\delta^*=1/2$.

Substantively, it turns out that minimax regret is achieved by randomizing over δ_1^* and δ_4^*. In words, the decision maker should toss a (biased) coin and then use rule δ_1^* if the coin came up head. The randomization parameter α^* changes with p and N in interesting ways:[3]

- For any given N, δ_4^* applies for $p\leq 1/2$ and its weight then decreases, with δ_1^* being attained for $p\geq p_N^*$, a value that is strictly below 1. Thus for every N, a sufficiently small but nonzero mass of missing data can be ignored. Although p_N^* converges to 1 at rate $N^{-1/2}$, it significantly differs from 1 for rather large N, so that δ_1^*, which was developed for fully observable data, exhibits quite some robustness to missing data.

- For any given $p\in(1/2,1)$, the randomization changes with N as follows: For N small enough, the presence of missing data is ignored, i.e. $\alpha^*=1$, but α^* converges to zero as N grows, so that the limit rule is approximated (but not attained) for large N. Again, the convergence of α^* to 0 is perhaps surprisingly slow; it is also nonuniform in p. It should be pointed out, however, that for any N, δ_4^* becomes similar to δ_1^* as $p\to 1$; thus, it does not follow that convergence of the treatment rule to its limit is "slow" (or nonuniform in p) in every interesting metric.

2.3 Value Function

By evaluating regret on the fictitious game's equilibrium path, one can find the minimax regret achievable under either treatment assignment rule.

Proposition 5 *For either treatment assignment scheme, the decision problem has minimax regret value $(1-p)/2$ if $p<p_N^*$ and*

$$\max_{a\in[1/2,1]}\left\{\begin{array}{c}(2p(a-1)+1)\\ \times\sum_{n<\frac{N^*}{2}}\left(\begin{array}{c}N^*\\n\end{array}\right)a^n(1-a)^{N^*-n}\end{array}\right\}$$

otherwise, where N^* is as in theorem 4. In particular, if $p\leq 1/2$, then the minimax regret value equals $(1-p)/2$ for any N.

As in Stoye (2006a) and Schlag (2006), learning only occurs with every other sample point. Unlike in those papers, learning is incomplete: As $N\to\infty$, regret does not converge to zero but to $(1-p)/2$. This reflects the fact that in the presence of missing data, even the asymptotic decision problem will generate positive regret. An especially unusual feature is that learning occurs only as long as $p\geq p_N^*\Leftrightarrow\alpha^*=1$. When this region of parameter space is left, regret "locks in" at its limiting value.[4] Recall that this occurs for some finite sample size for any $p<1$; what's more, it is the case for *any* sample size if $p\leq 1/2$. This insight generalizes Manski's (2007) finding that when at least half of the data are missing, minimax regret is independent of sample size. More generally, the presence of any missing data whatsoever means that the limiting decision quality is exactly attained for some finite N, and additional observations are useless beyond that threshold. At least from an econometrician's perspective, this finding is unexpected.

2.4 Optimal Sample Design

The preceding analysis took two different sample designs as given. They turn out to generate the same maximal expected regret whenever both are feasible, i.e. when N is even. An obvious question is whether this regret is optimal when sample design is itself a choice variable. The answer is in the affirmative.

[3] MATLAB code that evaluates α^* is available on the author's webpage at `http://homepages.nyu.edu/~js3909`.

[4] The intuition is that in the fictitious game, the switch to $\alpha^*<1$ marks the transition to a *pooling equilibrium* in which the signal generated by sample data is noninformative about the true state of the world. Hence, the decision maker ceases to learn from the signal.

To formalize this idea, let $h_n \equiv (t^i, y^i_{t1})^n_{i=1}$ denote the sample history up to realization n (with the understanding that $h_0 = \varnothing$). If assignment of treatments to sample points is a choice variable, this can be modelled by letting the decision maker choose a vector of mappings $\tau = (\tau_n)^N_{n=1}$, where $\tau_n(N, h_{n-1}) \in [0,1]$ specifies the probability of assigning treatment 1 to sample point n conditional on history h_{n-1} in a sample of overall size N. The randomized assignment scheme corresponds to $\tau^{\text{rand}}_n = 1/2, \forall n$, whereas one way to realize the stratified sample design is to set $\tau^{strat}_n \equiv \mathbb{I}\{n \text{ is even}\}$. Observe that the sampling scheme τ may depend on sample history, but that the decision maker has to specify it before seeing any sample points. This is a common formalization in econometrics because it corresponds to the concept of risk functions in statistical decision theory, but also because it is often realistic for the problems that economists consider. For example, assignment to job training is typically planned by the researcher but executed by caseworkers or other third parties. However, if one wanted to model an online problem, one might also want to allow the decision maker to re-optimize τ along the sample path, which does not in general lead to the same problem.[5]

Proposition 6 *Both τ^{rand} and τ^{strat} (when applicable) are minimax regret optimal assignment schemes.*

This result and proposition 3 in Stoye (2006a) jointly imply that if one faces a random sample from a population with a finite-valued covariate and can choose the sample design, then one can achieve minimax regret by using the randomized assignment scheme and applying theorem 4 separately across covariates. Compared to Manski (2007), this result applies to a narrower range of missing data scenarios but is more general on two other dimensions: p may exceed $1/2$, and the treatment scheme is defined even when some sample cells are empty.

2.5 Treatment Choice with Unknown Propensity Score

I now turn to the case where p is not known a priori but has to be learned from the data. Thus, assume that p can merely be restricted to lie in an interval $[\underline{p}, \overline{p}] \subseteq [0,1]$. The sample size N continues to refer to the number of observed units. This leaves open the question of how exactly learning about p happens as samples are realized. It turns out that this question is irrelevant: Minimax regret can be achieved by presuming that p equals $\underline{p}$.

Proposition 7 *Consider the setting of theorem 4, except that $p \in [\underline{p}, \overline{p}]$. Then minimax regret is achieved by setting $p = \underline{p}$ and applying δ^*.*

The intuition for this result is straightforward: As can be seen from proposition 6, minimax regret decreases in p. This is also intuitive since a higher p means that signals are representative of a larger part of the target population, and should thus be more informative. It implies that the worst case scenario is given by $p = \underline{p}$.

Proposition 4 has an unsettling implication: If $\underline{p} = 0$, minimax regret is achieved by setting $\delta^* = 1/2$ irrespective of the observed sample, i.e. by a "no-data rule." As observed by Savage (1954) and recently by Manski (2004), Schlag (2003), and Stoye (2006a), no-data rules are a frequent problem with the maximin utility criterion. Only Stoye (2006a) previously found similar problems to arise with minimax regret. Proposition 7 provides another, realistic problem that leads to a no-data minimax regret treatment rule. It suggests that the issue of no-data rules may not constitute the most compelling argument for minimax regret over maximin utility.

3 Summary and Outlook

This paper added to the recently growing literature on minimax regret and specifically to research by Manski (2007) and Stoye (2006a). It provided a joint generalization of much of those papers' analyses by considering a treatment choice problem where information is incomplete in two ways, firstly because of finite sample variation but also, and more fundamentally, because of missing data and hence incomplete identification of population distributions. The core finding is that results by Manski (2007) and Stoye (2006a) can be linked in a particular way: Each of them identifies a minimax regret treatment rule for a boundary case of the present problem, and a smooth transition between these solutions is generated by randomizing over them. This insight also strengthens the general finding that minimax regret tends to prescribe randomization, a point stressed by Schlag (2003, 2006). The result was extended by presenting the decision problem's value function, by allowing for unknown or partially known propensity scores, and by showing optimality of certain sample designs.

Many questions remain open on minimax regret treat-

[5] One might think that the difference cannot matter, because the definition of τ allows the decision maker to prescribe reactions to sample observations. In fact, this depends on how the decision maker reacts to the arrival of information, that is, on her updating rule. Under the most intuitive such rule, namely pointwise Bayesian updating of the state space, both minimax regret and maximin utility are dynamically inconsistent, meaning that the conjecture is false. What's more, the present choice of suppressing updating then appears not only realistic but sensible (e.g., Augustin 2003). Hanany and Klibanoff (2005) provide an updating rule that renders the conjecture true.

ment choice. For example, Stoye (2007) generalizes Manski (2007) in a different direction, namely by allowing for a multi-valued treatment. This generalization could be further extended by considering finite samples. The same remark holds for additional results that Stoye (2006a) presents with respect to covariates and the effect of restricting $\mathcal{S}$, as well as Manski's (2007) consideration of sample attrition that may vary by assigned treatment. When designing sample designs, one could also consider the possibility that outcomes experienced by the sample subjects are taken into account when evaluating the design. This possibility was ignored here for simplicity; it leads to intricate "bandit" problems as in Schlag (2003).

Acknowledgements

I thank the referees, the program chairs, Gary Chamberlain, and seminar audiences at Harvard/MIT and Pittsburgh for helpful comments.

A Proofs

Preliminaries Most proofs proceed by analyzing the following zero-sum game: (i) The decision maker (DM) chooses a statistical treatment rule $\delta : \theta \to [0,1]$, Nature chooses a mixed strategy $\sigma \in \Delta(\mathcal{S})$ over states. (ii) A neutral meta-player draws s according to σ, then θ according to s. (iii) DM's payoff is $\int R(\delta, s) d\sigma$. This game is useful because of the following fact (e.g., Berger 1985).

Lemma 8 *Assume that $\sigma^* \in \Delta(\mathcal{S})$ and δ^* are such that (δ^*, σ^*) is a Nash equilibrium of the above game, that is, $\delta^* \in \arg\min_{\delta \in \mathcal{D}} \int R(\delta, s) d\sigma^*$ and $\sigma^* \in \arg\max_{\sigma \in \Delta(\mathcal{S})} \int R(\delta^*, s) d\sigma$. Then δ^* is a minimax regret treatment rule.*

Proofs will, therefore, proceed by conjecturing and then verifying Nash equilibria of the fictitious game (as also in Schlag 2003, 2006, and Stoye 2006a, 2007).

Lemma 1 See Stoye (2006a, propositions 1 and 2).

Lemma 2 Consider first the randomized treatment assignment scheme. Assume that DM plays δ^*, then any $P(Y_{01}, Y_{00}, Y_{11}, Y_{10})$ in the support of σ^* must maximize $R(\delta^*, s)$. Expansion of $R(\delta^*, s)$ yields

$$\max\{(p\mu_{11} + (1-p)\mu_{10} - p\mu_{01} - (1-p)\mu_{00}) \times \mathbb{E}\left(\frac{1}{2} + \frac{p}{2(1-p)}\frac{I_N}{N}\right), (p\mu_{01} + (1-p)\mu_{00} - p\mu_{11} - (1-p)\mu_{10}) \times \mathbb{E}\left(\frac{1}{2} - \frac{p}{2(1-p)}\frac{I_N}{N}\right)\}.$$

Simple calculations show that distribution of I_N depends on $P(Y_{01}, Y_{00}, Y_{11}, Y_{10})$ only through (μ_{01}, μ_{11}). Thus $R(\delta^*, s)$ depends on $P(Y_{01}, Y_{00}, Y_{11}, Y_{10})$ only through $(\mu_{01}, \mu_{00}, \mu_{11}, \mu_{10})$. Furthermore, symmetry of the two components of the max-operator means that $(\mu_{01}, \mu_{00}, \mu_{11}, \mu_{10}) = (a, b, c, d)$ maximizes $R(\delta^*, s)$ iff $(\mu'_{01}, \mu'_{00}, \mu'_{11}, \mu'_{10}) \equiv (c, d, a, b)$ does. One can thus construct a best response to δ^* by finding some $(\mu^*_{01}, \mu^*_{00}, \mu^*_{11}, \mu^*_{10})$ that maximizes

$$(p\mu_{11} + (1-p)\mu_{10} - p\mu_{01} - (1-p)\mu_{00}) \times \mathbb{E}\left(\frac{1}{2} + \frac{p}{2(1-p)}\frac{I_N}{N}\right)$$

and presuming that Nature randomizes evenly between $(\mu^*_{01}, \mu^*_{00}, \mu^*_{11}, \mu^*_{10})$ and its symmetric counterpart. I will now find $(\mu^*_{01}, \mu^*_{00}, \mu^*_{11}, \mu^*_{10})$ and then verify that δ^* is a best response to Nature's strategy.

In the proof of proposition 1(ii) in Stoye (2006a), it is established that the distribution of I_N depends on (μ_{01}, μ_{11}) only through $\mu_{11} - \mu_{01}$. Without loss of generality, I therefore presume that $(\mu_{01}, \mu_{11}) = \left(\frac{1-\Delta}{2}, \frac{1+\Delta}{2}\right)$ for some $\Delta \in [-1, 1]$. Observe furthermore that since I_N is a sum of N realizations of an i.i.d. random variable, $\mathbb{E}(I_N/N) = \mathbb{E}I_1 = \frac{1}{2}(\mu_{11} - (1-\mu_{11})) - \frac{1}{2}(\mu_{01} - (1-\mu_{01})) = \Delta$. Thus, we can define $(\mu^*_{01}, \mu^*_{00}, \mu^*_{11}, \mu^*_{10})$ as maximizer of

$$(p\mu_{11} + (1-p)\mu_{10} - p\mu_{01} - (1-p)\mu_{00}) \times \left(\frac{1}{2} + \frac{p(\mu_{01} - \mu_{11})}{2(1-p)}\right).$$

Clearly this requires that $\mu^*_{10} = 1$, $\mu^*_{00} = 0$, and that $\Delta^* \equiv \mu^*_{11} - \mu^*_{01}$ maximize

$$(p\Delta + 1 - p) \times \left(\frac{1}{2} + \frac{p\Delta}{2(1-p)}\right) = \frac{1-p}{2} - \frac{p^2\Delta^2}{2(1-p)},$$

which obtains whenever $\Delta = 0 \Leftrightarrow \mu_{01} = \mu_{11}$. It follows that under Nature's best response, observations of Y_{01} and Y_{11} are uninformative, and the decision maker is indifferent between all treatment rules. In particular, δ^* is a best response.

The proof is essentially the same for stratified sampling. In that case, $\mathbb{E}I_N$ can be directly written as linear function of (μ_{01}, μ_{11}), so that proposition 1(ii) from Stoye (2006a) need not be invoked.

Lemma 3 Follows from Manski (2007, proposition 1); see also Stoye (2007, corollary 1).

Theorem 4 I restrict attention to the randomized treatment assignment scheme and also assume N to be odd; the extension to stratified sampling as well as even N follows along the lines of proposition 1 in Stoye (2006a). The core idea is to restrict DM's (pure) strategy space to $\{\delta_1^*, \delta_4^*\}$, rendering the game more tractable. Of course, it must be shown that equilibria of the simplified game are also equilibria of the original one. Thus, identify DM's strategy with $\alpha \in [0,1]$, the probability with which δ_1^* is played. As in lemma 2, the distribution of δ^* depends on s only through $\Delta \equiv \mu_{11} - \mu_{01}$. Nature will therefore pick $(\mu_{01}^*, \mu_{00}^*, \mu_{11}^*, \mu_{10}^*) \in [0,1]^4$ to maximize $R(\alpha, s)$, which can be expanded to

$$\begin{aligned}
&\max\{(p\mu_{11} + (1-p)\mu_{10} - p\mu_{01} - (1-p)\mu_{00}) \\
&\quad \times (1 - \alpha f_1(\mu_{11} - \mu_{01}) - (1-\alpha) f_0(\mu_{11} - \mu_{01})), \\
&\qquad (p\mu_{01} + (1-p)\mu_{00} - p\mu_{11} - (1-p)\mu_{10}) \\
&\qquad \times (\alpha f_1(\mu_{11} - \mu_{01}) + (1-\alpha) f_4(\mu_{11} - \mu_{01}))\},
\end{aligned}$$

where

$$f_i(d) \equiv \mathbb{E}(\delta_i^* | \mu_{11} - \mu_{01} = d).$$

Some observations from the proof of lemma 2 apply: The objective function is symmetric, so that to find best responses, one can restrict attention to maximizers of the first element. Such maximizers must have $(\mu_{10}, \mu_{00}) = (1, 0)$, and the optimization problem can be reduced to maximization over $\Delta \in [-1, 1]$ of

$$\phi(\Delta; p, \alpha) \equiv (p\Delta + 1 - p)(1 - \alpha f_1(\Delta) - (1-\alpha) f_4(\Delta)).$$

Notice that f_i and ϕ are differentiable in their arguments; this will be used as first-order conditions will be evaluated. To construct Nash equilibria, it will be assumed that Nature randomizes evenly over the maximizer such found and its symmetric counterpart. The new arguments relative to lemma 2 are as follows.

Step 1: By the same arguments that apply to the original game, the simplified game possesses Nash equilibria. These must fall into one of three classes:

(i) Separating equilibria: Assume $\Delta > 0$, then the better treatment is the one that has higher expected success in observable units. The sampling distribution is binomial and thus possesses a monotone likelihood ratio property. It follows that δ_1^* (respectively $\alpha = 1$) is a best response.

(ii) Pooling equilibria: Assume $\Delta = 0$, then the signal generated by the sample is uninformative. Any decision rule constitutes a best response to this. The equilibrium from lemma 2 is an example of this case.

(iii) Negatively separating equilibria: Assume $\Delta < 1$, then the sample generates an informative signal, but the decision maker wants to act *against* this signal. In the simplified game, her best response would therefore be δ_4^*, which is less sensitive to the signal than δ_1^*.

The first two cases have in common that DM's equilibrium strategy remains a best response in her unrestricted strategy space. Whenever the simplified game's equilibrium falls into one of these cases, it therefore is an equilibrium of the original game as well. This does not hold if the equilibrium is negatively separating, in which case the decision maker's unrestricted best response would be $\delta_5^* \equiv 1 - \delta_1^*$.

Step 2: I will now show that a negatively separating equilibrium cannot obtain. It follows that equilibria of the simplified game are either separating or pooling, and thus coincide with equilibria of the original game.

To show the claim, suppose that DM plays δ_4^*. This leads to a negatively separating equilibrium iff Nature's best response is some $\Delta^* < 0$. The accordingly constrained value of her response problem is

$$\begin{aligned}
&\sup_{\Delta \in [-1,0)} \phi(\Delta; p, 0) \\
= &\sup_{\Delta \in [-1,0)} (p\Delta + 1 - p)(1 - f_4(\Delta)).
\end{aligned}$$

For comparison, the problem of maximizing

$$\rho(\Delta; p) \equiv (p\Delta + 1 - p)(1 - f_3(\Delta))$$

was considered in lemma 2; recall it is solved by $\Delta = 0$ and has value $(1-p)/2$. Substitute the definitions of δ_3^* and δ_4^* into the definition of f_i to find

$$\begin{aligned}
f_3(\Delta) &= \mathbb{E}_{B(\Delta,N)} d^* \\
f_4(\Delta) &= \mathbb{E}_{B(\Delta,N)} d,
\end{aligned}$$

where

$$\begin{aligned}
d^* &= \frac{1}{2} + \frac{p(2n - N)}{2N(1-p)} \\
d &= \begin{cases} 0, & d^* < 0 \\ d^*, & 0 \le d^* \le 1 \\ 1, & d^* > 1 \end{cases}
\end{aligned}$$

and where $\mathbb{E}_{B(\Delta,N)}$ denotes expectation with respect to the distribution of n, which is binomial with parameters (Δ, N). From inspection of these, it is elementary that $f_4(\Delta)$ lies between $f_3(\Delta)$ and $1/2$ for any (Δ, p); specifically, $f_4(\Delta) \ge f_3(\Delta)$ whenever $\Delta < 0$. It follows that $\Delta \le 0 \Rightarrow \phi(\Delta; p, 0) \le \rho(\Delta; p)$. Hence,

$$\sup_{\Delta \in [-1,0)} \phi(\Delta; p, 0) \le \sup_{\Delta \in [-1,0)} \rho(\Delta; p) = (1-p)/2,$$

and this supremum is furthermore not attained on $[-1, 0)$. But $\phi(\Delta; p, 0) = (1-p)/2$, so $\Delta = 0$ is a strictly better response to δ_4^* than any $\Delta < 0$.

Step 3: It remains to characterize separating respectively pooling equilibria. The main tool for this will be evaluation of first-order conditions. For a separating equilibrium, one must have $0 \le \arg\max_\Delta \phi(\Delta; p, 1)$. Consider the partial derivatives

$$\begin{aligned}
\phi_\Delta(\Delta; p, 1) &= -f_1'(\Delta)(p\Delta + 1 - p) + p(1 - f_1(\Delta)) \\
\phi_{\Delta p}(\Delta; p, 1) &= (1 - \Delta) f_1'(\Delta) + 1 - f_1(\Delta) > 0.
\end{aligned}$$

Since the cross-derivative is positive, $\arg\max_\Delta \phi(\Delta; p, 1)$ increases in p in strong set order (that is, its smallest and largest element increase) by standard supermodularity arguments. Hence, the separating equilibrium can be maintained for $p > p_N^*$, where p_N^* is implicitly defined by

$$0 \in \arg \max_{\Delta \in [-1,1]} \phi(\Delta; p_N^*, 1).$$

An expression for p_N^* can be derived by inspecting the first-order condition:

$$\phi_\Delta(0, p_N^*; 1) = 0.$$

The previous expression for $\phi_\Delta(\Delta, p; 1)$ can be simplified at $\Delta = 0$. Write

$$\begin{aligned} f_1(\Delta) &= \Pr(I_N > N/2) \\ &= \sum_{n > N/2} \binom{N}{n} \left(\frac{1+\Delta}{2}\right)^n \left(\frac{1-\Delta}{2}\right)^{N-n}, \end{aligned}$$

which implies that (after some simplification)

$$f_1'(0) = 2^{-N} \sum_{n > N/2} \binom{N}{n} (2n - N) \equiv B.$$

Also observing that $f_1(0) = 1/2$, the first-order condition becomes

$$-(1-p)B + \frac{p_N^*}{2} = 0 \Longrightarrow p_N^* = \frac{2B}{2B+1}.$$

To see convergence of $p^* = \frac{2B}{2B+1}$ to 1, notice that B can be rewritten as

$$B = \mathbb{E}(|n| - N/2),$$

where n is the number of successes recorded in N independent coin tosses. The convergence rate of binomial distributions to the Normal immediately implies that $B = O\left(N^{1/2}\right)$ and hence that $1 - p_N^* = O(N^{-1/2})$.

Consider now the pooling equilibrium. This equilibrium requires that $\Delta = 0$ maximizes $\phi(\Delta, p; \alpha)$. A necessary condition for this is

$$\begin{aligned} 0 &= \phi_\Delta(0; p, \alpha) \\ &= (p \cdot 0 + 1 - p)\left(-\alpha f_1'(0) - (1-\alpha) f_4'(0)\right) \\ &\quad + p\left(1 - \alpha f_1(0) - (1-\alpha) f_4(0)\right). \end{aligned}$$

Some of the previous simplifications apply again; in particular, $f_4(0) = 1/2$. Substituting in for $f_4(\Delta) = \mathbb{E}_{B(\Delta,N)} d$, one finds that (after simplification)

$$f_4'(0) = 2^{-N} \sum_{n=0}^{N} \binom{N}{n} (2n - N)\, d \equiv A.$$

The first-order condition thus simplifies to

$$\begin{aligned} 0 &= -(1-p)(\alpha B + (1-\alpha)A) + \frac{p}{2} \\ \Longrightarrow \alpha^* &= \frac{\frac{p}{2(1-p)} - A}{B - A}. \end{aligned}$$

This yields an equilibrium iff $\alpha^* \in [0, 1]$. I will show below that $B > A$ and that $A \le \frac{p}{2(1-p)}$, with equality iff $p \le 1/2$. Hence, $\alpha^* \ge 0$ as required, and $\alpha^* = 0$ iff $p \le 1/2$, yielding the equilibrium from lemma 2. Furthermore, α^* equals 1 if $\frac{p}{2(1-p)} = B \Leftrightarrow p = \frac{2B}{2B+1}$, the condition identified for a separating equilibrium.

I conclude by filling the gaps in the preceding paragraph. To see that $A \le \frac{p}{2(1-p)}$, with equality iff $p \le 1/2$, observe that $f_4'(0) \le f_3'(0)$ because $f_4'(\Delta) \le f_3'(\Delta)$ was shown for $\Delta < 0$ in step 2, yet these derivatives are continuous. Hence $A \le f_3'(0)$, but

$$f_3'(0) = \frac{d}{d\Delta} \mathbb{E}_{B(\Delta,N)} \left(\frac{1}{2} + \frac{p(2n-N)}{2N(1-p)}\right) = \frac{p}{2(1-p)}$$

because $\mathbb{E}_{B(\Delta,N)}\left(\frac{2n-N}{N}\right) = \Delta$. For $p > 1/2$, one can minimally expand on arguments from step 2 to show $f_4'(0) < f_3'(0)$, hence $A < \frac{p}{2(1-p)}$.

To see $B > A$, take explicit derivatives of binomial expectations to find (after some simplification)

$$B = 2^{-N} \sum_{n=0}^{N} \binom{N}{n} (2n - N)\, \mathbb{I}\{d^* > 1/2\},$$

thus

$$B - A = 2^{-N} \sum_{n=0}^{N} \binom{N}{n} (2n - N) \left(\mathbb{I}\{d^* > 1/2\} - d\right),$$

but $(2n - N)\left(\mathbb{I}\{d^* > 1/2\} - d\right)$ is easily seen to be over nonnegative for any (n, N). Furthermore, the above sum is strictly positive whenever there exists n for which $\mathbb{I}\{d^* > 1/2\} - d \ne 0$, that is, whenever δ_1^* and δ_4^* do not agree. (They agree iff $p \ge N/(N+1)$, a number that is well above p^* for all N.)

Proposition 5 Follows by algebraically evaluating $\max_s R(\delta^*, s)$ using the above simplifications.

Proposition 6 Assume the decision maker can pick (τ, δ) and that the worst-case prior σ^* is as in proposition 1, then due to that prior's symmetry, the distribution of $(I_{n+1}|I_n)$ does not depend on τ_n. Hence, the decision maker is indifferent between all possible τ; in particular, the randomized design, in conjunction with δ^*, constitutes a best response. That σ^* remains a best response follows immediately from the proof of proposition 1. The conclusion extends to the stratified design because that design generates the same maximal regret as the randomized one, yet a zero-sum game cannot have two Nash equilibria with different values.

Proposition 7 The proof is just as in theorem 4, with the following adjustment: Extend the decision problem, and hence the fictitious game, by identifying the state space with $\mathcal{S} \times [0,1]$ with typical element (s,p). Assume DM sets $p = \underline{p}$ and then uses δ^* from proposition 4. Then by following steps from theorem 4, Nature's best-response problem can be reduced to

$$\max_{p \in [\underline{p}, \overline{p}], \Delta \in [-1,1]} \left\{ (p\Delta + 1 - p)\left(1 - \alpha f_1(\Delta) - (1-\alpha) f_4(\Delta)\right)\right\}.$$

The objective decreases in p – notice especially that since DM uses $p = \underline{p}$, $f_4(\Delta)$ is not a function of Nature's choice of p. Hence, Nature will choose $p = \underline{p}$. The remainder of the proof is unchanged.

References

[1] T. Augustin. On the Suboptimality of the Generalized Bayes Rule and Robust Bayesian Procedures from the Decision Theoretic Point of View — a Cautionary Note on Updating Imprecise Priors. In J.M. Bernard, T. Seidenfeld, and M. Zaffalon (Eds.), *ISIPTA 03: Proceedings of the Third International Symposium on Imprecise Probabilities and their Applications*. Carleton Scientific, 2003.

[2] J.O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer Verlag, 1985.

[3] B. Droge. Minimax Regret Analysis of Orthogonal Series Regression Estimation: Selection Versus Shrinkage. *Biometrika* 85:631-643, 1998.

[4] —. Minimax Regret Comparison of Hard and Soft Thresholding for Estimating a Bounded Normal Mean. *Statistics and Probability Letters* 76:83–92, 2006.

[5] Y.C. Eldar, A. Ben-Tal, and A. Nemirovski. Linear Minimax Regret Estimation of Deterministic Parameters with Bounded Data Uncertainties. *IEEE Transactions on Signal Processing* 52:2177-2188, 2004.

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

[7] E. Hanany and P. Klibanoff. Dynamically Consistent Updating of MaxMin EU and MaxMax EU Preferences. In F.G. Cozman, R. Nau, T. Seidenfeld (Eds.), *ISIPTA 05: Proceedings of the Fourth International Symposium on Imprecise Probabilities and Their Applications*. Carnegie Mellon University, 2005.

[8] K. Hirano and J.R. Porter. Asymptotics for Statistical Treatment Rules. Technical Report, University of Arizona and University of Wisconsin.

[9] C.F. Manski. Treatment Choice under Ambiguity Induced by Inferential Problems. *Journal of Statistical Planning and Inference* 105:67-82, 2002.

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

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

[12] —. *Social Choice with Partial Knowledge of Treatment Response*. Princeton University Press, 2005.

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

[14] D.B. Rubin. Estimating Causal Effects of Treatment in Randomized and Nonrandomized Studies. *Journal of Educational Psychology* 66:688-701, 1974.

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

[16] K.H. Schlag. How to Minimize Maximum Regret in Repeated Decision-Making. Technical Report, European University Institute, 2003.

[17] —. Eleven. Technical Report, European University Institute, 2006.

[18] J. Stoye. Minimax Regret Treatment Choice with Finite Samples. Technical Report, New York University, 2006a.

[19] —. Statistical Decisions Under Ambiguity. Technical Report, New York University, 2006b.

[20] —. Minimax Regret Treatment Choice with Missing Data and Many Treatments. *Econometric Theory* 23:190–199, 2007.