Chapter 15

Estimation of dynamic treatment regimes for complex outcomes: Balancing benefits and risks

Kristin A. Linn
Eric B. Laber
Leonard A. Stefanski

15.1 ▪ Introduction

Dynamic treatment regimes (DTRs) formalize clinical decision making as a sequence of decision rules, one per treatment decision, that map current patient information to a recommended treatment. A DTR is said to be optimal if it maximizes the expected value of a single scalar outcome when used to assign treatments to a population of interest. This notion of optimality has been used almost without exception in the statistics literature to dictate both estimation and evaluation of DTRs; yet, this framework oversimplifies the goals of clinical decision making which include balancing several potentially competing outcomes and accommodating heterogeneous patient preferences across these outcomes. For example, the tradeoff between treatment efficacy and tolerability has been studied within many disease-specific subfields including depression and other mental illnesses [37, 75, 269], HIV/AIDS [47], occupational HIV exposure [24], urinary conditions [74], chronic pain management [181], and cancer [89, 48, 206]. We offer a new framework that allows for much more general definitions of optimality in which the optimal regime maximizes a single scalar feature of a multivariate outcome distribution subject to (possibly many) constraints on other features of the outcome distribution. Examples include maximizing expected efficacy subject to constraints on side effect burden, maximizing the median of a quality of life measure subject to constraints on cost, and maximizing the expected time to an adverse event subject to constraints on the variance of the time to the adverse event. By varying the constraints, patients and clinicians can interactively explore tradeoffs across treatment strategies and select a treatment plan that best suits the individual needs of each patient.

The proposed methodology includes as a special case maximizing an expected outcome without any constraints, a setting for which many estimators already exist. Examples include Q-learning [435, 434, 243], interactive Q-learning [175], regularized Q-

learning [237, 61, 359, 129, 128], G-estimation [316], A-learning [241], regret-regression [142], and policy search methods [263, 463, 454, 453, 451]. One approach to dealing with competing outcomes within the framework of maximizing the expectation of a single scalar outcome is to form a composite outcome. Wang et al. [430] provide a case study using this approach and Lizotte et al. [202] try to use composite outcomes to learn about patient preferences. However, Laber et al. [177] point out several problems associated with constructing a high quality composite outcome: (i) a population-level composite outcome ignores individual preferences, (ii) patients may not know or may not be able to effectively communicate their preferences, and (iii) preferences may evolve over time. They also show that misspecification of a composite outcome can severely affect the quality of the estimated DTR. While preference elicitation can be used to improve the quality of a composite outcome, such elicitation is typically done separately from the construction of a DTR. Consequently, a patient or clinician may not be aware of how their reported preferences are affecting the estimated DTR.

In contrast, by posing the estimation of an optimal DTR as a constrained optimization problem we can (i) overtly incorporate competing outcomes and (ii) allow preference elicitation that is directly connected the estimated DTR by interactively changing the constraints and re-estimating the optimal DTR. For simplicity, in this chapter we restrict attention to the setting in which there is a bivariate outcome and data are collected in a two-stage sequential multiple assignment randomized trial (SMART) [186, 242]. This simple setting allows us to illustrate the basic ideas without excessive technical development. Generalizations to multiple outcomes and observational data are possible.

15.2 ▪ Optimal constrained treatment regimes

15.2.1 ▪ Notation and setup

We assume that data are generated in a two-stage SMART with binary treatments at each stage. Let $\{(X_{1i},A_{1i},X_{2i},A_{2i},Y_i,Z_i)\}_{i=1}^n$ denote independent and identically distributed copies of the trajectory (X_1,A_1,X_2,A_2,Y,Z), one for each of n subjects. Baseline subject information is contained in $X_1 \in \mathbb{R}^{p_1}$; $A_1 \in \{-1,1\}$ is the first-stage treatment; $X_2 \in \mathbb{R}^{p_2}$ contains interim information collected between the first- and second-stage treatments; $A_2 \in \{-1,1\}$ is the second-line treatment; $Y \in \mathbb{R}$ is the primary outcome of interest; and $Z \in \mathbb{R}$ is a second outcome of interest. For example, Y might be efficacy and Z a competing outcome such as side effect burden. We assume that Y and Z are continuous. Let $H_2 = (X_1^\intercal, A_1, X_2^\intercal)^\intercal$, so that H_2 includes all information available to a decision maker prior to assigning the second-stage treatment. In a SMART the distribution of A_1 is known and depends only on X_1, and the distribution of A_2 is known and depends only on H_2. We further assume that there exists $\eta > 0$ so that $\eta < P(A_1 = 1|X_1) < 1-\eta$ and $\eta < P(A_2 = 1|H_2) < 1-\eta$ with probability 1. The first-stage decision rule d_1 maps the support of X_1 to the space of possible first-stage treatments, and the second-stage decision rule d_2 maps the support of H_2 to the space of possible second-stage treatments. The pair of decision rules $d = (d_1,d_2)$ is a DTR, where a patient presenting with $X_1 = x_1$ at stage one would be assigned treatment $d_1(x_1)$, and a patient presenting with $H_2 = h_2$ would be assigned $d_2(h_2)$.

Let $\mathrm{pr}^d_{Y,Z}(y,z)$ denote the joint distribution of (Y,Z) when treatments are assigned according to the regime d, evaluated at (y,z). Similarly, let $\mathrm{pr}^d_Y(y)$ and $\mathrm{pr}^d_Z(z)$ denote the marginal distributions of Y and Z under regime d, evaluated at y and z, respectively. We

assume the following:

$$Y = m_Y(H_2) + A_2 c_Y(H_2) + \epsilon_Y,$$
$$Z = m_Z(H_2) + A_2 c_Z(H_2) + \epsilon_Z$$

for functions $\{m_Y(\cdot), c_Y(\cdot), m_Z(\cdot), c_Z(\cdot)\}$ and mean-zero error terms (ϵ_Y, ϵ_Z) that are independent of (H_2, A_2) and have joint distribution $F_{\epsilon_Y,\epsilon_Z}(\cdot,\cdot)$.

Let $M(\text{pr}^d_{Y,Z})$ denote a feature of the joint distribution function of (Y,Z) under d that we would like to maximize, and let $S(\text{pr}^d_{Y,Z})$ be an r-dimensional feature of this distribution that we would like to bound above. In addition, let $\mathscr{D}$ denote a class of regimes of interest and for $\mathbf{b}, \mathbf{c} \in \mathbb{R}^q$ write $\mathbf{b} \preceq \mathbf{c}$ to denote $b_j \leq c_j$ for $j = 1, \ldots, r$. If one exists, we define the optimal S-constrained regime, d^{opt}_S, as a solution to

$$\max_{d \in \mathscr{D}} M(\text{pr}^d_{Y,Z})$$
$$\text{subject to } S(\text{pr}^d_{Y,Z}) \preceq \varkappa, \tag{15.1}$$

where $\varkappa \in \mathbb{R}^r$ is a known vector of constants.

In practice, one might choose to simplify the constrained optimization problem in (15.1) by specifying an objective $M(\text{pr}^d_{Y,Z})$ that depends only on the marginal distribution of Y, i.e., $M(\text{pr}^d_{Y,Z}) = M(\text{pr}^d_Y)$. Similarly, the constrained quantity, $S(\text{pr}^d_{Y,Z})$, may be chosen to depend only on the marginal distribution of Z. In this case, problem (15.1) simplifies to

$$\max_{d \in \mathscr{D}} M(\text{pr}^d_Y)$$
$$\text{subject to } S(\text{pr}^d_Z) \preceq \varkappa. \tag{15.2}$$

To provide context for problem (15.2), consider a hypothetical SMART that aims to study treatment sequences for patients with depression, such as a simplified version of the STAR*D trial introduced in Chapter 2. Suppose remission is defined as having a depression score below a cutoff, say l^*, and that most patients can tolerate at most a side effect burden of u^*. Thus, investigators may be interested in identifying a treatment regime that maximizes the probability of remission subject to an average side effect burden that does not exceed the upper bound u^*. In our notation, this example corresponds to setting $M(\text{pr}^d_Y) = \text{pr}^d_Y(l^*)$, $S(\text{pr}^d_Z) = \int z \, d\,\text{pr}^d_Z(z)$, and $\varkappa = u^*$ in problem (15.2).

More generally, if Y is a measure of efficacy, one might specify $M(\text{pr}^d_Y)$ to be average efficacy, $\int y \, d\,\text{pr}^d_Y(y)$, the τth quantile of the efficacy distribution, $M(\text{pr}^d_Y) = \inf\{y : \text{pr}^d_Y(y) \geq \tau\}$, or the probability that Y exceeds a lower bound, say l^*, for which $M(\text{pr}^d_Y) = 1 - \text{pr}^d_Y(l^*)$. If Z is a measure of adverse effects, then one may specify the average adverse effect burden, $S(\text{pr}^d_Z) = \int z \, d\,\text{pr}^d_Z(z)$, as a constraint. Alternatively, using problem (15.1), $S(\text{pr}^d_{Y,Z})$ could be a two-dimensional constraint that considers both the variance of Y and the probability that Z exceeds some upper bound u^*, so that $S(\text{pr}^d_{Y,Z}) = \{\text{var}^d Y, 1 - \text{pr}^d_Z(u^*)\}^\intercal$, where $\text{var}^d Y$ denotes the variance of Y under regime d.

Remark 15.2.1. *Mean-optimal, single-outcome methods such as Q-learning [435, 434, 243], interactive Q-learning [175], G-estimation [316], A-learning [241], and policy search [263, 463, 454, 453, 451] solve 15.1 without the constraint and with $M(pr^d_{Y,Z}) = M(pr^d_Y)$ specified*

*as the average outcome under d, i.e., $M(pr^d_Y) = \int y\, dpr^d_Y(y)$. In the hypothetical depression SMART example, this would correspond to maximizing a single measure of efficacy, such as the negative depression symptom score (since lower scores indicate better symptom profiles). Using a subset of data from the STAR*D trial, [175] provide an illustrative example of such an analysis using interactive Q-learning.*

15.2.2 ▪ Estimating distributional summaries of Y and Z

If the underlying joint distribution of the data $(X_1, A_1, X_2, A_2, Y, Z)$ is known, constrained optimization methods could be used to solve problem (15.1). In most settings, however, the data generating distribution is unknown, and both the objective function $M(\text{pr}^d_{Y,Z})$ and constraints $S(\text{pr}^d_{Y,Z})$ must be estimated. Let $\widehat{\text{pr}}^d_{Y,Z}(y,z)$ denote an estimator of the joint cumulative distribution function $\text{pr}^d_{Y,Z}(y,z)$. If one exists, we define the estimated optimal S-constrained treatment regime, $\widehat{d}^{\text{opt}}_S$, as a solution to

$$\begin{aligned} &\max_{d\in\mathscr{D}} M(\widehat{\text{pr}}^d_{Y,Z}) \\ &\text{subject to } \; S(\widehat{\text{pr}}^d_{Y,Z}) \preceq \varkappa \end{aligned} \tag{15.3}$$

for known constant vector $\varkappa \in \mathbb{R}^r$. Broadly, solving problem (15.3) proceeds by estimating $\text{pr}^d_{Y,Z}(y,z)$, sampling from the estimated distribution function to estimate the objective and constraints, and finding the solution using a grid search or stochastic programming algorithm. For some functionals M and S, availability of an exact analytical form may negate the need for Monte Carlo approximation using samples from the estimated distribution. However, the methods developed in this chapter are computationally difficult in many settings, and thus it may be more convenient in practice to approximate the functionals using samples from the estimated distribution.

Remark 15.2.2. *We say problem (15.1) is* separable *if the objective and constraints depend only on the marginal distributions of Y and Z under regime d, i.e., $pr^d_Y(y)$ and $pr^d_Z(z)$, and not on the joint distribution $pr^d_{Y,Z}(y,z)$. Thus, problem (15.2) gives the separable form of problem 15.1. Consider the problem of maximizing the mean of Y, $M(pr^d_{Y,Z}) = M(pr^d_Y) = \int y\, dpr^d_Y(y)$ subject to bounds on the variance of Y and probability that Z exceeds an upper bound u^*, $S(pr^d_{Y,Z}) = S(pr^d_Y, pr^d_Z) = \{var^d Y,\ 1 - pr^d_Z(u^*)\}^\intercal$. Recalling our hypothetical depression trial example, Y might be a recoded depression score such that higher values indicate better symptom profiles, and Z might be a one-number summary of side effect burden. Then, this setup satisfies our definition of a separable constrained optimization problem. In this case, we need only estimate $pr^d_Y(y)$ and $pr^d_Z(z)$ instead of the joint distribution, $pr^d_{Y,Z}(y,z)$.*

Using the assumptions stated in Section 15.2.1, we now give the form of the marginal distributions of Y and Z under a given policy $d = (d_1, d_2)$. Let $d_2(h_2) = \text{sign}\{f_2(h_2)\}$ for some function $f_2(h_2)$, where we define $\text{sign}(x) = I(x \geq 0) - I(x < 0)$. Define the following: $F_{X_1}(\cdot)$ is the cumulative distribution function of vector X_1, $F_{X_2 \mid X_1, A_1}(\cdot \mid x_1, a_1)$ is the conditional cumulative distribution function of vector X_2 given $X_1 = x_1$ and $A_1 = a_1$, and $F_{\epsilon_Y}(\cdot)$ is the cumulative distribution function of ϵ_Y. In addition, define

$$J^d_Y(x_1, x_2, y) = F_{\epsilon_Y}(y - m_Y\{x_1, d_1(x_1), x_2\} - \text{sign}[f_2\{x_1, d_1(x_1), x_2\}]c_Y\{x_1, d_1(x_1), x_2\}).$$

We write $m_Y\{x_1, d_1(x_1), x_2\}$ and $c_Y\{x_1, d_1(x_1), x_2\}$ instead of $m_Y(h_2)$ and $c_Y(h_2)$ to emphasize the fact that under regime d, h_2 depends on x_1, $a_1 = d(x_1)$ and x_2. Then,

$$\mathrm{pr}_Y^d(y) = \int\int J_Y^d(x_1, x_2, y) dF_{X_2 \mid X_1, A_1}\{x_2 \mid x_1, d_1(x_1)\} dF_{X_1}(x_1).$$

Let $G_Y^{d_2}(\cdot,\cdot,\cdot \mid x_1, a_1)$ denote the joint conditional distribution of $\{m_Y(X_1, A_1, X_2), c_Y(X_1, A_1, X_2), f_2(X_1, A_1, X_2)\}$ given $X_1 = x_1$ and $A_1 = a_1$, so that

$$G_Y^{d_2}(s, t, w \mid x_1, a_1) = \int I(m_Y(x_1, a_1, x_2) \le s) I(c_Y(x_1, a_1, x_2) \le t) I(f_2(x_1, a_1, x_2) \le w) dF_{X_2 \mid X_1, A_1}(x_2 \mid x_1, a_1).$$

Then, one can show by interchanging the order of integration that

$$\int\int F_{\epsilon_Y}\{y - s - \mathrm{sign}(w)t\} dG_Y^{d_2}\{s, t, w \mid x_1, d_1(x_1)\} dF_{X_1}(x_1) = \int\int J_Y^d(x_1, x_2, y) dF_{X_2 \mid X_1, A_1}\{x_2 \mid x_1, d_1(x_1)\} dF_{X_1}(x_1),$$

where the right side is equivalent to $\mathrm{pr}_Y^d(y)$. Thus,

$$\mathrm{pr}_Y^d(y) = E\left[\int F_{\epsilon_Y}\{y - s - \mathrm{sign}(w)t\} dG_Y^{d_2}\{s, t, w \mid X_1, d_1(X_1)\}\right], \tag{15.4}$$

where the expectation is over the marginal distribution of X_1. Similarly, it can be shown that

$$\mathrm{pr}_Z^d(z) = E\left[\int F_{\epsilon_Z}\{z - u - \mathrm{sign}(w)v\} dG_Z^{d_2}\{u, v, w \mid X_1, d_1(X_1)\}\right], \tag{15.5}$$

where $F_{\epsilon_Z}(\cdot)$ is the marginal distribution function of ϵ_Z, and $G_Z^{d_2}(\cdot,\cdot,\cdot \mid x_1, a_1)$ is the joint conditional distribution of $\{m_Z(X_1, A_1, X_2), c_Z(X_1, A_1, X_2), f_2(X_1, A_1, X_2)\}$ given $X_1 = x_1$ and $A_1 = a_1$. Define $1_{x_1, a_1, x_2}(s, t, u, v, w)$ to be the product of the indicators $I(m_Y(x_1, a_1, x_2) \le s)$, $I(c_Y(x_1, a_1, x_2) \le t)$, $I(m_Z(x_1, a_1, x_2) \le u)$, $I(c_Z(x_1, a_1, x_2) \le v)$, and $I(f_2(x_1, a_1, x_2) \le w)$, and

$$G_{Y,Z}^{d_2}(s, t, u, v, w \mid x_1, a_1) = \int 1_{x_1, a_1, x_2}(s, t, u, v, w) dF_{X_2 \mid X_1, A_1}(x_2 \mid x_1, a_1).$$

Thus, $G_{Y,Z}^{d_2}(s, t, u, v, w \mid x_1, a_1)$ is the joint conditional distribution of

$$\{m_Y(X_1, A_1, X_2), c_Y(X_1, A_1, X_2), m_Z(X_1, A_1, X_2), c_Z(X_1, A_1, X_2), f_2(X_1, A_1, X_2)\}$$

given $X_1 = x_1$ and $A_1 = a_1$. The joint cumulative distribution of (Y, Z) under regime d and evaluated at (y, z) is

$$\mathrm{pr}_{Y,Z}^d(y, z) = E\left[\int F_{\epsilon_Y, \epsilon_Z}(y, z, s, t, u, v, w) dG_{Y,Z}^{d_2}\{s, t, u, v, w \mid X_1, d_1(X_1)\}\right], \tag{15.6}$$

where we define $F_{\epsilon_Y, \epsilon_Z}(y, z, s, t, u, v, w) = F_{\epsilon_Y, \epsilon_Z}\{y - s - \mathrm{sign}(w)t,\ z - u - \mathrm{sign}(w)v\}$ in the previous expression, and the outer expectation is with respect to X_1.

Expressions (15.4), (15.5), and (15.6) suggest that estimating $\mathrm{pr}_Y^d(y)$, $\mathrm{pr}_Z^d(z)$, and $\mathrm{pr}_{Y,Z}^d(y,z)$ can be accomplished by estimating the distributions $F_{\epsilon_Y}(\cdot)$, $F_{\epsilon_Z}(\cdot)$, $F_{\epsilon_Y,\epsilon_Z}(\cdot)$, $G_Y^{d_2}(\cdot,\cdot,\cdot \mid x_1,a_1)$, $G_Z^{d_2}(\cdot,\cdot,\cdot \mid x_1,a_1)$, and $G_{Y,Z}^{d_2}(\cdot,\cdot,\cdot,\cdot,\cdot \mid x_1,a_1)$. Let $\widehat{F}_{\epsilon_Y}(\cdot)$, $\widehat{F}_{\epsilon_Z}(\cdot)$, $\widehat{F}_{\epsilon_Y,\epsilon_Z}(\cdot)$, $\widehat{G}_Y^{d_2}(\cdot,\cdot,\cdot \mid x_1,a_1)$, $\widehat{G}_Z^{d_2}(\cdot,\cdot,\cdot \mid x_1,a_1)$, and $\widehat{G}_{Y,Z}^{d_2}(\cdot,\cdot,\cdot,\cdot,\cdot \mid x_1,a_1)$ denote estimators of the corresponding true distribution functions. Let E_n denote empirical expectation with respect to X_1. Then, for a given d, the following expressions are estimators of $\mathrm{pr}_Y^d(y)$, $\mathrm{pr}_Z^d(z)$, and $\mathrm{pr}_{Y,Z}^d(y,z)$:

$$\widehat{\mathrm{pr}}_Y^d(y) = E_n \int \widehat{F}_{\epsilon_Y}\{y - s - \mathrm{sign}(w)t\} d\widehat{G}_Y\{s,t,w \mid X_1, d_1(X_1)\}, \tag{15.7}$$

$$\widehat{\mathrm{pr}}_Z^d(z) = E_n \int \widehat{F}_{\epsilon_Z}\{z - u - \mathrm{sign}(w)v\} d\widehat{G}_Z\{u,v,w \mid X_1, d_1(X_1)\}, \tag{15.8}$$

$$\widehat{\mathrm{pr}}_{Y,Z}^d(y,z) = E_n \int \widehat{F}_{\epsilon_Y,\epsilon_Z}(y,z,s,t,u,v,w) d\widehat{G}_Z\{s,t,u,v,w \mid X_1, d_1(X_1)\}. \tag{15.9}$$

Approaches to estimating $G_Y^{d_2}$, $G_Z^{d_2}$, and $G_{Y,Z}^{d_2}$ We assume S-constrained decision rules at stage 2 can be expressed as $d_{2,S}(h_2) = \mathrm{sign}\{f_2(h_2)\}$ for some smooth function $f_2(\cdot)$. Thus, we arbitrarily assign treatment 1 at the second stage when $f_2(h_2) = 0$.

Recall from Section 15.2.1 that H_2 denotes the full history prior to treatment assignment at stage 2, i.e., $H_2 = (X_1^\intercal, A_1, X_2^\intercal)^\intercal$. Our strategy for modeling $G_{Y,Z}^{d_2}(\cdot,\cdot,\cdot,\cdot,\cdot \mid x_1,a_1)$ is to model the joint distribution of the standardized residuals,

$$\begin{aligned}
e_Y^m &= \frac{m_Y(H_2) - \mu_Y^m(X_1,A_1)}{\sigma_Y^m(X_1,A_1)}, \quad e_Y^c = \frac{c_Y(H_2) - \mu_Y^c(X_1,A_1)}{\sigma_Y^c(X_1,A_1)}, \\
e_Z^m &= \frac{m_Z(H_2) - \mu_Z^m(X_1,A_1)}{\sigma_Z^m(X_1,A_1)}, \quad e_Z^c = \frac{c_Z(H_2) - \mu_Z^c(X_1,A_1)}{\sigma_Z^c(X_1,A_1)}, \\
e^{f_2} &= \frac{f_2(H_2) - \mu^{f_2}(X_1,A_1)}{\sigma^{f_2}(X_1,A_1)},
\end{aligned} \tag{15.10}$$

that are obtained from mean and variance modeling of $m_Y(H_2)$, $c_Y(H_2)$, $m_Z(H_2)$, $c_Z(H_2)$, and $f_2(H_2)$ given X_1 and A_1 [54]. In (15.10), the mean functions are defined as

$$\begin{aligned}
\mu_Y^m(X_1,A_1) &= E\{m_Y(H_2) \mid X_1,A_1\}, \quad \mu_Y^c(X_1,A_1) = E\{c_Y(H_2) \mid X_1,A_1\}, \\
\mu_Z^m(X_1,A_1) &= E\{m_Z(H_2) \mid X_1,A_1\}, \quad \mu_Z^c(X_1,A_1) = E\{c_Z(H_2) \mid X_1,A_1\}, \\
\mu_{f_2}(X_1,A_1) &= E\{f_2(H_2) \mid X_1,A_1\},
\end{aligned}$$

and the standard deviation functions are defined as

$$\begin{aligned}
\sigma_Y^m(X_1,A_1) &= E[\{m_Y(H_2) - \mu_Y^m(X_1,A_1)\}^2 \mid X_1,A_1]^{1/2}, \\
\sigma_Y^c(X_1,A_1) &= E[\{c_Y(H_2) - \mu_Y^c(X_1,A_1)\}^2 \mid X_1,A_1]^{1/2}, \\
\sigma_Z^m(X_1,A_1) &= E[\{m_Z(H_2) - \mu_Z^m(X_1,A_1)\}^2 \mid X_1,A_1]^{1/2}, \\
\sigma_Z^c(X_1,A_1) &= E[\{c_Z(H_2) - \mu_Z^c(X_1,A_1)\}^2 \mid X_1,A_1]^{1/2}, \\
\sigma_{f_2}(X_1,A_1) &= E[\{f_2(H_2) - \mu_{f_2}(X_1,A_1)\}^2 \mid X_1,A_1]^{1/2}.
\end{aligned}$$

Our strategy for modeling $G_Y^{d_2}(\cdot,\cdot,\cdot \mid x_1,a_1)$ is to model the joint conditional distribution of the standardized residuals (e_Y^m, e_Y^c, e^{f_2}). If the constraints of the optimization problem are separable, then we would forego modeling the joint distribution of all the residuals in (15.10) for modeling the joint distribution of (e_Z^m, e_Z^c, e^{f_2}) to obtain an estimator of $G_Z^{d_2}(\cdot,\cdot,\cdot \mid x_1,a_1)$ instead of $G_{Y,Z}^{d_2}(\cdot,\cdot,\cdot,\cdot,\cdot \mid x_1,a_1)$.

In practice, clinical data are often expensive to obtain, and thus sample sizes are moderate or small. It may be useful to consider parametric models for $m_Y(H_2)$, $c_Y(H_2)$, $m_Z(H_2)$, $c_Z(H_2)$, and $f_2(H_2)$, as well as their mean and variance functions. For example, assuming linear working models throughout, the following steps may be used to estimate $m_Y(H_2) = m_Y(H_2;\beta_Y^m)$, $\mu_Y^m(X_1,A_1) = \mu_Y^m(X_1,A_1;\theta_Y^m)$ and $\sigma_Y^m(X_1,A_1) = \sigma_Y^m(X_1,A_1;\gamma_Y^m)$, thus obtaining estimated versions of the standardized residuals.

Algorithm for estimating e_Y^m

(i) Regress Y on H_2 and A_2 using the linear working model

$$Y = H_2^\intercal \beta_Y^m + A_2 H_2^\intercal \beta_Y^c + \epsilon_Y$$

to obtain the estimate $m_Y(H_2;\widehat{\beta}_Y^m) = H_2^\intercal \widehat{\beta}_Y^m$ of $m_Y(H_2)$. Note this also provides the estimate $c_Y(H_2;\widehat{\beta}_Y^c) = H_2^\intercal \widehat{\beta}_Y^c$ of $c_Y(H_2)$.

(ii) Estimate $\mu_Y^m(X_1,A_1;\theta_Y)$ by fitting a least squares regression using the working model

$$m_Y(H_2;\widehat{\beta}_Y^m) = H_1^\intercal \theta_{Y,0}^m + A_1 H_1^\intercal \theta_{Y,1}^m + \xi,$$

where ξ is a mean-zero error term and $\theta_Y^m = (\theta_{Y,0}^m, \theta_{Y,1}^m)$. Then, $\mu_Y^m(X_1,A_1;\widehat{\theta}_Y) = H_1^\intercal \widehat{\theta}_{Y,0}^m + A_1 H_1^\intercal \widehat{\theta}_{Y,1}^m$.

(iii) Using the estimated mean function, $\mu_Y^m(X_1,A_1;\widehat{\theta}_Y^m)$, from step (ii), fit a least squares regression using the working model

$$2\log\{m(H_{2i};\widehat{\beta}_Y^m) - \mu_Y^m(X_{1i},A_{1i};\widehat{\theta}_Y^m)\} = H_1^\intercal \gamma_{Y,0}^m + A_1 H_1^\intercal \gamma_{Y,1}^m + \delta,$$

where δ is a mean-zero error term and $\gamma_Y^m = (\gamma_{Y,0}^m, \gamma_{Y,1}^m)$. Then, $\sigma_Y^m(X_1,A_1;\widehat{\gamma}_Y^m) = \exp\{(H_1^\intercal \widehat{\gamma}_{Y,0}^m + A_1 H_1^\intercal \widehat{\gamma}_{Y,1}^m)/2\}^{1/2}$ is the resulting estimator of $\sigma_Y^m(X_1,A_1)$. In practice, an intercept term is often included in H_1, and the estimated intercept parameter may be biased by a scale factor. To correct for this bias, we adjust the estimated intercept so that the sample standard deviation of $\widehat{e}_{Y,i}^m$ is 1, where $\widehat{e}_{Y,i}^m$ is defined next.

(iv) Define the estimated standardized residuals as

$$\widehat{e}_{Y,i}^m = \frac{m_Y(H_{2i};\widehat{\beta}_Y^m) - \mu_Y^m(X_{1i},A_{1i};\widehat{\theta}_Y^m)}{\sigma_Y^m(X_{1i},A_{1i};\widehat{\gamma}_Y^{m,*})},$$

for $i = 1,\ldots,n$, where $\widehat{\gamma}_Y^{m,*}$ is the intercept-adjusted version of $\widehat{\gamma}_Y^m$ from step (iii).

Steps (i)–(iv) can be repeated with minor adjustments to obtain the remaining estimated standardized residuals $\{\widehat{e}_{Y,i}^c, \widehat{e}_{Z,i}^m, \widehat{e}_{Z,i}^c, \widehat{e}_i^{f_2}\}$. One convenient strategy for modeling the joint distribution of the standardized residuals is to use a copula; for example, in [200], the authors use a Gaussian copula. Random samples from the estimated copula can be transformed back to samples from $\widehat{G}_Y(\cdot,\cdot,\cdot \mid x_1,a_1)$, $\widehat{G}_Z(\cdot,\cdot,\cdot \mid x_1,a_1)$, or $\widehat{G}_{Y,Z}(\cdot,\cdot,\cdot \mid x_1,a_1)$ using the estimated mean and variance functions.

Estimating F_{ϵ_Y}, F_{ϵ_Z}, and $F_{\epsilon_Y,\epsilon_Z}$ Standard methods can be used to estimate the marginal distribution functions F_{ϵ_Y} and F_{ϵ_Z}; for example, it might be appropriate to use a parametric scale family or the empirical distribution function. Similarly, the joint distribution function $F_{\epsilon_Y,\epsilon_Z}$ could be estimated using a bivariate normal distribution, bivariate density estimation, or the empirical distribution function. Thus, we do not discuss modeling these error distributions further.

Sampling from estimated marginal distribution functions $\widehat{\mathrm{pr}}^d_Y$ and $\widehat{\mathrm{pr}}^d_Z$ Standard methods, e.g., inversion sampling, can be used to sample from $\widehat{\mathrm{pr}}^d_Y(y)$. However, because samples from $\widehat{\mathrm{pr}}^d_Y(y)$ (and $\widehat{\mathrm{pr}}^d_z(z)$) may need to be generated many times in order to solve problem (15.3), it may be necessary to use a cruder but more computationally efficient sampling scheme. Below we describe a simple sampling scheme based on discretization; the algorithm is described in terms of $\widehat{\mathrm{pr}}^d_Y(y)$ but applies with obvious modification to $\widehat{\mathrm{pr}}^d_Z(z)$. The algorithm is as follows:

(i) Partition $\mathbb{R}$ into M intervals $(-\infty, y_1], (y_1, y_2], \ldots, (y_{M-2}, y_{M-1}], (y_{M-1}, \infty)$.

(ii) For $j = 1, \ldots, M-1$ compute

$$\widehat{\mathrm{pr}}^d_Y(y_j) = E_n \int \widehat{F}_{\epsilon_Y}\{y_j - s - \mathrm{sign}(w)t\}\, d\widehat{G}_Z\{s, t, w \mid X_1, d_1(X_1)\},$$

and define $q^d_1 = \widehat{\mathrm{pr}}^d_Y(y_1)$, $q^d_M = 1 - \widehat{\mathrm{pr}}^d_Y(y_{M-1})$, and $q^d_j = \widehat{\mathrm{pr}}^d_Y(y_j) - \widehat{\mathrm{pr}}^d_Y(y_{j-1})$ for $j = 2, \ldots, M-1$.

(iii) Randomly sample from a multinomial distribution that takes the values $(y_1, (y_1 + y_2)/2, \ldots, (y_{M-2} + y_{M-1})/2, y_{M-1})$ with probabilities $(q^d_1, \ldots, q^d_M)$.

The above algorithm is trivial to implement and requires only $M-1$ evaluations of the estimated distribution function $\widehat{\mathrm{pr}}^d_Y$ regardless of the size of the randomly generated sample. Even if M is large, say of the same order as the generated sample, this is often much less expensive than inverting the estimated distribution function. A refinement of the above approach is to approximate the distribution function using either linear interpolation or a basis expansion and then to use inversion sampling with the approximated distribution function [261]. One potential difficulty is the choice of the endpoints in the partition used in step (i) in the previous algorithm. In settings where the support of Y (Z) is bounded across all DTRs, say to $[c, d]$, then $y_1 = c + \delta$ and $y_{M-1} = d - \delta$ for some $\delta > 0$ is a natural choice. However, in settings where the support of Y (Z) is unbounded across all DTRs some trial and error may be required to obtain a satisfactory choice. In the course of solving problem (15.3), if either q^d_1 or q^d_M becomes large then endpoints of the grid should be expanded. On the other hand, if any of the $q^d_j, 1 < j < M-1$ becomes large then the grid may be too coarse and may need to be refined.

Sampling from the estimated joint distribution function $\widehat{\mathrm{pr}}^d_{Y,Z}$ To sample from $\widehat{\mathrm{pr}}^d_{Y,Z}$ we use a bivariate extension of the discrete sampling scheme given for $\widehat{\mathrm{pr}}^d_{Y,Z}$. Let $\mathbb{R}^e$ denote the extended real line $\mathbb{R} \cup \{\pm\infty\}$ and for any two $u, v \in \mathbb{R}^e$ with $u < v$ define

$$\overline{\mathrm{ave}}(u, v) = \begin{cases} v, & u = -\infty, v < \infty, \\ (v+u)/2, & u > -\infty, v < \infty, \\ u, & u > -\infty, v = \infty, \\ \text{undefined}, & u = -\infty, v = \infty. \end{cases}$$

The algorithm is as follows:

(i) Partition $\mathbb{R}^2$ into M disjoint rectangles $r_1,\ldots,r_M$. For each $j=1,\ldots,M$, let $y_{j,1}<y_{j,2}$ and $z_{j,1}<z_{j,2}$ denote the four points in the extended real line, $\mathbb{R}\cup\{\pm\infty\}$, that determine rectangle r_j.

(ii) Use

$$\widehat{\mathrm{pr}}^d_{Y,Z}(y_{j,k},z_{j,k'}) = E_n\int \widehat{F}_{\epsilon_Y,\epsilon_Z}\{y_{j,k}-s-\mathrm{sign}(w)t, z_{j,k'}-u-\mathrm{sign}(w)v\} \times d\widehat{G}_Z\{s,t,u,v,w\mid X_1,d_1(X_1)\},$$

to determine the estimated probability mass inside rectangle r_j

$$q^d_j=\widehat{\mathrm{pr}}^d_{Y,Z}(y_{j,2},z_{j,2})-\widehat{\mathrm{pr}}^d_{Y,Z}(y_{j,1},z_{j,2})-\widehat{\mathrm{pr}}^d_{Y,Z}(y_{j,2},z_{j,1})+\widehat{\mathrm{pr}}^d_{Y,Z}(y_{j,1},z_{j,1}).$$

(iii) Randomly sample from a multinomial distribution that takes the paired value $\{\overline{\mathrm{ave}}(y_{j,1},y_{j,2}),\overline{\mathrm{ave}}(z_{j,1},z_{j,2})\}$ with probability q^d_j, $j=1,\ldots,M$.

As in the univariate setting, the choice of the partition is a potentially important tuning parameter. Monitoring the values of q^d_j, $j=1,\ldots,M$, can diagnose potential problems with the choice of partition.

15.3 ▪ Toy example

We illustrate the proposed method in the separable setting where the mean of Y is maximized subject to an upper bound on the mean of Z. To provide context, suppose Y is the negative of the final depression score measured at the end of a two-stage SMART, so that larger values indicate fewer depression symptoms. In addition, let Z represent the final measurement of global side effect burden, for which lower values are desirable. Thus, we consider problem (15.1) where $M(\mathrm{pr}^d_{Y,Z})=\int y\,d\mathrm{pr}^d_Y$ is the average negative depression score with the single constraint $S(\mathrm{pr}^d_{Y,Z})=\int z\,d\mathrm{pr}^d_Z=S(\mathrm{pr}^d_Z)$ being the average side effect burden score. For a grid of upper bounds on the average side effect burden $S(\mathrm{pr}^d_Z)$, say $\varkappa_l$ for $l=1,\ldots,L$, we estimate the optimal constrained policy using training data and display the resulting mean of Y when the estimated policy is applied to a population of test patients.

Patient trajectories (X_1,A_1,X_2,A_2,Y,Z) are generated independently from the model:

$$\begin{aligned} X_1\sim\mathrm{Normal}(1,1),\quad & A_t\sim\mathrm{Uniform}\{-1,1\},\quad t=1,2,\\ H_1=(1,X_1)^\intercal,\quad & X_2=H_1^\intercal\beta_{1,0}+A_1H_1^\intercal\beta_{1,1}+\eta(H_1,A_1)\xi,\\ \eta(H_1,A_1)=\exp(H_1^\intercal\gamma_0+A_1H_1^\intercal\gamma_1),\quad & \xi\sim\mathrm{Normal}(0,1),\\ H_2=(1,X_2)^\intercal,\quad & Y=H_2^\intercal\beta_{2,0,Y}+A_2H_2^\intercal\beta_{2,1,Y}+\epsilon_Y,\\ Z=H_2^\intercal\beta_{2,0,Z}+A_2H_2^\intercal\beta_{2,1,Z}+\epsilon_Z,\quad & (\epsilon_Y,\epsilon_Z)^\intercal\sim\mathrm{Normal}(0_2,\Sigma_{Y,Z}), \end{aligned}$$

where $\Sigma_{Y,Z}$ is a 2×2 matrix with 1s on the diagonal and 0.7s on the off diagonals. Thus, Y and Z are positively correlated, reflecting the scenario where higher efficacy is attained at the expense of higher side effect burden. Although greatly simplified, the data generated

from the model above represent data that would be observed from a two-stage randomized SMART. Variables X_1 and X_2 represent patient status summaries measured before treatment at the first and second stages, respectively. For example, X_1 may represent the depression score observed at baseline and X_2 the depression score following first-stage treatment. Thus, X_2 is an updated version of X_1 that also depends on first-stage treatment, A_1. We assume larger values of Y and smaller values of Z are desirable. These outcomes represent the depression and side effect scores observed at the conclusion of the trial, and are generated as linear functions of the second-stage depression score and second-stage randomized treatment, A_2. The remaining parameters of the generative model are set to

$$\begin{aligned} \beta_{1,0} &= (0.5, 0.75), \quad \beta_{1,1} = (0.25, 0.5), \\ \gamma_0 &= (0.25, -0.05), \quad \gamma_1 = (0.1, -0.05), \\ \beta_{2,0,Y} &= (30, 2), \quad \beta_{2,1,Y} = (5, -1.5), \\ \beta_{2,0,Z} &= (15, 1), \quad \beta_{2,1,Z} = (3, -0.5). \end{aligned}$$

These parameter values were chosen so that $d_2(h_2) = \text{sign}(h_2^\intercal \beta_{2,1,Y})$ is mean-optimal for Y for every second-stage history, h_2, but is nonoptimal for the mean of Z for at least some h_2. That is, the optimal regime estimated by our constrained framework will differ from optimal regimes estimated using unconstrained, mean-optimal frameworks such as Q-learning.

We restrict attention to estimation of linear decision rules at each stage. Let $d_1^{\text{opt}}(h_1) = \text{sign}(h_1^\intercal \eta_1^{\text{opt}})$ and $d_2^{\text{opt}}(h_2) = \text{sign}(h_2^\intercal \eta_2^{\text{opt}})$ denote the true optimal linear decision rules, and let $\widehat{d}_1^{\text{opt}}(h_1) = \text{sign}(h_1^\intercal \widehat{\eta}_1^{\text{opt}})$ and $\widehat{d}_2^{\text{opt}}(h_2) = \text{sign}(h_2^\intercal \widehat{\eta}_2^{\text{opt}})$ denote their estimated counterparts. To find an approximate solution to problem (15.3) we used a version of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm implemented in the freely available Gnu Scientific Library [132] with log-barrier function to represent the constraint [211]. To improve solution quality, we used five multiple randomly generated starting values. The complete algorithm is as follows:

Constrained interactive Q-learning algorithm:

(i) Maximize $\widehat{E}^{\eta}(Y)$ with respect to $\eta = (\eta_1, \eta_2)$ without any constraint on $\widehat{E}^{\eta}(Z)$. Denote the solution by $\widehat{\eta}^{Y\text{-opt}}$.

(ii) Estimate $E^{\widehat{\eta}^{Y\text{-opt}}}(Z)$; denote the estimate by $\widehat{E}^{\widehat{\eta}^{Y\text{-opt}}}(Z)$.

  a. If $\widehat{E}^{\widehat{\eta}^{Y\text{-opt}}}(Z) < \varkappa$, stop. Set $\widehat{\eta}^{\text{opt}} = \widehat{\eta}^{Y\text{-opt}}$.

  b. If $\widehat{E}^{\widehat{\eta}^{Y\text{-opt}}}(Z) \geq \varkappa$, proceed to step (iii); $\widehat{\eta}^{Y\text{-opt}}$ does not correspond to a feasible regime because the constraint on the mean of Z is not satisfied.

(iii) Minimize $\widehat{E}^{\eta}(Z)$ with respect to $\eta = (\eta_1, \eta_2)$. Denote the solution by $\widehat{\eta}^{Z\text{-opt}}$.

(iv) Estimate $E^{\widehat{\eta}^{Z\text{-opt}}}(Z)$; denote the estimate by $\widehat{E}^{\widehat{\eta}^{Z\text{-opt}}}(Z)$.

  a. If $\widehat{E}^{\widehat{\eta}^{Z\text{-opt}}}(Z) \geq \varkappa$, stop. No feasible regime exists because the constraint is not satisfied for the optimal solution for the mean of Z.

  b. If $\widehat{E}^{\widehat{\eta}^{Z\text{-opt}}}(Z) < \varkappa$, proceed to step (v).

(v) Estimate η^{opt} by solving

$$\max_{\eta} \widehat{E}^{\eta}(Y) + \lambda_K \log\{\varkappa - \widehat{E}^{\eta}(Z)\} \tag{15.11}$$

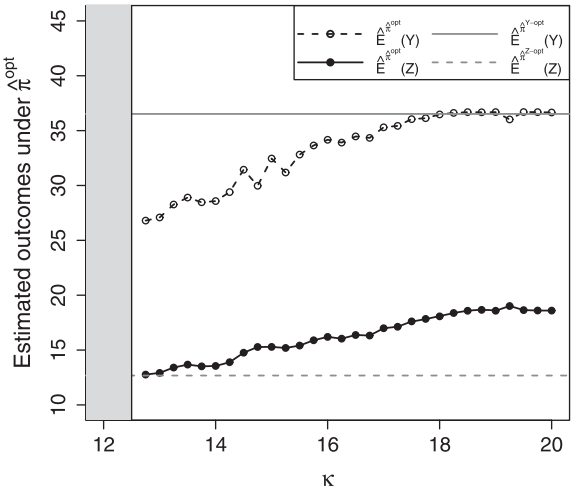

Figure 15.1. *Estimated mean outcomes for a range of constraints on Z. The solid gray line is the estimated mean of Y under the estimated unconstrained optimal regime. The dashed gray line is the estimated mean of Z under the regime that maximizes $E^d(Z)$. The black dashed line with empty circles shows the estimated mean of Y under the estimated optimal constrained regime; the solid black line with filled circles shows the estimated mean of Z under the estimated optimal constrained regime. No feasible regime existed for κ values in the shaded region.*

for a small positive value of $\lambda_K > 0$. To improve numerical stability, we solved using a sequence of decreasing $\lambda_1 > \lambda_2 > \cdots > \lambda_K$ of the form $\lambda_j = \lambda_{j-1}/4$ and used the solution at λ_{j-1} as the starting value at λ_j, $j = 2,\ldots,K$. We set $\widehat{\eta}^{\text{opt}}$ to be the solution at λ_K. In our experiments we set $K = 5$.

We found that the starting penalty value λ_1 was an important tuning parameter. In the simulated experiments considered here we choose λ_1 so that

$$\widehat{E}^{\eta^{Y-\text{opt}}}(Y) - E_n Y = \lambda_1 \log\left\{\kappa - \widehat{E}^{\eta^{Z-\text{opt}}}(Z)\right\}.$$

The motivation for this choice is that the log-penalty should be of the same order of magnitude as the potential gain in the mean of Y over treatment randomization.

Results are presented in Figure 15.1 for a sequence of κ from 12 to 20 in increments of 0.25. The algorithm stopped at step (iv)a for $\kappa = 12.5$ and below; this is represented by the shaded region. We display the estimated means of Y and Z under the estimated optimal constrained policy, $\widehat{d}^{\text{opt}}$, using a dashed line with open circles and a solid line with closed circles, respectively. We also plot the unconstrained estimated optimal mean of Y as a solid gray line, and the estimated optimal mean of Z under no restrictions on the distribution of Y is represented by the dashed gray line. The tradeoff between outcomes Y and Z is apparent in Figure 15.1: small values of κ lead to near optimal values of the mean of Z but poor values of the mean of Y and large values of κ lead to near optimal values in the mean of Y but poor values in the mean of Z.

Figures are one way to help decision makers understand the inherent tradeoffs across outcomes *at the population level.* With modifications to the methods presented here, it is possible to make a plot analogous to Figure 15.1 that is conditional on a patient's current characteristics, thereby producing a decision aid for an individual patient and facilitating communication between patients and clinicians about individual preferences. We are

currently working on this extension. Historical data on patient side effect tolerance in combination with Figure 15.1 may guide practitioners in the decision process by providing an estimate of the average efficacy that might be attained should they implement a constrained regime based on the historical tolerance data. Alternatively, illustrations such as Figure 15.1 could be used as exploratory tools to inform future research. For example, if the results provide evidence that acceptable levels of efficacy are only attainable at the cost of unacceptable side effect burden, followup studies might include less burdensome treatment options.

Statistical inference for the estimated quantities in Figure 15.1 is another extension that merits attention. In the next section, we briefly describe the inference challenges that arise in this problem. Figure 15.1 would be much more informative with confidence bands around each of the four lines. For example, suppose at $\varkappa = 16$ that 95% confidence intervals overlap for the gray solid line and dashed line with open circles overlap. Similarly, suppose 95% confidence intervals overlap for the gray dashed line and solid line with closed circles at $\varkappa = 16$. This would suggest the estimated constrained regime with $\varkappa = 16$ may perform as well as both the unconstrained mean-optimal regime for Y and the unconstrained mean optimal regime for Z.

15.4 ▪ Future work

Using data to inform clinical decision making is complicated by multiple, possibly competing, outcomes. We presented constrained estimation as one approach for accommodating competing outcomes. Furthermore, the proposed framework is capable of handling features of the outcome distribution that are much more general than the mean, which has been the focus of the literature to date. This chapter was intended to serve as a sort of proof of concept with the hope that it might spark additional research in this direction. Here we outline a few extensions that we feel are important. First, theoretical developments including consistency and inference are needed. The constraints make this a nonregular problem, complicating inference so that standard methods for inference, e.g., bootstrap or series approximations, cannot be applied without modification [12]. Second, robust and computationally efficient methods are needed to implement the method in moderate or high dimensional settings. While our algorithm performed well in the setting considered, it would be interesting to explore how this method would scale with the number of constraints and the complexity of the functional determining the objective and constraints. A third direction for future work is to extend the proposed method to settings with many (e.g., more than two) time points. A challenge in this setting is building high quality yet parsimonious models for the distributions of the outcomes under arbitrary DTRs. We are currently working on some of these extensions.

15.5 ▪ Appendix: List of notation

$X_1 \in \mathbb{R}^{p_1}$	Baseline covariates.
$A_1 \in \{-1, 1\}$	First-stage randomized treatment.
$X_2 \in \mathbb{R}^{p_2}$	Covariates measured after assignment to first-stage treatment but prior to second-stage treatment randomization.
$H_2 \in \mathbb{R}^{p_1+1+p_2}$	Second-stage history vector that includes X_1, A_1, and X_2.
$A_2 \in \{-1, 1\}$	Second-stage randomized treatment.
$Y \in \mathbb{R}$	Primary outcome.
$Z \in \mathbb{R}$	Secondary outcome.

$d \in \mathscr{D}$	A pair of decision rules within the class $\mathscr{D}$ that is a function of observed first-stage covariates x_1 and second-stage history h_2, $d = \{d_1(x_1), d_2(h_2)\}$.
$\text{pr}^d_{Y,Z}(y,z)$	Joint distribution of (Y,Z) when treatments are assigned according to the regime d, evaluated at (y,z).
$\text{pr}^d_Y(y)$	Marginal distribution of Y under regime d, evaluated at y.
$\text{pr}^d_Z(z)$	Marginal distribution of Z under regime d, evaluated at z.
$m_Y(H_2)$	Function that models the main effect of H_2 on Y.
$c_Y(H_2)$	Function that models the interaction effect of H_2 with A_2 on Y.
$m_Z(H_2)$	Function that models the main effect of H_2 on Z.
$c_Z(H_2)$	Function that models the interaction effect of H_2 with A_2 on Z.
ϵ_Y	Mean-zero error term for the model of Y on H_2 and A_2.
ϵ_Z	Mean-zero error term for the model of Z on H_2 and A_2.
$F_{\epsilon_Y}(\cdot)$	Distribution function of ϵ_Y.
$F_{\epsilon_Z}(\cdot)$	Distribution function of ϵ_Z.
$F_{\epsilon_Y,\epsilon_Z}(\cdot,\cdot)$	Joint distribution of ϵ_Y and ϵ_Z.
$M(\text{pr}^d_{Y,Z})$	Univariate feature of $\text{pr}^d_{Y,Z}$ to be maximized.
$S(\text{pr}^d_{Y,Z})$	r-dimensional feature of $\text{pr}^d_{Y,Z}$ to be bounded above by $\varkappa \in \mathbb{R}^r$.
$F_{X_1}(\cdot)$	Cumulative distribution function of vector X_1.
$F_{X_2 \mid X_1,A_1}(\cdot \mid x_1,a_1)$	Conditional cumulative distribution function of vector X_2 given $X_1 = x_1$ and $A_1 = a_1$.
$d_{2,S}(h_2)$	S-constrained second-stage decision rule.
$f_2(H_2)$	Smooth function whose sign determines the S-constrained second-stage decision rule for history H_2.
$J^d_Y(x_1,x_2,y)$	Distribution of ϵ_Y evaluated at $y - m_Y\{x_1,d_1(x_1),x_2\} - \text{sign}[f_2\{x_1,d_1(x_1),x_2\}]c_Y\{x_1,d_1(x_1),x_2\}$.
$G^{d_2}_Y(\cdot,\cdot,\cdot \mid x_1,a_1)$	Joint conditional distribution of $m_Y(X_1,A_1,X_2)$, $c_Y(X_1,A_1,X_2)$, and $f_2(X_1,A_1,X_2)$ given $X_1 = x_1$ and $A_1 = a_1$.
$G^{d_2}_{Y,Z}(\cdot,\cdot,\cdot,\cdot,\cdot) \mid x_1,a_1)$	Joint conditional distribution of $m_Y(X_1,A_1,X_2)$, $c_Y(X_1,A_1,X_2)$, $m_Z(X_1,A_1,X_2)$, $c_Z(X_1,A_1,X_2)$, and $f_2(X_1,A_1,X_2)$ given $X_1 = x_1$ and $A_1 = a_1$.
$\mu^m_Y(X_1,A_1)$	Conditional mean of $m_Y(H_2)$ given X_1 and A_1.
$\mu^c_Y(X_1,A_1)$	Conditional mean of $c_Y(H_2)$ given X_1 and A_1.
$\mu^m_Z(X_1,A_1)$	Conditional mean of $m_Z(H_2)$ given X_1 and A_1.
$\mu^c_Z(X_1,A_1)$	Conditional mean of $c_Z(H_2)$ given X_1 and A_1.
$\mu_{f_2}(X_1,A_1)$	Conditional mean of $f_2(H_2)$ given X_1 and A_1.
$\sigma^m_Y(X_1,A_1)$	Square root of the conditional variance of $m_Y(H_2) - \mu^m_Y(X_1,A_1)$ given X_1 and A_1.
$\sigma^c_Y(X_1,A_1)$	Square root of the conditional variance of $c_Y(H_2) - \mu^c_Y(X_1,A_1)$ given X_1 and A_1.
$\sigma^m_Z(X_1,A_1)$	Square root of the conditional variance of $m_Z(H_2) - \mu^m_Z(X_1,A_1)$ given X_1 and A_1.
$\sigma^c_Z(X_1,A_1)$	Square root of the conditional variance of $c_Z(H_2) - \mu^c_Z(X_1,A_1)$ given X_1 and A_1.
e^m_Y	Standardized residual of $m_Y(H_2)$.
e^c_Y	Standardized residual of $c_Y(H_2)$.

e_Z^m	Standardized residual of $m_Z(H_2)$.
e_Z^c	Standardized residual of $c_Z(H_2)$.
e^{f_2}	Standardized residual of $f_2(H_2)$.

Acknowledgments
This work was supported the following grants: NIH P01 CA142538 (E. B. Laber); Department of Natural Resources grant PR-W-F14AF00171 (E. B. Laber); NSF grant DMS-1406456 (L. A. Stefanski); and NIH grants R01 CA085848 and P01 CA142538 (L. A. Stefanski).