Minimax-regret treatment rules with many treatments

Matthew A. Masten[1]

Abstract

Statistical treatment rules map data into treatment choices. Optimal treatment rules maximize social welfare. Although some finite sample results exist, it is generally difficult to prove that a particular treatment rule is optimal. This paper develops asymptotic and numerical results on minimax-regret treatment rules when there are many treatments. I first extend a result of Hirano and Porter (Econometrica 77:1683–1701, 2009) to show that an empirical success rule is asymptotically optimal under the minimax-regret criterion. The key difference is that I use a permutation invariance argument from Lehmann (Ann Math Stat 37:1–6, 1966) to solve the limit experiment instead of applying results from hypothesis testing. I then compare the finite sample performance of several treatment rules. I find that the empirical success rule performs poorly in unbalanced designs, and that when prior information about treatments is symmetric, balanced designs are preferred to unbalanced designs. Finally, I discuss how to compute optimal finite sample rules by applying methods from computational game theory.

1 Introduction

The purpose of empirical work is often to inform decision-making. Manski (2000, 2004, 2007a, 2009) argued that our statistical methods should reflect the underlying decision problem. He proposed using Wald's (1950) statistical decision theory framework to formally analyze methods for converting sample data into policy

✉ Matthew A. Masten
matt.masten@duke.edu

[1] Department of Economics, Duke University, Durham, USA

decisions. From a Bayesian perspective, Chamberlain (2000) and Dehejia (2005) also argue for the relevance of decision theory to econometrics.

In this paper, I consider statistical decision problems when there are many possible treatments to choose from. In the first part, I show that an empirical success rule which assigns everyone to the treatment with the largest estimated welfare is locally asymptotically minimax optimal under regret loss, a result proved for the two treatment case by Hirano and Porter (2009). As in Hirano and Porter (2009), this asymptotic approach allows the distribution of the data to be arbitrary with unbounded support, whereas most existing finite sample results require the distribution of the data to have bounded support (e.g., (Schlag, 2003, 2006; Stoye, 2009a)).[1] In the second part, I examine the performance of various treatment rules in finite samples. In particular, I show that the empirical success rule performs poorly when the sampling design is highly unbalanced—when some treatments are purposely given a larger proportion of subjects than other treatments. My computations also suggest that balanced designs are preferred to unbalanced designs. I end by discussing how to numerically compute optimal treatment rules by applying results from computational game theory.

Several papers (e.g., (Manski, 2004; Schlag, 2006; Stoye, 2009a, 2012; Tetenov, 2012; Hirano & Porter, 2009)) have analyzed statistical decision problems using the minimax-regret criterion when there is a binary treatment. The first part of this paper extends a result of Hirano and Porter (2009) to the many treatment case. In the second part of this paper, I show how to use the proof strategy of Stoye (2009a) to numerically compute optimal treatment rules when the analytical solution appears intractable.

Few papers discuss minimax-regret treatment rules with many treatments. Prior to 2013, I am only aware of two previous results: First, Stoye (2007b) derives population level treatment rules for more than two treatments (where ambiguity arises due to missing data), but does not consider finite sample rules. Second, a series of papers by Bahadur (1950), Bahadur and Goodman (1952), and Lehmann (1966) showed that the empirical success rule is minimax-regret optimal with many treatments, assuming the data comes from a known parametric family satisfying monotone likelihood ratio in a scalar parameter and under a balanced sampling design (see appendix B for further details). In the first part of this paper, I show how this result may be used to extend a result of Hirano and Porter (2009). There has been more work since 2013: Manski and Tetenov (2016) extend the large deviations analysis of Manski (2004) to multiple treatments to derive a finite sample bound on the maximum regret of the empirical success rule. They then use that result to study the choice of sampling design. Appendix B.3 of Kitagawa and Tetenov (2017) discusses an extension of empirical welfare minimization to multiple treatments (also see (Kitagawa & Tetenov, 2018)). Kallus (2018) and Zhou et al. (2023) derive regret bounds for treatment rules with more than two treatments that also incorporate covariate information.

The asymptotic results here and in Hirano and Porter (2009) apply to arbitrary sampling designs,[2] and hence do not allow us to compare specific designs. Likewise,

---

[1] The finite sample results in section 3 of Tetenov (2012), which allow for unbounded support, are an exception.

[2] Although I require the design to be 'asymptotically balanced'; see section 2.5.

the Bahadur et al results do not apply to unbalanced designs. Regardless of the number of treatments, the finite-sample minimax-regret treatment rule for an unbalanced design is currently unknown. Nonetheless, unbalanced designs are important in practice. First, when there are a large number of treatments, it may be difficult or impossible to gather data on all treatments. Moreover, the costs of treatment may differ, in which case we need to trade off potential gains from a balanced design versus the different costs of treatment. Finally, the traditional statistics literature on experimental design, based on power analysis of hypothesis tests, sometimes recommends unbalanced designs. Depending on the a priori information available, this recommendation may not be optimal when the minimax-regret treatment choice criterion is used instead. My computational results suggest that, without a priori restrictions on treatment response, balanced designs are preferred to unbalanced designs. My results also suggest that designs which do not commit to an allocation in advance, and instead allocate subjects with equal probability to all treatments, are preferred to balanced designs.

Although results on binary treatments are insightful, policy-makers often have to choose between many different options. In these cases, previous research provides insufficient guidance for decision-making. While analytical finite-sample optimality results are preferred, this paper shows that asymptotic and numerical results can be a useful substitute when analytical results are unavailable.

2 Asymptotics for statistical treatment rules with many treatments

Finite sample optimality results are often difficult to derive. For years, asymptotic theory has been used instead when finite sample result are unavailable. Much of the foundational work on asymptotics followed Wald's (1945) statistical decision function approach, culminating in Le Cam's (1986) magnum opus. In this general formulation, statistics is viewed as a formal tool for making decisions with finite sample data, where the decision maker incurs real losses if she makes a suboptimal decision. This view was seemingly forgotten along the way, as researchers focused on convenient choices of loss functions which led to now-standard work on estimation problems and hypothesis tests. Recent work by Chamberlain (2000), Dehejia (2005), and Manski (2000, 2004), has renewed interest in the decision theory view of statistics.

In particular, Hirano and Porter (2009) applied Le Cam's (1986) local asymptotic theory to the comparison of statistical treatment rules when there are two treatments to choose from.[3] This approach allows them to derive locally asymptotically optimal rules under weaker assumptions than needed to derive exact finite-sample results, as in Stoye (2009a). This asymptotic approach allows

---

[3] Instead of using local asymptotics, Otsu (2008) applies the large deviations approach to compare treatment rules. Also see Manski (2004), who uses a large deviation result to derive finite sample bounds on maximal regret.

the data to have unbounded support, and the sampling design may be anything which point identifies the parameters.

In this section, I extend one of Hirano and Porter's (2009) results to show that an empirical success rule is asymptotically optimal under the minimax-regret criterion when there are an arbitrary, but finite, number of treatments. Specifically, letting $\delta^*_{k,N}$ denote the proportion of people to be assigned to treatment k, and letting $\mathcal{M}_k = \{s \in \{1, \ldots, K\} : w(\hat{\theta}_{k,N}) = w(\hat{\theta}_{s,N})\}$, I show that

$$\delta^*_{k,N} = \begin{cases} 1/|\mathcal{M}_k| & \text{if } w(\hat{\theta}_{k,N}) \geq w(\hat{\theta}_{s,N}) \text{ for all } s \neq k \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

is locally asymptotically minimax optimal under regret loss, where $w(\hat{\theta}_{k,N})$ is a 'best' estimate of the welfare achieved by treatment k. Such a function which maps data into an allocation of treatments to individuals is called a treatment rule.

As in Hirano and Porter (2009), I use Le Cam's limits of experiments framework. This framework splits the problem of deriving asymptotically optimal rules into four steps: (1) establish that the data generating process converges to a 'limit experiment', where one observes a single draw from a specific distribution, often a mean-shifted normal, (2) show that no sequence of rules can do better than the optimal rule in the limit experiment, a result called the asymptotic representation theorem, (3) derive the optimal rule in the limit experiment, and (4) construct a sequence of rules which converges to the optimal rule from step 3. The key difference between my result and that of Hirano and Porter is step 3: solving the limit experiment. Since they consider only two treatments, they are able to apply finite sample optimality results from hypothesis testing theory, namely the Neyman–Pearson lemma. For more than two treatments, I instead apply results of Bahadur (1950), Bahadur and Goodman (1952), and Lehmann (1966) on picking the normal population with the largest mean, which use permutation invariance arguments.

The rest of this section is organized as follows: In section 2.1 I specify the setup of the statistical decision problem. In section 2.2 I derive the distribution of plug-in estimators of welfare based on estimators like the MLE. I then state an asymptotic representation theorem in section 2.3. Consequently, in section 2.4, I derive the optimal treatment rule in the limit experiment. Finally, I show in section 2.5 that the plug-in rule matches the optimal treatment rule in the limit experiment, and hence is locally asymptotically optimal.

2.1 Setup

I begin by describing the general setup of a statistical decision problem used in Manski (2004). I then specialize that setup to the case of finitely many treatments and describe the data generating processes under consideration.

2.1.1 General statistical decision theory setup

A treatment $t \in \mathcal{T}$ can be applied at the individual level to members of some population. When individual i receives treatment t, she experiences the outcome $Y_i(t) \in \mathcal{Y}$. Let P_t denote the distribution of outcomes in the population that would occur if we assigned everyone to treatment t. Assume P_t is in a parametric family of distributions, so that for each treatment t, $P_t = P_{\theta_t}$ for some finite vector θ_t and a known function $P_{\theta_t} = P(\theta_t)$. Let $\theta = \{\theta_t : t \in T\}$. θ is called the *state of the world*.

Example 1 A simple example is to let the density of P_t be the location model $f(y - \theta_t)$, where f is a known density function, symmetric about zero. In this model, $\theta_t \in \mathbb{R}$ is the median of P_t. θ_t is also the mean of P_t, if it exists.

Suppose outcome distributions are ranked by a scalar mapping W, called the *welfare function*. Larger values of welfare are preferred. For example, we may rank outcome distributions by their average outcome: $W(P_t) = \mathbb{E}_{\theta_t}[Y_i(t)]$. Let $w(\theta_t) = W(P_{\theta_t})$ denote the welfare achieved when all people are assigned to treatment t and the true state of the world for treatment t is θ_t. If we knew θ, then the welfare $w(\theta_t)$ would be known for all treatments t and hence to maximize welfare we would solve

$$\sup_{t \in \mathcal{T}} \; w(\theta_t). \tag{2}$$

Unfortunately, we do not know the true state of the world θ. To learn about θ, we gather sample data ω which lies in some sample space Ω. Given the sample data, we make a decision about what distribution of treatments t we will assign in the population. Denote this distribution by $\delta(t \mid \omega)$. This $\delta(\cdot \mid \omega)$ is a density function on $\mathcal{T}$ for all ω. That is,

$$\int_{\mathcal{T}} \delta(t \mid \omega) \, d\nu(t) = 1,$$

where ν is a σ-finite measure on $\mathcal{T}$. Call $\delta(\cdot \mid \cdot) \in \mathcal{D}$ a *statistical treatment rule*.

Let Q_θ denote the sampling distribution of the data ω when the true state of the world is θ. Following Wald (1950), we evaluate statistical treatment rules according to their mean performance across repeated sampling. Performance is measured using a function $L(\delta, \theta)$, called a *loss function*. The mean loss of a rule δ is called the *risk*:

$$R(\delta, \theta) = \int_{\Omega} L(\delta(\cdot \mid \omega), \theta) \, dQ_\theta(\omega).$$

Although many loss functions may be considered, I focus on *regret loss*, defined as follows. For a given dataset ω, the rule δ yields the welfare

$$W[\delta(\cdot \mid \omega), \theta] = \int_{\mathcal{T}} \delta(t \mid \omega) w(\theta_t) \, d\nu(t).$$

Define the *regret* from the rule δ at state θ to be

$$\text{Regret}[\delta(\cdot \mid \omega), \theta] = U_\theta^* - W[\delta(\cdot \mid \omega), \theta],$$

where

$$U_\theta^* = \sup_{t \in \mathcal{T}} w(\theta_t)$$

is the maximal welfare if we knew the state of the world was θ. Then the risk of δ under regret loss is

$$
\begin{aligned}
R(\delta, \theta) &= \int_\Omega L(\delta(\cdot \mid \omega), \theta) \, dQ_\theta(\omega) \\
&= \int_\Omega \text{Regret}[\delta(\cdot \mid \omega), \theta] \, dQ_\theta(\omega) \\
&= U_\theta^* - \int_\Omega W[\delta(\cdot \mid \omega), \theta] \, dQ_\theta(\omega) \\
&= U_\theta^* - \int_\Omega \left\{ \int_\mathcal{T} \delta(t \mid \omega) w(\theta_t) \, d\nu(t) \right\} dQ_\theta(\omega) \\
&= U_\theta^* - \int_\mathcal{T} \mathbb{E}_\omega[\delta(t \mid \omega)] w(\theta_t) \, d\nu(t).
\end{aligned}
$$

Since θ is unknown, the risk $R(\delta, \theta)$ cannot be used directly to evaluate statistical treatment rules. Two common ways of eliminating θ are: (1) averaging risk over θ with respect to some distribution $\pi(\theta)$, or (2) looking at the worst case θ. I focus on worst case analysis. Since regret is bad, larger values of risk are bad. Hence the worst case risk is

$$\sup_\theta R[\delta, \theta].$$

Define a *finite sample minimax-regret treatment rule* δ^* as a solution to

$$\inf_\delta \sup_\theta R[\delta, \theta].$$

2.1.2 Finitely many treatments

Deriving finite sample minimax-regret treatment rules for an arbitrary set of treatments $\mathcal{T}$ is quite challenging. Most previous work has focused on the binary treatment case. In this paper, I consider the case where $\mathcal{T}$ is a finite set of K distinct treatments, $\mathcal{T} = \{t_1, \ldots, t_K\}$. In this case, $\delta(\cdot \mid \omega)$ is a probability mass function, with $\delta_k(\omega) \equiv \delta(t_k \mid \omega)$ denoting the percentage of the population to be assigned to treatment t_k given data ω. Note that $\sum_{k=1}^K \delta_k(\omega) = 1$ must hold for all ω. From here on, I suppress dependence on ω and just write δ_k instead.

For simplicity, I assume there is a unique solution to equation (2). That is, there is a unique optimal treatment. Under this assumption, the infeasible optimal treatment rule is

$$\delta_k^* = \mathbb{1}(w(\theta_k) > w(\theta_s) \text{ for all } s \neq k)$$

where I let $\theta_k = \theta_{t_k}$. It is helpful to rewrite the regret loss function using this infeasible optimal rule:

$$\text{Regret}[\delta, \theta] = \max\{w(\theta_1), \dots, w(\theta_K)\} - \sum_{k=1}^{K} w(\theta_k)\delta_k$$

$$= \sum_{k=1}^{K} w(\theta_k)[\mathbb{1}(w(\theta_k) > w(\theta_s) \text{ for all } s \neq k) - \delta_k].$$

Note that, when $K = 2$, this loss function simplifies to that in Hirano and Porter (2009). The risk is then

$$R[\delta, \theta] = \sum_{k=1}^{K} w(\theta_k)[\mathbb{1}(w(\theta_k) > w(\theta_s) \text{ for all } s \neq k) - \mathbb{E}\delta_k].$$

Next I specify the sampling process. Here I let $N = n_1 + \cdots + n_K$ denote the total number of observations.

Assumption 1 (Data generating process)

1. For each k we observe a random sample of size n_k of data from P_{θ_k}. This sample is independent of the other datasets.
2. $n_k/N \to \lambda_k \in (0, 1)$ as $N \to \infty$.

Throughout this paper I also assume the parameters θ_k are point identified. See Song (2014) for some related asymptotic results in the partially identified case

Assumption 2 (Identification) Each θ_k is point identified.

Assumptions 1 and 2 together let us use the data from sample k to consistently estimate θ_k.

2.2 Distribution of plug-in estimators of welfare under local alternatives

The purpose of this paper is to give conditions under which the plug-in rule

$$\delta_{k,N}^* = \begin{cases} 1/|\mathcal{M}_k| & \text{if } w(\hat{\theta}_{k,N}) \geq w(\hat{\theta}_{s,N}) \text{ for all } s \neq k \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

is locally asymptotically minimax under regret loss, where $\hat{\theta}_{k,N}$ is a 'best regular' estimator of θ_k and $\mathcal{M}_k = \{s \in \{1, \dots, K\} : w(\hat{\theta}_{k,N}) = w(\hat{\theta}_{s,N})\}$. Many treatment rules will be consistent, in the sense that they asymptotically select the optimal treatment:

$$\lim_{N \to \infty} \mathbb{E}\delta_{k,N} = \begin{cases} 1 \text{ if } k = \underset{k}{\operatorname{argmax}} w(\theta_k) \\ 0 \text{ otherwise.} \end{cases}$$

Local asymptotic theory allows us to more finely distinguish between any two treatment rules, by considering a sequence of parameter values such that the best treatment is not clear, even asymptotically. This 'local sequence' prevents the decision problem from becoming trivial asymptotically. To this end, I consider parameter sequences of the form

$$\left(\theta_0 + \frac{h_1}{\sqrt{N}}, \dots, \theta_0 + \frac{h_K}{\sqrt{N}} \right), \qquad h = (h_1, \dots, h_K), \tag{3}$$

where θ_0 is, without loss of generality, such that $w(\theta_0) = 0$.[4] Thus, under this sequence, all treatments are eventually equivalent (since w is continuous by assumption 3 below). The parameters h_k are called local parameters.

Assumption 3 (Model regularity)

1. For each k, $\theta_k \in \Theta$ where Θ is an open subset of the Euclidean space $\mathbb{R}^{d_\theta}$. $\theta_0 \in \Theta$ satisfies $w(\theta_0) = 0$.
2. The class $\{P_\theta : \theta \in \Theta\}$ is differentiable in quadratic mean (QMD): There exists a vector of measurable functions $\dot{\ell}_\theta$, called the score functions, such that

$$\int \left[\sqrt{p_{\theta+h}} - \sqrt{p_\theta} - \frac{1}{2}h'\dot{\ell}_\theta \sqrt{p_\theta} \right]^2 d\mu = o(\|h\|^2)$$

   as $h \to 0$, where p_θ denotes the density of P_θ with respect to the measure μ. Let the information matrix $I_0 = \mathbb{E}_{\theta_0}[\dot{\ell}_{\theta_0} \dot{\ell}'_{\theta_0}]$ be nonsingular.
3. $w(\theta)$ is continuous, and is differentiable at θ_0.

It is not necessary that the distribution of outcomes under each treatment lies in the same parametric family, or that θ_k all have the same dimensions, but it simplifies the exposition to do so. This generality follows from assumption 1.1, which says that all the samples are jointly independent, and the fact that the joint distribution of independent normals is the multivariate normal distribution. Also see remark 1 below.

Let $\hat{\theta}_{k,N}$ be a best regular estimator of θ_k, meaning that

---

[4] This can be achieved by simply adding the same constant to all welfare values, which will not affect the decision problem.

$$\sqrt{N}\left(\hat{\theta}_{k,N} - \left[\theta_0 + \frac{h_k}{\sqrt{N}}\right]\right) \overset{h_k}{\rightsquigarrow} \mathcal{N}(0, \lambda_k^{-1} I_0^{-1}) \tag{4}$$

for all h_k. For example, $\hat{\theta}_{k,N}$ can be the MLE of θ_k. Let $\hat{\boldsymbol{\theta}}_N = (\hat{\theta}'_{1,N}, \dots, \hat{\theta}'_{K,N})'$, $\mathbf{w}(\hat{\boldsymbol{\theta}}_N) = (w(\hat{\theta}_{1,N}), \dots, w(\hat{\theta}_{K,N}))'$,

$$\dot{w} = \frac{\partial w}{\partial \theta}\Big|_{\theta=\theta_0}, \quad \dot{\mathbf{w}}' = \begin{pmatrix} \dot{w}' & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \dot{w}' \end{pmatrix}, \quad \text{and} \quad \mathbf{I}_0 = \begin{pmatrix} \lambda_1 I_0 & 0 & \cdots & 0 \\ 0 & \lambda_2 I_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K I_0 \end{pmatrix}.$$

Then we have the following result on the distribution of plug-in estimators of welfare under the sequence of local alternatives (3).

Proposition 1 *Suppose assumptions* 1, 2, *and* 3 *hold. Then, for every h,*

$$\sqrt{N}\mathbf{w}(\hat{\boldsymbol{\theta}}_N) \overset{h}{\rightsquigarrow} \mathcal{N}(\dot{\mathbf{w}}'h, \dot{\mathbf{w}}'\mathbf{I}_0^{-1}\dot{\mathbf{w}}),$$

where $\hat{\theta}_{k,N}$ are best regular estimators of $\theta_{k,N}$.

This proof, along with all others, is given in appendix A. It follows by applying Le Cam's third lemma to the asymptotic linear representations of $\sqrt{N}\mathbf{w}(\hat{\boldsymbol{\theta}}_N)$ (obtained via the delta method and since $\hat{\boldsymbol{\theta}}_N$ is best regular) and the log-likelihood ratio (which satisfies local asymptotic normality due to the QMD assumption).

In particular, this proposition gives

$$\sqrt{N}w(\hat{\theta}_{k,N}) \overset{h_k}{\rightsquigarrow} \mathcal{N}(\dot{w}'h_k, \lambda_k^{-1}\dot{w}'I_0^{-1}\dot{w})$$

for each k. It's important that we have obtained the asymptotic distribution under the sequence of local alternatives, and not under the fixed 'true' distribution P_{θ_0}. Indeed, as used in the proof of this proposition, standard asymptotic theory shows that $\sqrt{N}w(\hat{\theta}_{k,N})$ converges to a normal distribution centered at zero under P_{θ_0}. Proposition 1 is analogous to lemma 3 of Hirano and Porter (2009).

2.3 The asymptotic representation theorem

In the previous section I derived the limiting distribution of plug-in estimators of welfare under a sequence of local alternatives. In this section, I first scale the risk and loss functions to keep them nontrivial asymptotically. I then state an asymptotic representation theorem, which formalizes the notion that no sequence of treatment rules can be better than the best treatment rule in the limit experiment.

To prevent regret loss from going to zero asymptotically, I scale it by $\sqrt{N}$:

$$\sqrt{N} \text{ Regret}\left[\delta, \theta_0 + \frac{h}{\sqrt{N}}\right]$$

$$= \sum_{k=1}^{K} \sqrt{N}w\left(\theta_0 + \frac{h_k}{\sqrt{N}}\right)\left[\mathbb{1}\left(\sqrt{N}w\left(\theta_0 + \frac{h_k}{\sqrt{N}}\right) > \sqrt{N}w\left(\theta_0 + \frac{h_s}{\sqrt{N}}\right) \text{ for all } s \neq k\right) - \delta_k\right]$$

$$\rightarrow \sum_{k=1}^{K} \dot{w}'h_k[\mathbb{1}(\dot{w}'h_k > \dot{w}'h_s \text{ for all } s \neq k) - \delta_k] \qquad \text{as } N \rightarrow \infty$$

$$\equiv L_\infty(\delta, h),$$

where the third line follows since

$$\sqrt{N}w\left(\theta_0 + \frac{h_k}{\sqrt{N}}\right) \rightarrow \dot{w}'h_k \quad \text{as } N \rightarrow \infty,$$

which holds by a Taylor expansion and since $w(\theta_0) = 0$.

Scaled finite sample risk under the local sequence is

$$\sqrt{N}R_N\left[\delta_N, \theta_0 + \frac{h}{\sqrt{N}}\right]$$

$$= \sqrt{N}\mathbb{E}_{\theta_0 + h/\sqrt{N}} \text{ Regret}\left[\delta_N, \theta_0 + \frac{h}{\sqrt{N}}\right]$$

$$= \sum_{k=1}^{K} \sqrt{N}w\left(\theta_0 + \frac{h_k}{\sqrt{N}}\right)\left[\mathbb{1}\left(\sqrt{N}w\left(\theta_0 + \frac{h_k}{\sqrt{N}}\right) > \sqrt{N}w\left(\theta_0 + \frac{h_s}{\sqrt{N}}\right) \text{ for all } s \neq k\right) - \beta_{k,N}(h)\right],$$

where I have defined

$$\beta_{k,N}(h) = \mathbb{E}_{\theta_0 + h/\sqrt{N}}\delta_{k,N}.$$

Assumption 4 (Pointwise convergence) The rule δ_N is such that for each component k and each h, $\beta_{k,N}(h)$ converges to some limit $\beta_k(h)$.

Under this pointwise convergence assumption, scaled finite sample risk converges to asymptotic risk as follows:

$$\sqrt{N}R_N\left[\delta_N, \theta_0 + \frac{h}{\sqrt{N}}\right] \rightarrow \sum_{k=1}^{K} \dot{w}'h_k[\mathbb{1}(\dot{w}'h_k > \dot{w}'h_s \text{ for all } s \neq k) - \beta_k(h)].$$

Theorem 1 (Asymptotic representation theorem) *Suppose assumptions* 1–4 *hold. Then for each k there exists a function* $\delta_k : \mathbb{R}^{\dim(h_k)} \rightarrow [0, 1]$ *such that for every* h_k,

$$\beta_k(h) = \int \delta_k(\Delta_k) \, d\mathcal{N}(\Delta_k \mid h_k, \lambda_k^{-1}I_0^{-1}),$$

and $\sum_{k=1}^{K} \delta_k(\Delta_k) = 1$ *for all* $\Delta = (\Delta_1, \dots, \Delta_K)$.

Assumption 3 implies that $\{P_{\theta_1}^{n_1} \otimes \cdots \otimes P_{\theta_K}^{n_K} : \theta \in \Theta^K\}$ converges to the limit experiment $\{\mathcal{N}(h_1, \lambda_1^{-1}I_0^{-1}) \otimes \cdots \otimes \mathcal{N}(h_K, \lambda_K^{-1}I_0^{-1})\}$ (see Van der Vaart (1998) chapter 9 for a formal discussion of convergence of experiments). The asymptotic representation theorem states that for any rule $\delta_N = (\delta_{1,N}, \ldots, \delta_{K,N})$ which has a limit in the sense of assumption 4, there exists a rule $\delta = (\delta_1, \ldots, \delta_K)$ in the limit experiment whose risk $R_\infty(\delta, h)$ equals the limiting risk of δ_N. We say that δ_N is *matched* by δ in the limit experiment. This theorem is a special case of theorem 9.3 on page 127 of Van der Vaart (1998) (see also theorem 15.1 on page 215 and proposition 7.10 on page 98 for similar special cases), and hence its proof is omitted.

2.4 The optimal treatment rule in the limit experiment

Because of the asymptotic representation theorem, no rule can do better than the best rule in the limit experiment, which I derive in this section. In the limit experiment, $\{\mathcal{N}(h_1, \lambda_1^{-1}I_0^{-1}) \otimes \cdots \otimes \mathcal{N}(h_K, \lambda_K^{-1}I_0^{-1})\}$, we observe a single draw $\Delta = (\Delta_1', \ldots, \Delta_K')'$ from the distribution

$$\mathcal{N}\left(\begin{pmatrix} h_1 \\ \vdots \\ h_K \end{pmatrix}, \begin{pmatrix} \lambda_1^{-1}I_0^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_K^{-1}I_0^{-1} \end{pmatrix}\right).$$

The risk of a rule δ in this limit experiment is

$$R_\infty(\delta, h) \equiv \sum_{k=1}^{K} \dot{w}'h_k[\mathbb{1}(\dot{w}'h_k > \dot{w}'h_s \text{ for all } s \neq k) - \beta_k(h)]$$

where

$$\beta_k(h) = \int \delta_k(\Delta_k) \, d\mathcal{N}(\Delta_k \mid h_k, \lambda_k^{-1}I_0^{-1}).$$

A rule δ^* is minimax optimal in this experiment if it solves

$$\inf_\delta \sup_h R_\infty(\delta, h).$$

Because of the form of the limit risk R_∞, the (infeasible) optimal choice of treatments in the limit experiment is

$$\delta_k = \begin{cases} 1 & \text{if } \dot{w}'h_k > \dot{w}'h_s \text{ for all } s \neq k \\ 0 & \text{otherwise.} \end{cases}$$

In the main result of this section, I show that if we replace the unknown mean h_k by the observed realization Δ_k, we obtain a minimax optimal treatment rule:

$$\delta_k^* = \begin{cases} 1 & \text{if } \dot{w}'\Delta_k > \dot{w}'\Delta_s \text{ for all } s \neq k \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Because the limit risk R_∞ only depends on h through the linear combinations $\dot{w}'h_k$, for all k, the random variable $\dot{w}'\Delta$ is a sufficient statistic in the following sense.

Lemma 1 *For any rule $\delta_k(\Delta)$ which is a function of the entire vector Δ, there exists a rule $\tilde{\delta}_k(\dot{w}'\Delta)$ which is a function of only $\dot{w}'\Delta$, and yet achieves the same risk as δ_k.*

This result is a kind of complete class theorem. Hence it suffices to consider the limit experiment where we observe a single draw $\dot{w}'\Delta$ from the distribution

$$\mathcal{N}\left(\begin{pmatrix} \dot{w}'h_1 \\ \vdots \\ \dot{w}'h_K \end{pmatrix}, \begin{pmatrix} \lambda_1^{-1}\dot{w}'I_0^{-1}\dot{w} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1}\dot{w}'I_0^{-1}\dot{w} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K^{-1}\dot{w}'I_0^{-1}\dot{w} \end{pmatrix}\right).$$

The following assumption states that, asymptotically, the sample sizes are equal.

Assumption 5 (Asymptotically balanced samples) $\lambda_1 = \cdots = \lambda_K = 1/K$.

This assumption implies that the only differences between treatments in the limit experiment are their means h_k. Consequently, the question of finding the optimal treatment rule is simply that of finding the optimal rule when the goal is to pick the normal population with the largest mean, when all populations have equal variance. This problem was solved in a series of papers by Bahadur (1950), Bahadur and Goodman (1952), and Lehmann (1966).

Theorem 2 *The rule δ^* defined in equation* (5) *is minimax optimal in the limit experiment*:

$$\sup_h R_\infty(\delta^*, h) = \inf_\delta \sup_h R_\infty(\delta, h).$$

This result follows immediately from the above discussion and the results of Bahadur, Goodman, and Lehmann, which I discuss in appendix B. Their results use permutation invariance arguments. This is quite different from the approach in Hirano and Porter (2009), who apply results from hypothesis testing. In particular, they use the Neyman-Pearson lemma; see Van der Vaart (1998) proposition 15.2 on page 217. The hypothesis testing approach does not appear to generalize to the case with more than two treatments.

Theorem 2 relies on assumption 5 to ensure equal variances in the limit. Practically, this means that rules which approximate δ^* (see section 2.5) can only be guaranteed to be optimal when the sample sizes in finite samples are roughly equal. Such rules may have poor finite sample performance when the sample sizes are dramatically different. Relaxing this assumption has proven to be quite difficult analytically. In the binary treatment case, Hirano and Porter (2009) do not require an assumption like 5. Nonetheless, their main results are similar to theorem 2, in that they also show that an empirical success rule is asymptotically optimal. This empirical success rule may also have poor finite sample performance

when sample sizes are dramatically different (see section 3), and hence this calls into question the value of the local asymptotic approximation for these cases.[5]

Remark 1 As mentioned earlier, the results generalize to allow different parametric models across treatments. In this case, assumption 5 must be modified to require the variances $\lambda_k^{-1} \dot{w}_k' I_{0,k}^{-1} \dot{w}_k$ to be equal for all k, where $I_{0,k}$ is the information matrix corresponding to the kth treatment, evaluated at the centering point $\theta_{0,k}$, and $w_k(\theta_k) = W(P_k(\theta_k))$.

2.5 Local asymptotic minimaxity of the plug-in rule

Theorem 2 shows that the optimal decision rule in the limit experiment is

$$\delta_k^* = \begin{cases} 1 & \text{if } \dot{w}' \Delta_k > \dot{w}' \Delta_s \text{ for all } s \neq k \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Proposition 1 shows that $\sqrt{N} w(\hat{\theta}_{k,N})$ has an asymptotic normal distribution with mean $\dot{w}' h_k$ under the sequence of local alternatives; i.e., the same distribution as $\dot{w}' \Delta_k$. This suggests that an optimal rule might be obtained by replacing $\dot{w}' \Delta_k$ with $\sqrt{N} w(\hat{\theta}_{k,N})$ in δ^*. In this section, I show that this plug-in rule,

$$\delta_{k,N}^* = \begin{cases} 1/|\mathcal{M}_k| & \text{if } w(\hat{\theta}_{k,N}) \geq w(\hat{\theta}_{s,N}) \text{ for all } s \neq k \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

matches the optimal rule in the limit experiment, and hence that this plug-in rule is locally asymptotically minimax under regret loss.

Let $\mathcal{D}$ denote the set of all sequences of rules δ_N that converge in the sense of assumption 4. The following result shows that the minimax value in the limit experiment is an asymptotic risk bound.

Lemma 2 (Asymptotic minimax bound) *Suppose assumptions* 1–5 *hold. Then for all* $\delta_N \in \mathcal{D}$,

$$\sup_J \liminf_{N \to \infty} \sup_{h \in J} \sqrt{N} R_N \left[\delta_N, \theta_0 + \frac{h}{\sqrt{N}} \right] \geq \inf_\delta \sup_h R_\infty(\delta, h),$$

where the outer supremum over J is taken over all finite subsets J of $\mathbb{R}^{\dim(h)}$.

We call any δ_N which achieves the lower bound a *locally asymptotically minimax* rule. The following theorem is the main asymptotic result of this paper.

Theorem 3 *Suppose assumptions* 1–5 *hold. Let* $\hat{\theta}_N$ *be a best regular estimator*, *as described by equation* (4). *Then the plug-in rule* δ_N^* *defined in* (1) *is locally asymptotically minimax*:

$$\sup_J \liminf_{N \to \infty} \sup_{h \in J} \sqrt{N} R_N \left[\delta_N^*, \theta_0 + \frac{h}{\sqrt{N}} \right] = \inf_\delta \sup_h R_\infty(\delta, h),$$

where the outer supremum over J is taken over all finite subsets J of $\mathbb{R}^{\dim(h)}$.

Compared to the finite sample results discussed in appendix B, this asymptotic result has several advantages. It does not require each treatment distribution to lie in the same parametric class (although I have assumed this in the exposition for simplicity). It does not require sample sizes to be exactly balanced, although the approximation will likely be poor if the sample size is far from being balanced. It does not require the parameters θ_k to be scalar, and does not require the distribution of data to have monotone likelihood ratio.

2.6 Discussion

In this section, I have shown that, when there are a finite number of treatments, the rule which assigns everyone to the treatment with the largest estimated welfare is locally asymptotically minimax under regret loss. This extends one of the results in Hirano and Porter (2009), and relies on applying permutation invariance results by Bahadur (1950), Bahadur and Goodman (1952), and Lehmann (1966) in the limit experiment, instead of results from hypothesis testing. One limitation of this result is that I required the sample sizes to be asymptotically balanced, which was not required by Hirano and Porter when there are only two treatments. This requirement suggests that the empirical success rule defined in equation (1) may perform poorly when the sample size is far from balanced. More generally, the performance of the rule (1) depends on how well the limit experiment approximates the actual finite-sample distribution of $\hat{\theta}_N$. I leave a further exploration of the quality of the finite-sample approximation to future research. Likewise, I leave the problem of asymptotically unbalanced samples to future research.

3 Numerical computation of finite-sample minimax-regret treatment rules

Finite-sample minimax-regret treatment rules are often difficult to derive analytically. When analytical results are not available, we can instead numerically compare the performance of any proposed treatment rules, and also compute optimal treatment rules. In this section, I first compare selected treatment rules. I show that optimal rules for certain sampling designs may perform quite poorly for other designs. In particular, the empirical success rule does poorly for unbalanced designs. I next show how optimal

rules can be computed. Minimax-regret rules can be thought of as solutions to a fictitious game between nature and the decision-maker. Consequently, numerical techniques used to compute equilibria of games can be used to compute minimax-regret rules.

3.1 Setup

The general setup is as in section 2. Here I describe several additional assumptions I use for the numerical finite-sample results, and I describe the sampling designs I consider throughout the section.

3.1.1 Additional assumptions

Assume the welfare function is the population mean outcome: $W(P_{\theta_t}) = \mathbb{E}_{\theta_t}[Y_i(t)] \equiv m_\theta(t)$, where $\theta = \{\theta_t : t \in \mathcal{T}\}$. Assume outcomes have an arbitrary distribution on a common bounded support, which is normalized to $[0, 1]$. Using Schlag's (2003) 'binary randomization' technique, described below, it suffices to only consider the case where outcomes are binary, $Y_i(t) \in \mathcal{Y} = \{0, 1\}$, where 1 is a success and 0 is a failure. Under this assumption, the population mean function is just the proportion of individuals in the population who achieve a success, $m_\theta(t) = \mathbb{P}[Y_i(t) = 1]$. Since outcomes are binary, the state of the world is fully described by the vector $\theta = (\mu_1, \dots, \mu_K)$ of means $\mu_k = m_\theta(t_k)$. Assume Θ is the product space $[0, 1]^K$.

Our sample data ω has the form $\{(Y_i, k_i)\}_{i=1}^N$. Let N_k denote the number of observations with treatment k. Each Y_i is an independent draw from the Bernoulli distribution with mean μ_{k_i}. Let n_k denote the number of successes among subjects with treatment k. If the number of observations N_k of each treatment is non-random, the vector of success counts $(n_1, \dots, n_K)$ is a sufficient statistic for the sample data, and so from now on we view ω as producing this vector. If the number of observations N_k is random, then we augment the vector of success counts with the realized number of observations of each treatment.

Remark 2 Schlag (2003) showed that by performing a 'binary randomization' it suffices to only consider binary outcomes. This technique is described as follows. As above, we assume outcomes have an arbitrary distribution with a common bounded support, normalized to $[0, 1]$. Replace each observation $Y_i \in [0, 1]$ by $\tilde{Y}_i \in \{0, 1\}$, obtained by making a single draw of a Bernoulli(Y_i) random variable. Now apply a treatment rule $\tilde{\delta}$ to the binary data $\tilde{Y}_i$. Let δ denote the overall rule, including both the binary randomization step and the application of $\tilde{\delta}$. It turns out that if $\tilde{\delta}$ is minimax-regret optimal for binary data, then δ is minimax-regret optimal for the original data.

3.1.2 Sampling designs

Thus far I have not fully described how the data ω is gathered. I consider two kinds of sampling designs. For simplicity, I describe them in terms of just two treatments. Call the pair (N_1, N_2) of sample sizes an *allocation*.

- **Ex ante known allocation**. The researcher chooses N_1 and N_2. Among all the possible combinations of subjects which achieve this allocation, one of them is chosen at random with equal probability. When $N_1 = N_2$, we say the design is *balanced*. Otherwise, it is *unbalanced*.[6]
- **Ex ante unknown allocation**. The researcher chooses the total sample size $N = N_1 + N_2$. Individuals are independently assigned to treatment 1 or 2 with *equal* probability. Thus, before performing the experiment, any pair (N_1, N_2) such that $N_1 + N_2 = N$ is possible. In particular, it is possible that all subjects are assigned to the same group and we make no observations in the other group. For this design, I assume the decision-maker commits to a decision rule δ before the allocation (N_1, N_2) is revealed.[7]

The ex ante unknown allocation is easy to implement, but it may lead to extremely unbalanced samples ex post. This feature makes it intuitively unattractive to many researchers, who consequently prefer an ex ante known balanced design. On the other hand, traditional design of experiments based on analysis of power often recommends an ex ante known unbalanced design.[8] For example, suppose there are two treatments with normally distributed outcomes and known variances but unknown means. Then the power of a *t*-test for a difference in the means is maximized by making more observations of the treatment with a larger variance. In section 3.2, I consider optimal sampling designs based instead on minimizing maximum regret.

3.2 Comparison of selected rules

In this section, I consider several cases where the analytical optimal rule is unknown. I compare several 'reasonable' rules by numerically calculating the maximal regret associated with each rule. These are the empirical success rule, various Bayes' rules, and Stoye's (2009a) rule, which is minimax-regret optimal for the ex ante unknown allocation and two treatments.

I begin with ex ante known allocations with an unbalanced design and two treatments. The main findings are that the empirical success rule is not optimal for unbalanced designs, and that balanced designs are preferred to unbalanced designs. Moreover, the empirical success rule does worse as the design becomes even more unbalanced, and its maximal regret may actually increase with the sample size. The same results hold for three treatments. With ex ante unknown allocations and three treatments, the obvious extension of Stoye's rule is not optimal. Finally, my

---

[6] Note that a balanced allocation requires a total sample size that is a multiple of the number of treatments, K.

[7] Without commitment, we are back in the situation of looking at decision rules for an ex ante known allocation.

[8] For example, see List et al. (2011).

calculations here suggest that an ex ante unknown design is preferred to any ex ante known design.

3.2.1 Ex ante known allocation, unbalanced design

Suppose there are two treatments with sample sizes $N_1 \neq N_2$. I consider three different treatment rules.[9] All rules have the form

$$\delta_2 = \begin{cases} 1 & \text{if } I_{21} > 0 \\ 1/2 & \text{if } I_{21} = 0 \\ 0 & \text{if } I_{21} < 0, \end{cases} \qquad \delta_1 = 1 - \delta_2.$$

Define the comparison numbers I_{21} as follows.

- **Empirical success**: Define

$$I_{ij}^{\text{ES}} = \frac{n_i}{N_i} - \frac{n_j}{N_j}.$$

- **Stoye's rule**: Define

$$
\begin{aligned}
I_{ij} &= n_i - n_j - \frac{N_i - N_j}{2} \\
&= N_i \left(\frac{n_i}{N_i} - \frac{1}{2} \right) - N_j \left(\frac{n_j}{N_j} - \frac{1}{2} \right).
\end{aligned}
$$

  Note that when $N_j = 0$, we choose i if and only if the sample success proportion n_i/N_i is greater than 1/2. Thus 1/2 is the a priori mean.
- **Squared error minimax**: Define

$$I_{ij}^{\text{M}} = \frac{(1/2)\sqrt{N_i} + n_i}{\sqrt{N_i} + N_i} - \frac{(1/2)\sqrt{N_j} + n_j}{\sqrt{N_j} + N_j}.$$

Note that all rules are equivalent when $N_1 = N_2$.[10] "Stoye's rule" is the minimax-regret optimal rule for $K = 2$ and an ex ante unknown allocation, as shown by Stoye (2009a). The squared error minimax rule is derived as follows. Consider group 1. We observe $n_1 \sim \text{Bin}(\mu_1, N_1)$. With a Beta($\alpha, \beta$) prior on μ_1, the Bayes estimator of μ_1 under squared error loss is

---

[9] Another rule, suggested by Schlag (2006), is to randomly drop data until sample sizes are equal. I do not consider this rule here.

[10] This case, the ex ante known allocation with a balanced design, is what Stoye (2009a) called "matched pairs." In this case, his proposition 1(i) shows that the empirical success rule is finite sample minimax-regret optimal, when outcomes are binary. His proposition 2 then extends this result to outcomes with bounded support by using the the binary randomization technique. Here my focus is on unbalanced designs, in which case the finite sample optimal rule is unknown.

Table 1 Maximal regret for **Stoye's rule**. Row is N_1, Column is N_2

	1	2	3	4	5	10	15	20
1	0.125	0.148	0.174	0.198	0.209	0.263	0.288	0.309
2	0.148	0.087	0.105	0.134	0.160	0.226	0.266	0.287
3	0.174	0.105	0.071	0.082	0.109	0.197	0.239	0.270
4	0.198	0.134	0.082	0.061	0.067	0.163	0.218	0.249
5	0.209	0.160	0.109	0.067	0.054	0.136	0.193	0.233
10	0.263	0.226	0.197	0.163	0.136	0.038	0.091	0.143

Table 2 Maximal regret for **Empirical success rule**

	1	2	3	4	5	10	15	20
1	0.125	0.125	0.129	0.136	0.143	0.167	0.182	0.192
2	0.125	0.087	0.074	0.077	0.078	0.090	0.099	0.105
3	0.129	0.074	0.071	0.064	0.064	0.060	0.066	0.071
4	0.136	0.077	0.064	0.061	0.058	0.051	0.052	0.052
5	0.143	0.078	0.064	0.058	0.054	0.047	0.044	0.043
10	0.167	0.090	0.060	0.051	0.047	0.038	0.035	0.033

$$\hat{\mu}_1^{B} = \frac{\alpha + n_1}{\alpha + \beta + N_1}$$
$$= \left(\frac{\alpha + \beta}{\alpha + \beta + N_1} \right) \frac{\alpha}{\alpha + \beta} + \left(\frac{N_1}{\alpha + \beta + N_1} \right) \frac{n_1}{N_1}.$$

Rule δ_1^{M} uses the prior $\alpha = \beta = (1/2)\sqrt{N_1}$ for treatment 1 and $\alpha = \beta = (1/2)\sqrt{N_2}$ for treatment 2.[11] This corresponds to the minimax optimal rule for estimating μ_1 and μ_2 separately, each under squared error loss. Thus, in rule δ^{M}, we first estimate μ_1 and μ_2 separately by their Bayes' estimators under squared error loss and a particular prior. These estimators are just the posterior means, a consequence of using squared error loss. The rule then picks the treatment with the largest Bayes' estimator. The Bayes' estimator biases the sample proportion toward 1/2, the a priori mean. For example, if $N_1 = 1$,

$$\hat{\mu}_1^{B}(n_1) = \begin{cases} 1/4 & \text{if } n_1 = 0 \\ 3/4 & \text{if } n_1 = 1. \end{cases}$$

For large N_1, the Bayes estimators are approximately equal to the sample proportions; $\hat{\mu}_1^{B} \approx n_1/N_1$. Despite the fact that the beta-prior Bayes estimators were derived for a different purpose, they lead to treatment rules that often perform better than the

---

[11] I also considered the uniform prior, $\alpha = \beta = 1$, and the Jeffreys prior, $\alpha = \beta = 1/2$. These rules tend to perform slightly worse than δ^{M}, but better than empirical success or Stoye's rule.

Table 3 Maximal regret for **Squared error minimax rule**	1	2	3	4	5	10	15	20
1	0.125	0.148	0.105	0.097	0.108	0.106	0.108	0.111
2	0.148	0.087	0.105	0.094	0.067	0.082	0.064	0.071
3	0.105	0.105	0.071	0.082	0.067	0.054	0.052	0.068
4	0.097	0.094	0.082	0.061	0.067	0.057	0.048	0.045
5	0.108	0.067	0.067	0.067	0.054	0.051	0.045	0.043
10	0.106	0.082	0.054	0.057	0.051	0.038	0.036	0.036

empirical success rule.[12] Tables 1, 2, and 3 show maximal regret for various sample sizes.

Among the rules considered, Stoye's rule performs the worst. This outcome underlines the importance of the sampling design for evaluating optimality. Stoye's rule is optimal if we commit to it before seeing the allocation, but it is not optimal if we allow ourselves to condition on the realized allocation. The empirical success rule performs better, but the beta-prior Bayes rule typically performs best. In particular, it improves upon the empirical success rule significantly more as the the allocation becomes more unbalanced; that is, as $N_1 - N_2$ gets large. This latter fact follows since the maximal regret of the empirical success rule is *increasing* in N_2, when N_1 is fixed (likewise if we fix N_2). In fact, most of the rules have this property (including the uniform prior and Jeffreys prior Bayes rules, not shown). Only the squared error minimax rule seems to mostly escape it. This result is related to the fact that the probability the empirical success rule makes a correct selection is decreasing in the sample size of the best population.[13] Thus, neither Stoye's rule nor the empirical success rule are minimax-regret optimal with ex ante known unbalanced allocations.[14] The apparently good performance of the beta-prior Bayes rules suggests the true optimal rule may look similar.

These results also shed light on Manski's example from Stoye (2009a). Suppose $N = 1100$, $N_1 = 1000$, $n_1 = 550$, $N_2 = 100$, $n_2 = 99$. Then Stoye's rule chooses treatment 1. The empirical success rule and the beta-prior Bayes rule both choose treatment 2. Thus the counterintuitive result that Stoye's rule chooses treatment 1 with this data may just reflect the fact that Stoye's rule is not optimal for an ex ante known unbalanced allocation.

Tables 1, 2, and 3 allow us to examine optimal sample size allocation. Equal sample size allocations are essentially always preferred to unequal allocations. For example, suppose $N = 6$. Compare the allocation $N_1 = 5$, $N_2 = 1$ to the allocation $N_1 = 3$, $N_2 = 3$. Maximal regret for empirical success is cut in half, from 0.143 to

---

[12] Abughalous and Miescke (1989) also discuss Bayes rules in this two treatment, unbalanced allocation setting, under 0-1 loss and monotone permutation invariance loss.

[13] For example, see Tong and Wetzell (1979) and also Bofinger (1985).

[14] This does not mean that they are not admissible, however. Note for example when $N_1 = 3$, $N_2 = 2$, the empirical success rule is the best of the five. For normal data, Miescke and Park (1999) prove that the empirical success rule is admissible for regret loss. I am not aware of a proof for binary outcomes, however.

0.071. Maximal regret for Stoye's rule goes from 0.209 to 0.071. Maximal regret for the squared error minimax rule goes from 0.108 to 0.071. For an equal allocation, $N_1 = N_2$, increasing total sample size N always lowers maximal regret. For $N = 10$, all rules deliver a regret of 5.5%. For $N = 11$, regret is below 5%.[15] The sampling cost and importance of decreasing regret by a small percent will lead to an optimal total sample size.

Next suppose there are three treatments. As in the two treatment case, I computed maximal regret for the three-treatment generalization of the three rules considered above. All rules have the form

$$\delta_1 = \begin{cases} 1 & \text{if } I_{12} > 0, I_{13} > 0 \\ 1/2 & \text{if } I_{12} > 0, I_{13} = 0 \\ 1/2 & \text{if } I_{12} = 0, I_{13} > 0 \\ 1/3 & \text{if } I_{12} = 0, I_{13} = 0 \\ 0 & \text{if } I_{12} < 0 \text{ or } I_{13} < 0 \end{cases}$$

for some pairwise comparison numbers I_{ij} defined as in the previous section on two treatments. δ_2 and δ_3 are defined analogously. Since Bayes rules form a complete class, and Bayes rules can be defined in terms of pairwise comparisons, we can restrict attention to rules with the form above.

Table 4 shows the maximal regret for various allocations. The findings here are similar to the two treatment case—the beta-prior Bayes rule does best, then the empirical success rule, and then Stoye's rule. Note, in particular, the poor performance of Stoye's rule when the sample sizes are most unequal. For example, when $N_3 = 1, N_1 = 8, N_2 = 10$.

3.2.2 Ex ante unknown allocations

Thus far I have only considered ex ante known allocations. In this section, I briefly consider ex ante unknown allocations. For these allocations, Stoye (2009a) derived the minimax-regret optimal rule for $K = 2$. I show that the obvious generalization of this rule is not minimax-regret optimal. Furthermore, I show that ex ante unknown allocations appear to be preferred to ex ante known allocations. Intuitively, committing to an allocation in advance gives nature an advantage in choosing her least favorable prior.

Table 5 shows maximal regret for Stoye's rule and the beta-prior Bayes rule using the Jeffreys prior, both described in the previous section. The Jeffreys Bayes rule beats Stoye's rule when $N = 4$ and $N = 6$, while Stoye's rule wins when $N = 8$ and $N = 10$. Although this shows that Stoye's rule is not optimal, its maximal regret is not much larger than that of the Jeffreys Bayes rule in the cases where it loses.

---

[15] This is Schlag's (2006) "Eleven tests needed for a recommendation" result. (By 'tests' he means 'observations'.)

Table 4 Maximal regret for three treatment rules when $K = 3$ under ex ante known allocations

Maximal regret for Stoye's rule, $N_3 = 1$. Row is N_1. Column is N_2

	1	5	10
1	0.21038	0.26136	0.29936
2	0.22425	0.27451	0.30632
4	0.25568	0.29727	0.32303
8	0.28856	0.32076	0.34099

Stoye's rule, $N_3 = 6$

	1	5	10
1	0.27477	0.31094	0.33341
2	0.28357	0.24675	0.282
4	0.30436	0.12883	0.19417
8	0.3262	0.12695	0.13582

Empirical success rule, $N_3 = 1$

	1	5	10
1	0.21038	0.23344	0.26903
2	0.21038	0.18586	0.21394
4	0.22385	0.17891	0.19081
8	0.25715	0.18844	0.19701

Empirical success rule, $N_3 = 6$

	1	5	10
1	0.24224	0.18394	0.19396
2	0.19292	0.10747	0.11161
4	0.18161	0.083565	0.077335
8	0.19035	0.076931	0.067696

Squared error minimax rule, $N_3 = 1$

	1	5	10
1	0.21038	0.17716	0.17596
2	0.22425	0.17374	0.17304
4	0.16702	0.14075	0.13733
8	0.17132	0.13505	0.13813

Squared error minimax rule, $N_3 = 6$

	1	5	10
1	0.19248	0.15101	0.14574
2	0.16659	0.10185	0.11062
4	0.14677	0.10185	0.077224
8	0.13941	0.077254	0.067684

Table 5 Maximal regret when the allocation is ex ante unknown and $K = 3$

Total sample size =	1	2	3	4	6	8	10
Stoye's rule	0.3333	0.2268	0.1860	0.1630	0.1335	0.1158	0.1037
Jeffreys Bayes rule	0.3333	0.2268	0.1860	0.1574	0.1324	0.1169	0.1050

When $K = 2$ and N is even,[16] the ex ante unknown allocation design and the ex ante known balanced design lead to identical values of regret.[17] Neither is preferred over the other. This conclusion no longer holds when $K = 3$. Tables 4 and 5 provide a counter-example. When $N = 3$, the ex ante known balanced allocation $N_1 = N_2 = N_3 = 1$ yields minimal maximal regret 0.21038 (due to the Bahadur–Goodman–Lehmann theorem; see appendix B). Stoye's rule under an ex ante unknown allocation, however, has a smaller maximal regret of 0.1860. The optimal rule for an ex ante unknown allocation may do even better than 0.1860.

When $K = 3$, ex ante unknown allocations also appear to perform better than ex ante known unbalanced allocations. For example, when $N = 4$, Stoye's rule under an ex ante unknown allocation has maximal regret 0.1630, while Stoye's rule under the ex ante known allocation $N_1 = 2, N_2 = 1, N_3 = 1$ has maximal regret 0.2243. Likewise, when $N = 6$, Stoye's rule under an ex ante unknown allocation has maximal regret 0.1335, while Stoye's rule under the ex ante known allocation $N_1 = 4, N_2 = 1, N_3 = 1$ has maximal regret 0.2557.

3.3 Computation of optimal rules

Although analytical finite-sample optimal treatment rules are often unknown, in this section I show how they can be computed numerically. I first consider a simple, but naive approach to numerically calculating optimal rules. I illustrate it by computing optimal rules for various unbalanced allocations. These calculations suggest that there exists an optimal rule which mixes at more than just one realization of the data, a distinct feature of minimax rules (e.g., (Manski, 2007b; Manski & Tetenov, 2007)). I next discuss a more sophisticated approach based on results from computational game theory.

Consider the two treatment case. With binary outcomes, the sample space is finite. Thus any treatment rule can be written as a finite vector specifying the fraction allocated to treatment 1 versus 2 for each element of the sample space. Specifically, the sample space for the sufficient statistics n_k is

---

[16] Required so that a perfect balanced design is possible.

[17] See proposition 6 of Stoye (2007a).

Table 6 Some minimax-regret optimal rules for unbalanced allocations

Column is n_2. Row is n_1. The cells are $\lambda^*_{n_1,n_2} = \mathbb{P}[\delta \text{ selects treatment 2} \mid \omega = (n_1, n_2)]$

$N_1 = 1, N_2 = 2$	0	1	2		
0	0.3343	1	1		
1	0	0	0.6657		

$N_1 = 2, N_2 = 3$	0	1	2	3	
0	0.494	1	1	1	
1	0	0	1	1	
2	0	0	0	0.506	

$N_1 = 1, N_2 = 3$	0	1	2	3	
0	0.1266	1	1	1	
1	0	0	0	0.8733	

$N_1 = 1, N_2 = 4$	0	1	2	3	4
0	0.0008	0.886	1	1	1
1	0	0	0	0.114	0.9992

$N_1 = 1, N_2 = 5$	0	1	2	3	4	5
0	0.1872	0.5668	1	1	1	1
1	0	0	0	0	0.4332	0.8128

$$\Omega = \{(0,0), (0,1), \ldots, (0, N_2),$$
$$(1,0), (1,1), \ldots, (1, N_2),$$
$$\vdots$$
$$(N_1, 0), (N_1, 1), \ldots, (N_1, N_2)\}.$$

Any treatment rule δ is completely defined by $(N_1 + 1) \cdot (N_2 + 1)$ constants

$$\lambda_{ij} \equiv \mathbb{P}[\delta \text{ selects treatment 2} \mid \omega = (i,j)].$$

The probability that δ selects treatment 2, prior to gathering the data, is

$$\mathbb{E}\delta_2 = \sum_{i,j} \lambda_{ij} \mathbb{P}(\omega = (i,j))$$
$$= \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} \lambda_{ij} \binom{N_1}{i} a^i (1-a)^{N_1 - i} \binom{N_2}{j} b^j (1-b)^{N_2 - j}, \quad (6)$$

where a and b are the unknown probabilities of success from treatments 1 and 2, respectively. Regret is

$$R(\delta, \theta) = R[\{\lambda_{ij}\}, (a,b)]$$
$$= \max\{a, b\} - (a\mathbb{E}\delta_1 + b\mathbb{E}\delta_2). \quad (7)$$

Table 7 Maximal regret for the optimal rule, along with several other rules

$N_1, N_2 =$	1,2	1,3	1,4	1,5	2,3
Stoye's rule	0.1482	0.1740	0.1975	0.2088	0.1055
Empirical success	0.1250	0.1293	0.1361	0.1428	0.0741
Squared error minimax	0.1482	0.1055	0.0971	0.1083	0.1055
Optimal	0.0986	0.0921	0.0918	0.0911	0.0732

The minimax-regret problem is

$$\min_{\lambda_{ij} \in [0,1]} \max_{(a,b) \in [0,1]^2} R[\{\lambda_{ij}\}, (a, b)].$$

For small sample sizes, we can solve the minimax-regret problem by using nonlinear optimization packages like KNITRO to solve a nested optimization problem, where we first solve the inner problem, and then solve the outer problem. I implemented this approach for several cases. The solutions are displayed in table 6. Note that there may not be a unique optimal rule, and therefore the numerical solutions are just one possible optimal rule.

These calculations suggest several things. The rules are monotonic in n_2 holding n_1 fixed (likewise if we swap n_1 and n_2). That is, for each fixed n_1, the probability that we choose treatment 2 is increasing in n_2. For the first three cases, mixing only occurs in the extreme cases where either there are no successes for both treatments, or there are only successes for both treatments. In the second two cases, when $N_1 = 1$ and $N_2 = 4$ or 5, the rule mixes for more than just those extreme cases.

Table 7 lists the value of maximal regret for the three rules considered previously, and the optimal rules computed above. Reassuringly, maximal regret is strictly decreasing in sample size, although the benefits of increasing sample size on a treatment which already has most of the observations are quite small.

Although these initial computations are helpful, more sophisticated numerical approaches have been developed for solving these kinds of computational problems. Schlag (2003, 2006) and Stoye (2009a) reconsidered Wald's (1945) game theory technique for deriving exact, analytical finite sample results. Stoye (2009a) describes this approach in detail (see also (Berger, 1985)), so I will only briefly review the main idea. Under conditions that are satisfied here, any minimax-regret rule δ^* is equivalent to a Bayes rule with respect to some prior π^* on Θ, called the *least favorable prior*. We envision a fictional game between the decision-maker, who must choose the rule δ^*, and nature, who chooses the prior π^*. It is well known that δ^* is a minimax-regret rule if (δ^*, π^*) is a Nash equilibrium of this game. Thus it suffices to derive a Nash equilibrium of this game. Schlag (2003, 2006) and Stoye (2009a) use this idea to derive analytical results, but these proofs often rely on symmetry properties of the setup, which make them difficult to extend to more general settings, such as allowing for many treatments with differing costs, or for unbalanced sampling designs. Rather than analytically finding equilibria, we can compute equilibria numerically, and hence numerically compute optimal treatment rules. I sketch the game to be solved numerically in chapter 2 of Masten (2013), and I discuss the

computational methods that can be used below. I leave actual implementation of these methods to future research, however.

Several papers ((Judd, Renner, and Schmedders, 2012; Kubler & Schmedders, 2010), and (Borkovsky, Doraszelski, and Kryukov, 2010)) discuss computational techniques for solving for equilibria of games with continuous strategies, with a particular emphasis on computing *all* equilibria, when there are multiple equilibria. This feature is particularly important in statistical decision theory: two player games often have multiple equilibria, and hence it is likely the game considered here will also have multiple equilibria. If there are many qualitatively different optimal rules, the minimax-regret criterion will suffer from the same problem that Bayes rules users face when choosing priors.[18] The presence of multiple minimax-regret treatment rules will require a method for choosing a particular rule.

These methods compute equilibria by solving the system of first order conditions for the two players subject to any required constraints (e.g., Judd (1998) pages 162–165). Consequently, these methods can be seen as simply numerical methods for solving systems of equations. The first approach is called all-solutions homotopy. The idea behind homotopy methods is to start with a system of equations $g(x) = 0$ whose solutions are known, and to then continuously transform that system into the system of interest $f(x) = 0$, whose solutions are desired. Assuming $g(x) = 0$ has as many solutions as $f(x) = 0$, each solution of $g(x) = 0$ will map into a solution of $f(x) = 0$. Judd et al. (2012) discuss the theory, application, and implementation of this method; also see Borkovsky et al. (2010). The second approach is called the Gröbner basis method. This approach is based on a result in algebraic geometry which lets us reduce the problem of solving for all solutions to a system of equations to the problem of finding all solutions to just a single equation with a single unknown. Kubler and Schmedders (2010) describe this approach and illustrate its application in economics.

Both approaches require the system to be a polynomial, which is satisfied here.[19] The Gröbner basis method has the nice feature that when the system of interest is polynomial with rational coefficients, it proves *exact* analytical solutions, not just numerical approximations. If we restrict the decision-maker's λ_{ij}'s to be rational, then the first-order conditions of the statistical decision game are rational, since binomial coefficients are rational. The downside of Gröbner basis methods is that it can only handle smaller systems of equations, compared to all-solutions homotopy. One limitation of the numerical approach is that as the sample size increases, so does the number of equations in the system, eventually making the problem infeasible under either method. Nonetheless, at such sample sizes, the asymptotic approximation results of section 2 are likely to be reasonable, and can be used instead.

---

[18] Although all optimal rules yield the same value of maximal regret, and hence the decision-maker is technically indifferent between them.

[19] In general, homotopy methods whose only goal is to find *one* solution apply to arbitrary systems. Only the all-solutions homotopy imposes the polynomial restriction.

3.4 Discussion

The results in this section show that when no analytical finite-sample optimality results are available, numerical computations can be a useful substitute. I showed that optimality depends importantly on the sampling design. The empirical success rule, while optimal for balanced designs, performs poorly in highly unbalanced designs. My results suggest that, when choosing the sampling design, all treatments should be treated as symmetrically as possible. For example, ex ante known allocations with balanced designs yield lower regret for many rules compared to ex ante known allocations with unbalanced designs. Likewise, ex ante unknown allocations are preferred to ex ante known allocations.

Much future work remains. In particular, implementing the computational methods discussed will help analyze more complicated settings, such as larger sample sizes, more than three treatments, and asymmetric costs of treatment.

A Proofs

Proof of proposition 1 Assumption 3.2 implies that the sequence of experiments

$$\{P_{\theta_1}^{n_1} \otimes \cdots \otimes P_{\theta_K}^{n_K} : \theta \in \Theta^K\}$$

is locally asymptotically normal with norming rate $\sqrt{N}$. Let

$$P_{N,h} = P_{\theta_0 + h_1/\sqrt{N}}^{n_1} \otimes \cdots \otimes P_{\theta_0 + h_K/\sqrt{N}}^{n_K}.$$

where $h = (h_1', \ldots, h_K')'$. Then

$$\log \frac{dP_{N,h}}{dP_{\theta_0}^N} = h'\Delta_{N,\theta_0} - \frac{1}{2}h'\mathbf{I}_0 h + o_{P_{\theta_0}}(1),$$

where

$$\Delta_{N,\theta_0} = \begin{pmatrix} \Delta_{N,\theta_0,1} \\ \vdots \\ \Delta_{N,\theta_0,K} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{\lambda_1}}{\sqrt{n_1}} \sum_{i=1}^{n_1} \dot{\ell}_{\theta_0}(X_{1i}) \\ \vdots \\ \frac{\sqrt{\lambda_K}}{\sqrt{n_K}} \sum_{i=1}^{n_K} \dot{\ell}_{\theta_0}(X_{Ki}) \end{pmatrix}, \quad \text{and} \quad \mathbf{I}_0 = \begin{pmatrix} \lambda_1 I_0 & 0 & \cdots & 0 \\ 0 & \lambda_2 I_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K I_0 \end{pmatrix}.$$

Here $\Delta_{N,\theta_0} \overset{\theta_0}{\rightsquigarrow} \mathcal{N}(0, \mathbf{I}_0)$. The corresponding limit experiment consists of observing K independent normally distributed variables with means h_k and variances $\lambda_k^{-1} I_0^{-1}$.

Since $\hat{\theta}_{k,N}$ is a best regular estimator of θ_k, the converse part of lemma 8.14 of Van der Vaart (1998) implies that the estimator sequence $\hat{\theta}_{k,N}$ satisfies the expansion

$$\sqrt{N}(\hat{\theta}_{k,N} - \theta_0) = \lambda_k^{-1} I_0^{-1} \Delta_{N,\theta_0,k} + o_{P_{\theta_0}}(1).$$

By the delta method,

$$\sqrt{N}[w(\hat{\theta}_{k,N}) - w(\theta_0)] = \sqrt{N} w(\hat{\theta}_{k,N})$$
$$= \dot{w}' \lambda_k^{-1} I_0^{-1} \Delta_{N,\theta_0,k} + o_{P_{\theta_0}}(1),$$

where recall that

$$\dot{w} = \frac{\partial w}{\partial \theta}\Big|_{\theta = \theta_0}.$$

Recall that $\hat{\theta}_N = (\hat{\theta}'_{1,N}, \dots, \hat{\theta}'_{K,N})'$, $\mathbf{w}(\hat{\theta}_N) = (w(\hat{\theta}_{1,N}), \dots, w(\hat{\theta}_{K,N}))'$, and

$$\dot{\mathbf{w}}' = \begin{pmatrix} \dot{w}' & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \dot{w}' \end{pmatrix}.$$

Then the previous results show that

$$\sqrt{N} \mathbf{w}(\hat{\theta}_N) = \dot{\mathbf{w}}' \mathbf{I}_0^{-1} \Delta_{N,\theta_0} + o_{P_{\theta_0}}(1),$$

where note that

$$\mathbf{I}_0^{-1} = \begin{pmatrix} \lambda_1^{-1} I_0^{-1} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1} I_0^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K^{-1} I_0^{-1} \end{pmatrix}.$$

Hence, by Slutsky's theorem (this allows us to drop the $o_{P_{\theta_0}}(1)$ terms) and a multi-variate CLT, we have

$$\left(\sqrt{N} \mathbf{w}(\hat{\theta}_N), \quad \log \frac{dP_{N,h}}{dP_{\theta_0}^N} \right) = \left(\dot{\mathbf{w}}' \mathbf{I}_0^{-1} \Delta_{N,\theta_0}, \quad h' \Delta_{N,\theta_0} - \frac{1}{2} h' \mathbf{I}_0 h \right) + o_{P_{\theta_0}}(1)$$

$$\overset{\theta_0}{\rightsquigarrow} \mathcal{N}\left(\begin{pmatrix} 0 \\ -\frac{1}{2} h' \mathbf{I}_0 h \end{pmatrix}, \begin{pmatrix} \mathbb{E}_{\theta_0} \psi_{\theta_0} \psi'_{\theta_0} & \mathbb{E}_{\theta_0} \psi_{\theta_0} h' \dot{\ell}_{\theta_0} \\ \mathbb{E}_{\theta_0} \psi'_{\theta_0} h' \dot{\ell}_{\theta_0} & h' \mathbf{I}_0 h \end{pmatrix} \right)$$

$$= \mathcal{N}\left(\begin{pmatrix} 0 \\ -\frac{1}{2} h' \mathbf{I}_0 h \end{pmatrix}, \begin{pmatrix} \dot{\mathbf{w}}' \mathbf{I}_0^{-1} \dot{\mathbf{w}} & \dot{\mathbf{w}}' h \\ \dot{\mathbf{w}}' h & h' \mathbf{I}_0 h \end{pmatrix} \right)$$

where $\psi_{\theta_\theta} = \dot{\mathbf{w}}' \mathbf{I}_0^{-1} \dot{\ell}_{\theta_0}$, and we have defined

$$\dot{\ell}_{\theta_0} = (\sqrt{\lambda_1} \dot{\ell}_{\theta_0}(X_1)', \dots, \sqrt{\lambda_K} \dot{\ell}_{\theta_0}(X_K)').$$

Then $\mathbb{E}_{\theta_0}\dot{\boldsymbol{\ell}}_{\theta_0}\dot{\boldsymbol{\ell}}'_{\theta_0} = \mathbf{I}_0$, where note that the off-diagonal entries are zero since the K samples are independent and the expected score is zero. Hence the third line follows since

$$
\begin{aligned}
\mathbb{E}_{\theta_0}\psi_{\theta_0}\psi'_{\theta_0} &= \mathbb{E}_{\theta_0}\dot{\mathbf{w}}'\mathbf{I}_0^{-1}\dot{\boldsymbol{\ell}}_{\theta_0}(\dot{\mathbf{w}}'\mathbf{I}_0^{-1}\dot{\boldsymbol{\ell}}_{\theta_0})' \\
&= \dot{\mathbf{w}}'\mathbf{I}_0^{-1}\mathbb{E}_{\theta_0}[\dot{\boldsymbol{\ell}}_{\theta_0}\dot{\boldsymbol{\ell}}'_{\theta_0}]\mathbf{I}_0^{-1}\dot{\mathbf{w}} \\
&= \dot{\mathbf{w}}'\mathbf{I}_0^{-1}\dot{\mathbf{w}}
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}_{\theta_0}\psi'_{\theta_0}h'\dot{\boldsymbol{\ell}}_{\theta_0} &= \dot{\mathbf{w}}'\mathbf{I}_0^{-1}\mathbb{E}_{\theta_0}[\dot{\boldsymbol{\ell}}_{\theta_0}\dot{\boldsymbol{\ell}}'_{\theta_0}]h \\
&= \dot{\mathbf{w}}'\mathbf{I}_0^{-1}\mathbf{I}_0 h \\
&= \dot{\mathbf{w}}'h.
\end{aligned}
$$

Thus, by Le Cam's third lemma (see example 6.7, page 90 of (Van der Vaart, 1998)):

$$
\begin{aligned}
\sqrt{N}\mathbf{w}(\hat{\theta}_N) &\overset{h}{\rightsquigarrow} \mathcal{N}\big(0 + \dot{\mathbf{w}}'h, \dot{\mathbf{w}}'\mathbf{I}_0^{-1}\dot{\mathbf{w}}\big) \\
&= \mathcal{N}\left(\begin{pmatrix} \dot{w}'h_1 \\ \vdots \\ \dot{w}'h_K \end{pmatrix}, \begin{pmatrix} \lambda_1^{-1}\dot{w}'I_0^{-1}\dot{w} & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1}\dot{w}'I_0^{-1}\dot{w} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K^{-1}\dot{w}'I_0^{-1}\dot{w} \end{pmatrix}\right).
\end{aligned}
$$

$\square$

Proof of theorem 2 This result follows from the results in appendix B, as summarized in corollary S1. Note that assumption 5 ensures that the variances are equal in the limit experiment, as needed to apply those results. $\square$

The following lemma restates the parts of lemma 4 of Hirano and Porter (2009) which are relevant here, for reference.

Lemma S1 (Hirano and Porter (2009) *lemma* 4) *Suppose assumptions* 1–4 *hold. Let* δ_N *be a sequence of treatment assignment rules which is matched by a rule* δ *in the limit experiment, as given by the asymptotic representation theorem. Let J be a finite subset of* $\mathbb{R}^{\dim(h)}$. *If*

$$
\lim_{N\to\infty}\sqrt{N}R_N\left[\delta_N, \theta_0 + \frac{h}{\sqrt{N}}\right] = R_\infty(\delta, h)
$$

holds for all $h \in \mathbb{R}^{\dim(h)}$, *then*

$$
\sup_J \liminf_{N\to\infty} \sup_{h\in J} \sqrt{N}R_N\left[\delta_N, \theta_0 + \frac{h}{\sqrt{N}}\right] = \sup_h R_\infty(\delta, h),
$$

where the outer supremum over J is taken over all finite subsets J of $\mathbb{R}^{\dim(h)}$.

Proof of lemma 2 By the asymptotic representation theorem, $\delta_N \in \mathcal{D}$ is matched by a rule $\tilde{\delta}$, meaning that $\mathbb{E}_{\theta_0+h/\sqrt{N}}\delta_{k,N} \to \mathbb{E}_h\tilde{\delta}_k$ for all h. Consequently, by our risk function calculations in section 2.3, we have

$$\lim_{N\to\infty} \sqrt{N}R_N\left[\delta_N, \theta_0 + \frac{h}{\sqrt{N}}\right] = R_\infty(\tilde{\delta}, h)$$

for all h. Thus

$$\sup_J \liminf_{N\to\infty} \sup_{h\in J} \sqrt{N}R_N\left[\delta_N, \theta_0 + \frac{h}{\sqrt{N}}\right] = \sup_h R_\infty(\tilde{\delta}, h)$$
$$\geq \inf_\delta \sup_h R_\infty(\delta, h),$$

where the first line follows by lemma 4 of Hirano and Porter (2009). □

Proof of theorem 3 By proposition 1,

$$\begin{aligned}
\beta_{k,N}^*(h) &\equiv \mathbb{E}_{\theta_0+h/\sqrt{N}}\delta_{k,N}^* \\
&= P_{\theta_0+h/\sqrt{N}}\left(\sqrt{N}w(\hat{\theta}_{k,N}) > \sqrt{N}w(\hat{\theta}_{s,N}) \text{ for all } s \neq k\right) \\
&\to \mathbb{P}_h(\dot{w}'\Delta_k > \dot{w}'\Delta_s \text{ for all } s \neq k) \qquad\qquad \text{as } N \to \infty \\
&= \mathbb{E}_h[\delta_k^*(\Delta)] \\
&= \int \delta_k^*(\Delta_k)\, d\mathcal{N}(\Delta_k \mid h_k, \lambda_k^{-1}I_0^{-1}) \\
&\equiv \beta_k^*(h)
\end{aligned}$$

for all h. Thus, by our risk calculations in section 2.3, δ_N^* is matched in the limit experiment by δ^*; that is,

$$\lim_{N\to\infty} \sqrt{N}R_N\left[\delta_N^*, \theta_0 + \frac{h}{\sqrt{N}}\right] = R_\infty(\delta^*, h)$$

for all h. Then

$$\sup_J \liminf_{N\to\infty} \sup_{h\in J} \sqrt{N}R_N\left[\delta_N^*, \theta_0 + \frac{h}{\sqrt{N}}\right] = \sup_h R_\infty(\delta^*, h)$$
$$= \inf_\delta \sup_h R_\infty(\delta, h),$$

where the outer supremum over J is taken over all finite subsets of $\mathbb{R}^{\dim(h)}$. The first line follows by lemma 4 of Hirano and Porter (2009). The second line follows since δ^* is optimal in the limit experiment, by theorem 2. □

B The finite-sample results of Bahadur, Goodman, and Lehmann

In this appendix I reproduce several results from Lehmann (1966), who built on Bahadur (1950) and Bahadur and Goodman (1952). Theorem 2 in the present paper is an immediate corollary of these results. None of the results in this section are new (except for corollary S2, which is a minor extension); I include them here for reference and accessibility. I have also made a few notational changes. The basic approach is to first show that the empirical success rule is uniformly best within the class of permutation invariant rules, and then show that this implies that it is a minimax rule among all possible rules. Although theorem 2 only requires this result to hold for normally distributed data, the result holds more generally, and so I present the general result here.

Suppose there are K treatments. From each treatment group $k = 1, \ldots, K$ we observe a vector of outcomes X_k. Let P_θ denote the joint distribution of the data $X = (X_1, \ldots, X_K)$, where $\theta = (\theta_1, \ldots, \theta_K)$ and $\theta_k \in \mathbb{R}$. Let $L(k, \theta) = L_k(\theta)$ denote the loss from assigning everyone to treatment k. Let

$$G = \{g_1, \ldots, g_J\}$$

be a finite group of transformations.

Example 2 (Permutation group) Let $\pi : \{1, \ldots, K\} \to \{1, \ldots, K\}$ be a function which permutes the indices of θ. So $\pi(k)$ denotes the new index which k is mapped into. $\pi^{-1}(k)$ denotes the old index which maps into the new kth spot. Let $\{\pi_j\}$ be the set of all J permutations of indices. The set of transformations

$$g_j(\theta) = (\theta_{\pi_j^{-1}1}, \ldots, \theta_{\pi_j^{-1}K})$$

which permute the indices of θ is called the *permutation group*. For example, let $\theta = (\theta_1, \theta_2, \theta_3)$. Consider the particular permutation π_j with $\pi_j(1) = 2$, $\pi_j(2) = 1$, and $\pi_j(3) = 3$. Then

$$g_j(\theta) = (\theta_2, \theta_1, \theta_3).$$

Assume from now on that G is the permutation group. Let δ denote an arbitrary treatment rule. We say δ is *invariant under G* if

$$g_j[(\delta_1(X), \ldots, \delta_K(X))] = (\delta_1(g_j X), \ldots, \delta_K(g_j X)) \quad \text{for all } j \text{ and all } X.$$

Assumption S1 The following hold.

1. (Invariant distribution). P_θ is invariant under G. That is, the distribution of $g_j(X)$ is $P_{g_j(\theta)}$ for all j.[20]

2. (Invariant loss). $L(k, \theta)$ is invariant under G. That is, if $L((1, \ldots, K), \theta)$ is the vector of losses then

$$L(g_j((1, \ldots, K)), g_j(\theta)) = L((1, \ldots, K), \theta)$$

   for all j.

3. (Independence). $X = (X_1, \ldots, X_K)$ has a probability density of the form

$$h_\theta(T) = C(\theta) f_{\theta_1}(T_1) \cdots f_{\theta_K}(T_K),$$

   with respect to a σ-finite measure ν, where $T_k = T_k(X_k)$, are real valued statistics, and $T = (T_1, \ldots, T_K)$.

4. (MLR). f_{θ_k} has monotone (non-decreasing) likelihood ratio in T_k for each k.

5. (Monotone loss). Larger parameter values are weakly preferred. That is, the loss function L satisfies

$$\theta_i < \theta_j \quad \Rightarrow \quad L_i(\theta) \geq L_j(\theta).$$

Example 3 (Independent normal observations with equal variances) Let $X_k \sim \mathcal{N}(\theta_k, \sigma^2)$, and X_k is independent of all other observations. Then this data satisfies assumptions (1), (3), and (4), with $T_k = X_k$.

Example 4 (Regret loss) The regret loss function

$$L(k, \theta) = \max_i \ \theta_i - \theta_k$$

satisfies (2) and (5).

Let

$$\mathcal{M} = \{k : T_k \geq T_j \text{ for all } j\}$$

denote the set of treatments which yield the largest observed statistics T_k. Define the empirical success rule by

$$\delta_k^*(X) = \begin{cases} 1/|\mathcal{M}| & \text{if } T_k \geq T_j \text{ for all } j \\ 0 & \text{if } T_k < T_j \text{ for some } j. \end{cases} \tag{8}$$

---

[20] Compare the assumption of invariance to the assumption that P_θ is exchangeable. Exchangeability is stronger: it says that $g_j(X) \sim P_\theta$ for all j—we do not have to permute the parameters. Also note that this invariant distribution assumption is violated as soon as we allow unequal sample sizes. For example, if $T_0 \sim \mathcal{N}(\mu_0, \sigma^2/N_0)$ and $T_1 \sim \mathcal{N}(\mu_1, \sigma^2/N_1)$, then permuting the indices yields $T_0 \sim \mathcal{N}(\mu_0, \sigma^2/N_1)$ and $T_1 \sim \mathcal{N}(\mu_1, \sigma^2/N_0)$, which is not true when $N_0 \neq N_1$.

When the T_k are continuously distributed, it is not necessary to define what δ^* does in the event of ties, since they occur with probability zero. When T_k are discretely distributed, however, ties may occur with positive probability.

Theorem S1 (Bahadur–Goodman–Lehmann) *Suppose assumption S1 holds. Then the empirical success rule* (8) *uniformly over θ minimizes the risk $R(\delta, \theta)$ among all rules δ based on T which are invariant under G.*

Proof First note that δ^* is invariant under G. By lemma 1 of Lehmann (1966) (see below), it suffices to prove that, for all θ, δ^* minimizes risk averaged over all permutations,

$$r(\delta, \theta) \equiv \frac{1}{J} \sum_{j=1}^{J} R(\delta, g_j \theta),$$

among all procedures (not just invariant ones). The risk of a rule δ is

$$R(\delta, \theta) = \sum_{k=1}^{K} L_k(\theta) \mathbb{E}[\delta_k(T)]$$

$$= \sum_{k=1}^{K} L_k(\theta) \int \delta_k(t) h_\theta(t) \, dv(t).$$

Average risk is

$$r(\delta, \theta) = \frac{1}{J} \sum_{j=1}^{J} R(\delta, g_j \theta)$$

$$= \frac{1}{J} \sum_{j=1}^{J} \sum_{k=1}^{K} L_k(g_j \theta) \int \delta_k(t) h_{g_j \theta}(t) \, dv(t)$$

$$= \frac{1}{J} \sum_{k=1}^{K} \int \left[\sum_{j=1}^{J} L_k(g_j \theta) h_{g_j \theta}(t) \right] \delta_k(t) \, dv(t)$$

$$= \frac{1}{J} \sum_{k=1}^{K} \int A_k(t) \delta_k(t) \, dv(t),$$

where we defined

$$A_k(t) = \sum_{j=1}^{J} L_k(g_j \theta) h_{g_j \theta}(t).$$

Suppose without loss of generality that $\theta_1 < \cdots < \theta_K$. We will show that, for a fixed t, $A_k(t)$ is minimized over k by any i such that

$$t_i = \max_k t_k$$

Consequently, the average risk $r(\delta, \theta)$ is minimized for any rule δ which puts $\delta_k(t) = 0$ whenever $t_k < \max t_i$. Finally, since the lemma requires the procedure δ^* to be invariant under G, we break ties between maxima at random with equal probability.

Thus all we have to do is show that for a fixed t,

$$A_i(t) = \min_k A_k(t) \qquad \text{if and only if} \qquad t_i = \max_k t_k.$$

A_i minimized means $A_i \le A_k$ for all k, or $A_i - A_k \le 0$ for all k. Thus consider the difference

$$A_i - A_k = \sum_{j=1}^{J} [L_i(g_j\theta) - L_k(g_j\theta)] h_{g_j\theta}(t).$$

We will look at each term in the sum separately. Consider a permutation g defined by π such that $\pi^{-1}i < \pi^{-1}k$. By our ordering of the parameters, this implies that $\theta_{\pi^{-1}i} < \theta_{\pi^{-1}k}$. Let g' be the permutation obtained from g by swapping $\pi^{-1}i$ and $\pi^{-1}k$.

The contribution to $A_i - A_k$ from the terms corresponding to g and g' is

$$[L_i(g\theta) - L_k(g\theta)]h_{g\theta}(t) + [L_i(g'\theta) - L_k(g'\theta)]h_{g'\theta}(t).$$

Since the loss function is invariant,

$$L_i(g\theta) = L_k(g'\theta) \qquad \text{and} \qquad L_k(g\theta) = L_i(g'\theta).$$

So the previous equation reduces to

$$[L_i(g\theta) - L_k(g\theta)][h_{g\theta}(t) - h_{g'\theta}(t)].$$

The first piece is positive (nonnegative at least) since $\theta_{\pi^{-1}k} > \theta_{\pi^{-1}i}$. Consider the sign of the second term. By the form of the density,

$$h_{g\theta}(t) = C(g\theta)f_{\theta_{\pi^{-1}1}}(t_1) \cdots f_{\theta_{\pi^{-1}K}}(t_K)$$

Since g' is the same permutation as g, except that j_i and j_k are switched, $h_{g'\theta}(t)$ is the same as above, except that $f_{\theta_{\pi^{-1}i}}$ is evaluated at t_k and $f_{\theta_{\pi^{-1}k}}$ is evaluated at t_i. Thus, when we compare the difference $h_{g\theta}(t) - h_{g'\theta}(t)$ to zero, we can divide by all the shared terms (including the constant C, which is invariant), which all cancel, leaving just the $\pi^{-1}i$ and $\pi^{-1}k$ terms remaining.

Thus

$$h_{g\theta}(t) - h_{g'\theta}(t) \ge (\le)0 \qquad \Leftrightarrow \qquad f_{\theta_{\pi^{-1}i}}(t_i)f_{\theta_{\pi^{-1}k}}(t_k) \ge (\le)f_{\theta_{\pi^{-1}k}}(t_i)f_{\theta_{\pi^{-1}i}}(t_k)$$

$$\Leftrightarrow \qquad \frac{f_{\theta_{\pi^{-1}k}}(t_k)}{f_{\theta_{\pi^{-1}i}}(t_k)} \ge (\le)\frac{f_{\theta_{\pi^{-1}k}}(t_i)}{f_{\theta_{\pi^{-1}i}}(t_i)}$$

$$\Leftrightarrow \qquad t_k \ge (\le)t_i.$$

The last line follows since $\theta_{\pi^{-1}k} \geq \theta_{\pi^{-1}i}$ and f_θ has MLR. Thus, when $t_k \leq t_i$, the second term is negative, and hence the entire term is negative.

Now, *every* permutation g_π must have either $\pi^{-1}i < \pi^{-1}k$ or $\pi^{-1}i > \pi^{-1}k$. So if we consider the set of all permutations with $\pi^{-1}i < \pi^{-1}k$ then we can uniquely pair each one of these with a permutation with $\pi^{-1}k < \pi^{-1}i$. This is what we did above. By doing so, we have considered *every* term in the summation which defines $A_i - A_k$. Thus, when $t_k \leq t_i$, every term in the sum has the desired sign, and hence

$$A_i - A_k \leq 0,$$

as desired. $\square$

Lemma S2 (*Lehmann* ([1966](#)) *lemma* 1) *Fix θ. A necessary and sufficient condition for an invariant procedure δ^* to minimize the risk $R(\delta, \theta)$ among all invariant procedures is that it minimizes the average risk*

$$r(\delta, \theta) = \frac{1}{J} \sum_{j=1}^{J} R(\delta, g_j\theta)$$

among all procedures (not just the invariant ones).

Proof

1. (Sufficient). Let δ^* be an invariant procedure which minimizes average risk $r(\delta, \theta)$ among all procedures. Let δ' be any other invariant procedure. Now, the risk function of any invariant procedure is constant over each orbit[21]

   $$\{g_j\theta : j = 1, \ldots, J\}.$$

   Thus

   $$R(\delta', \theta,) = r(\delta', \theta) \qquad \text{and} \qquad R(\delta^*, \theta) = r(\delta^*, \theta),$$

   since both δ' and δ^* are invariant. By the assumption that δ^* minimizes average risk among all procedures,

   $$r(\delta^*, \theta) \leq r(\delta', \theta).$$

   Hence

   $$R(\delta^*, \theta) \leq R(\delta', \theta).$$

2. (Necessary). Suppose that δ^* minimizes risk $R(\delta, \theta)$ among all invariant procedures. Let δ' be any procedure. There exists an invariant procedure δ'' such that

   $$r(\delta', \theta) = r(\delta'', \theta).$$

---

[21] See Berger ([1985](#)) page 396 for proof.

For example, we can take

$$\delta'' = \frac{1}{J} \sum_{j=1}^{J} g_j[\delta'(g_j^{-1}X)].$$

Thus we have

$$r(\delta', \theta) = r(\delta'', \theta) = R(\delta'', \theta)$$
$$\geq R(\delta^*, \theta) = r(\delta^*, \theta).$$

□

Finally, this result implies certain optimality properties among all rules, not just permutation invariant ones. To this end, suppose all rules are ranked by a complete and transitive ordering $\succeq$.

Example 5 The minimax criterion orders rules as follows: $\delta' \succeq \delta$ if

$$\sup_{\theta} R(\delta', \theta) \leq \sup_{\theta} R(\delta, \theta).$$

This ordering satisfies all the assumptions of Lehmann's lemma 2 below.

Lemma S3 (*Lehmann* (1966) *lemma 2*)

1. *If the ordering $\succeq$ is such that*

(a) *$\delta' \succeq \delta$ implies $g\delta g^{-1} \succeq g\delta' g^{-1}$, and*
(b) *for any finite set of rules $\delta^{(i)}, i = 1, \ldots, r, \delta^{(i)} \succeq \delta$ for all i implies $\frac{1}{r} \sum_{i=1}^{r} \delta^{(i)} \succeq \delta$.*

   then given any procedure δ there exists an invariant procedure δ' such that $\delta' \succeq \delta$.
2. *Suppose that in addition to (a) and (b) above, the ordering satisfies*

(c) *$R(\delta', \theta) \leq R(\delta, \theta)$ for all θ implies $\delta' \succeq \delta$.*

   Then, if there exists a procedure δ^* that uniformly minimizes the risk among all invariant procedures, δ^* is optimal with respect to the ordering $\succeq$. That is, $\delta^* \succeq \delta$ for all δ.

Proof

1. Note that

$$\delta' = \frac{1}{r} \sum_{j=1}^{r} g_j \delta g_j^{-1}$$

is invariant and, by (a) and (b), at least as good as δ.

2. Let δ be any rule. By part 1, there exists an invariant δ' that is preferred to δ. δ^* uniformly minimizes risk among all invariant procedures. In particular, it has uniformly smaller risk than δ'. Thus, by (c), $\delta^* \succeq \delta' \succeq \delta$.

$\square$

Thus, combining the BHL theorem with Lehmann's lemma 2 yields the following result.

Corollary S1 *The empirical success rule* (8) *is minimax optimal under regret loss in the experiment where, for all k, $X_k \sim \mathcal{N}(\theta_k, \sigma^2)$ and X_k is independent of all other observations.*

As a final remark, Schlag (2003) showed that when outcomes have an arbitrary distribution on a common bounded support, it suffices to only consider binary outcomes by performing a 'binary randomization'. This technique can be used to extend the Bahadur-Goodman-Lehman result as follows.

Corollary S2 *Let X_k be a vector $(X_{k,1}, \dots, X_{k,n_k})$ of n_k scalar observations. Suppose $n_1 = \dots = n_K$. Suppose the parameter space for the distribution of $X_{k,i}$ is the set of all distributions with common bounded support, for all $i = 1, \dots, n_k$, for all k. Then the empirical success rule* (8) *is minimax optimal under any loss function satisfying assumptions* (2) *and* (5).

References

Abughalous, M., & Miescke, K. J. (1989). On selecting the largest success probability under unequal sample sizes. *Journal of Statistical Planning and Inference, 21*, 53–68.

Bahadur, R. (1950). On a problem in the theory of k populations. *The Annals of Mathematical Statistics, 21*, 362–375.

Bahadur, R., & Goodman, L. (1952). Impartial decision rules and sufficient statistics. *The Annals of Mathematical Statistics, 23*, 553–562.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis, 2nd ed.*, Springer Verlag.

Bofinger, E. (1985). Monotonicity of the probability of correct selection or are bigger samples better? Journal of the Royal Statistical Society. *Series B (Methodological), 47*, 84–89.

Borkovsky, R. N., Doraszelski, U., & Kryukov, Y. (2010). A user's guide to solving dynamic stochastic games using the homotopy method. *Operations Research, 58*, 1116–1132.

Chamberlain, G. (2000). Econometrics and decision theory. *Journal of Econometrics, 95*, 255–283.

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

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

Judd, K. L. (1998). *Numerical Methods in Economics*, The MIT press.

Judd, K. L., Renner, P., & Schmedders, K. (2012). Finding all pure-strategy equilibria in dynamic and static games with continuous strategies, *Working paper*.

Kallus, N. (2018). Balanced policy evaluation and learning, *Advances in Neural Information Processing Systems*, 31.

Kitagawa, T., & Tetenov, A. (2017). Who should be treated? empirical welfare maximization methods for treatment choice, *cemmap Working Paper (CWP24/17)*.

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

Kubler, F., & Schmedders, K. (2010). Tackling multiplicity of equilibria with Gröbner bases. *Operations research, 58*, 1037–1050.

Le Cam, L. M. (1986). *Asymptotic Methods in Statistical Decision Theory*. New York: Springer-Verlag.

Lehmann, E. (1966). On a theorem of Bahadur and Goodman. *The Annals of Mathematical Statistics, 37*, 1–6.

List, J. A., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? *Some Simple Rules of Thumb for Optimal Experimental Design, Experimental Economics, 14*, 439–457.

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

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

Schlag, K. H. (2006). Eleven-Tests needed for a recommendation, Working Paper.

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

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

Manski, C. F. (2009). Diversified treatment under ambiguity. *International Economic Review, 50*, 1013–1041.

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

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

Masten, M. A. (2013). Equilibrium Models in Econometrics, Ph.D. thesis, Northwestern University.

Miescke, K. J., & Park, H. (1999). On the natural selection rule under normality. *Statistics & Decisions, Supplemental Issue No., 4*, 165–178.

Otsu, T. (2008). Large deviation asymptotics for statistical treatment rules. *Economics Letters, 101*, 53–56.

Schlag, K. H. (2003). How to minimize maximum regret in repeated decision-making, Working Paper.

Song, K. (2014). Point decisions for interval-identified parameters. *Econometric Theory, 30*, 334–356.

Stoye, J. (2007a). Minimax-regret treatment choice with finite samples and missing outcome data, in *Proceedings of the Fifth International Symposium on Imprecise Probability: Theories and Applications*, ed. by M. Z. G. de Cooman, J. Veinarová.

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

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

Stoye, J. (2009b). Web Appendix to "Minimax regret treatment choice with finite samples", Mimeo.

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

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

Tong, Y., & Wetzell, D. (1979). On the behaviour of the probability function for selecting the best normal population. *Biometrika, 66*, 174–176.

Van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

Wald, A. (1945). Statistical decision functions which minimize the maximum risk. *The Annals of Mathematics, 46*, 265–280.

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

Zhou, Z., Athey, S., & Wager, S. (2023). Offline multi-action policy learning: Generalization and optimization. *Operations Research, 71*, 148–183.