PAC-Bayesian Treatment Allocation Under Budget Constraints*

Daniel F. Pellatt†

December 3, 2022
Link to latest version

Abstract

This paper considers the estimation of treatment assignment rules when the policy maker faces a general budget or resource constraint. Utilizing the PAC-Bayesian framework, we propose new treatment assignment rules that allow for flexible notions of treatment outcome, treatment cost, and a budget constraint. For example, the constraint setting allows for cost-savings, when the costs of non-treatment exceed those of treatment for a subpopulation, to be factored into the budget. It also accommodates simpler settings, such as quantity constraints, and doesn't require outcome responses and costs to have the same unit of measurement. Importantly, the approach accounts for settings where budget or resource limitations may preclude treating all that can benefit, where costs may vary with individual characteristics, and where there may be uncertainty regarding the cost of treatment rules of interest. Despite the nomenclature, our theoretical analysis examines frequentist properties of the proposed rules. For stochastic rules that typically approach budget-penalized empirical welfare maximizing policies in larger samples, we derive non-asymptotic generalization bounds for the target population costs and sharp oracle-type inequalities that compare the rules' welfare regret to that of optimal policies in relevant budget categories. A closely related, non-stochastic, model aggregation treatment assignment rule is shown to inherit desirable attributes.

JEL Classification: C01, C14, C44, C51
Keywords: Budget and quantity constraints, penalized empirical welfare, randomized experiments, treatment assignment, statistical learning.

1 Introduction

This paper proposes new statistical decision rules for treatment assignments under a general budget or resource constraint. A key objective in the empirical analysis of treatment data is identifying policies that result in the most beneficial outcomes. There is a large literature (e.g. Manski (2004) and Hirano and Porter (2009)) that examines how to determine which policies are optimal to implement in the absence of constraints such as one on policy cost. In practice, however, policy makers are rarely free from constraints when it comes to the policies they may enact. Several recent papers in the econometrics literature, including Kitagawa and Tetenov (2018), Athey and Wager (2021), and Mbakop and Tabord-Meehan (2021), consider the treatment

*The author is grateful to his advisor Yixiao Sun for feedback and support on this paper. Kaspar Wuthrich, Richard Carson, Ross Starr, and Nikolay Kudrin have also provided helpful comments and useful discussion.

†Department of Economics, University of California, San Diego. Address: 9500 Gilman Drive, La Jolla, CA 92093. Email: dpellatt@ucsd.edu

estimation problem from an empirical welfare maximization (EWM) perspective that allows for arbitrary constraints on the functional form of the decision rule. However, these papers do not address general budget constraints nor cost uncertainty that varies with the characteristics of individual agents. For example, while Kitagawa and Tetenov (2018) consider quantity constraints via random rationing, this treats costs as fixed and hence cannot identify which policies most efficiently balance cost vs. outcome trade-offs when costs vary with individual characteristics.

Here we focus on the setting where costs may be uncertain, current resource limitations may preclude treating all that can benefit, and where individual characteristics can influence treatment responses and costs. Compared to the unconstrained setting, the theoretically optimal treatment rule involves population objects that are more difficult to estimate and analyze in concert. For example, Bhattacharya and Dupas (2012) show that under a quantity constraint, which is simpler than the setting with variable costs, the optimal rule is to assign treatment when the conditional average treatment effect exceeds its $(1 - c)$th quantile. Here c is the maximal proportion of treatments assignable under the constraint. As a result, it can be difficult to evaluate properties of interest for proposed approaches and each existing approach has limitations.

The contributions of the paper are as follows. First, we propose new treatment rules that expand the tool set available to policy makers in the budget constrained setting. Second, we show they possess several potential benefits in terms of theoretical guarantees, the variety of settings in which they can be applied, and ease of estimation. Third, we show expert knowledge can be incorporated when the policy maker has non-data-dependent insights into the problem. However, the ability to integrate expert knowledge is a secondary feature of the approach. In our primary implementation we assume no such knowledge.

PAC-Bayesian analysis applies the probably approximately correct learning framework to objects of interest that involve probability distributions over model or parameter families. These objects can include, for example, treatment rules formed by aggregating over a family of potential rules. Our work can be seen as extending the PAC-Bayesian learning approach to the treatment setting in a way that incorporates a secondary cost objective. This motivates the proposed rules and allows us to derive generalization bounds for the costs and oracle-type inequalities for the welfare regret of proposed rules. Here, the welfare regret associated with a treatment rule is the loss in expected welfare of the decision rule relative to the theoretically optimal decision rule (cf. Manski (2004)). To work within the regret framework, we also derive the form of a theoretically optimal treatment policy if the data generating process (DGP) were known under a general budget constraint.

Individualized treatment policies under budget restrictions are of interest in a variety of settings. Often policy makers with limited resources face uncertainty regarding the costs and benefits of potential policies where this uncertainty is driven due to the fact that costs and benefits vary with the individual characteristics of those who decide to participate in a program. For example, Finkelstein et al. (2012) examine outcomes such as health care utilization and self-reported health measures following a randomized expansion of household access to Medicaid in Oregon. A policy maker may be interested in identifying policies to maximize a well-defined weighted average of such outcomes given a binding expenditure constraint. The government has control over eligibility rules defined on characteristics such as age, income, and the number of children in a household that directly influence expected cost and cost uncertainty.

Insecticide-treated nets (ITNs) for protection against malaria in regions of Africa represent another common example. Lengeler (1998), for instance, documents reductions in child mortality while Kuecken et al. (2014) document returns to education related to ITN provisions.

Teklehaimanot et al. (2007) estimate the cost of providing an ITN to every at-risk individual in sub-Saharan Africa to be 2.5 billion dollars. However, government and aid funding was below that level at the time of the study. Bhattacharya and Dupas (2012) look at a treatment policy estimator under quantity constraints derived from data from a randomized experiment assigning ITNs to rural households in Kenya. They use average costs in combination with government funding estimates to determine the quantity constraint. Our approach makes it possible to target policies in such a way as to account for cost heterogeneity (e.g. different distribution channels) and hence improve efficiency and achieve a higher overall outcome level.

Beyond aid and social safety net policies, the budget constrained treatment assignment problem can also arise in a commercial context for firms considering potentially costly promotions aimed at obtaining new customers. For instance, Sun et al. (2021) recently proposed a budget constrained treatment estimator aimed at determining which customers should be offered trial access to a premium service. They seek to use customers' individual characteristics to discriminate against making offers to customers likely to heavily utilize the service in the trial (high cost) while being unlikely to use the service after the trial period expires. Rather than the simple notion of not wanting to implement a policy that leads to long-term losses, many companies will also face a short-term constraint on how much they can "lose" in the trial phase to gain market share. For other firms, like Uber which is considered in Sun et al. (2021), a deeper issue may arise. Increasing sales or trial offers may fundamentally alter the firm's cost structure (e.g., increasing driver compensation to induce enough new drivers to work to handle the increased number of trips).

The rules we develop start from a user-specified family of (non-stochastic) treatment models $\mathcal{F}$ that map an individual's covariates that are observable pre-treatment to the $\{0, 1\}$ treatment indicator space. Rather than choosing the model that maximizes the empirical welfare in $\mathcal{F}$, for example, we instead consider stochastic treatment rules derived from $\mathcal{F}$ and a measure of budget penalized empirical welfare. Given an individual's pre-treatment covariates, their treatment probability is calculated as an exponentially weighted average over the treatments specified by members of $\mathcal{F}$. The treatment probability is similar to a weighted majority vote taken over $\mathcal{F}$. The exponential weighting received by members of the model family is greatest for models with a large budget-penalized empirical welfare. The magnitude of the penalization term related to cost is determined by a parameter u that modulates the trade-off between maximizing welfare and reducing costs. Any choice for u will correspond to a different maximal empirical budget, with $u = 0$ corresponding to an unlimited budget (no constraint). Typically, for larger sample sizes, the rule is unlikely to assign identical covariates to different treatments unless there are subsets of the model family with similarly high values of penalized welfare that prescribe different treatments. We also consider closely related, non-stochastic, model aggregation treatment rules that aggregate over $\mathcal{F}$ to make treatment decisions.

Utilizing a PAC-Bayesian framework, under reasonable conditions we show that for a set of u values, in large samples, with high probability we obtain increasingly accurate estimates of the target population costs associated with corresponding stochastic treatment rules. We can use these estimates to select u or, alternatively, u can be chosen via cross validation. At the same time, with u chosen in either manner, with high probability the resulting rule achieves a welfare regret comparable to that of the best models in the model family that have a similar target population cost. Starting from a set of budget penalty parameters, the policy maker can trace out good estimates of the feasible target population budgets, select the parameter associated with one of these estimates, and obtain a treatment rule with desirable regret properties. Regarding the non-stochastic, model aggregation treatment rules, we show that they inherit desirable properties from the stochastic rules. We also consider the setting where u is chosen to meet a predetermined target population budget level. The procedure in this case is still reasonably motivated, as the rule minimizes an upper bound on the target population regret among rules that satisfy an empirical budget constraint. However, the generalization bounds for the target population cost and the oracle-type inequalities in this case become more complex to interpret.

The remainder of the paper is organized as follows. Section 2 discusses related literature and papers with alternative budget constrained treatment estimators. Section 3 details the statistical setting, treatment model formulation, and initial properties useful for later results. Section 4 provides theoretical motivation for the proposed treatment rules, utilizing the PAC-Bayesian analysis framework to examine (frequentist) properties of the proposed rules. Section 5 conducts a simulation experiment and discusses implementation and estimation. Section 6 concludes.

2 Related Literature

The topic of budget constrained treatment allocation is the subject of a small but growing literature. Sun et al. (2021) and Wang et al. (2018) empirically implement treatment rules starting from the notion of a theoretically optimal rule. They estimate unknown population level objects that appear in the optimal rules and then plug in the empirical counterparts to the corresponding theoretical formulas to obtain rules. The standard drawback of this sort of approach is that the estimation technique doesn't directly target policies that maximize the welfare problem of interest. For example, the regressions utilized to fit the conditional average treatment and cost functions in Wang et al. (2018) might yield parameters that are most accurate in regions of the covariate space that are less important for distinguishing individuals with a high outcome-to-cost ratios in the population. Wang et al. (2018) also consider a second method that shares similarities with the approach taken by Huang and Xu (2020). These approaches add the budget constraint to the outcome-weighted treatment learning approach considered, for example, in Zhao et al. (2012). These approaches work from optimization problems that directly target an empirical version of the problem of interest.

One drawback of the aforementioned techniques is a lack of theoretical insight regarding the true target population cost and risk attributes of the proposed rules. Sun (2021) adapts the EWM setting of Kitagawa and Tetenov (2018) to account for a general budget constraint. She considers a conservative rule that will satisfy the budget constraint asymptotically. She also considers a modified rule where a Lagrange multiplier parameter is capped during estimation. This will, asymptotically, approach the welfare of the budget constrained welfare maximizing policy among the user-specified model class. This methodology extends the arbitrary form features of EWM to the budget constraint setting. However, the rules involve a non-convex estimation procedure that may become difficult if the model class includes more flexible functional forms. While our methodology sacrifices some ability to satisfy functional form constraints due to its stochastic nature, one benefit is that we can take advantage of Bayesian estimation machinery as discussed in Section 5. Lastly, although the modified rules of Sun (2021) will approach the optimal rule within the original budget constraint, it is worth noting that the modified rule may violate that budget constraint. One benefit of our approach is we can compare our rules to those with the highest welfare among rules with the same target population cost as the proposed rules.

In a broader context, this paper contributes to a growing literature on statistical treatment rules in econometrics, including Manski (2004), Dehejia (2005), Hirano and Porter (2009), Bhattacharya and Dupas (2012), Kitagawa and Tetenov (2018), Viviano (2019), and Athey and Wager (2021). This literature has overlap with additional fields including statistics and machine learning. For examples, see Qian and Murphy (2011) and Beygelzimer and Langford (2009), respectively. Additional references and a discussion of the links between these fields can be found in Athey and Wager (2021). We also note that, in the machine learning literature, London and Sandler (2019) utilize a PAC-Bayesian approach to policy estimation for the logged bandit feedback problem which is closely related to treatment policy estimation.

Lastly, our analysis and proposed treatment rules are heavily influenced by the PAC-Bayesian machine learning literature. Seminal works in this area include Shawe-Taylor and Williamson (1997), McAllester (1999b), McAllester (1999a), Seeger (2002), and McAllester (2003b). In particular, we utilize techniques stemming from Catoni (2007), Lever et al. (2010), Maurer (2004), Germain et al. (2015), and Alquier et al. (2016). The theoretical contribution of our paper is, first, to modify and adapt relevant tools and generalization bounds to the treatment choice setting. We also develop the incorporation of a secondary objective or loss function (the treatment cost cost) into the analysis that yields informative oracle-type inequalities and generalization bounds relevant to the constrained budget setting.

3 Setup and Assumptions

3.1 Statistical setting and policy maker's problem

We consider the setting where a policy maker has data consisting of observations $Z_i = (Y_i, C_i, D_i, X_i)$ for $i = 1, \ldots, n$. Here, $X_i \in \mathcal{X} \subset \mathbb{R}^{d_x}$, where $d_x \in \mathbb{N}$, denotes a vector of covariates for individual or unit i observed prior to treatment assignment, $Y_i \in \mathbb{R}$ is unit i's outcome that is observed after treatment assignment, $C_i \in \mathbb{R}$ is the cost incurred and $D_i \in \{0, 1\}$ is a treatment assignment indicator that is 1 if unit i was assigned the treatment and is zero otherwise. C_i may be uncertain at the time of treatment assignment and is allowed to be observed after treatment assignment.

To account for heterogeneous treatment responses and costs, we work from a potential outcomes and costs framework. For unit i and for $j \in \{0, 1\}$, let $Y_{i,j}$ and $C_{i,j}$ denote the outcome and cost, respectively, that would have been observed if unit i had been assigned $D_i = j$. Ignoring the index i, we can relate the observed outcome and cost to their potential outcomes and costs by writing

$$Y = Y_1 D + Y_0 (1 - D), \quad C = C_1 D + C_0 (1 - D). \tag{1}$$

The following assumption formalizes this setting. It also includes conditions needed to identify properties related to potential outcomes and costs when they are not observed directly in sample data.

Assumption 3.1 *(i) Random Sample: Let Q be the joint distribution of $(Y_0, Y_1, C_0, C_1, D, X)$, where $Y_0, Y_1, C_0, C_1 \in \mathbb{R}$, $D \in \{0, 1\}$, $X \in \mathcal{X} \subseteq \mathbb{R}^{d_x}$. Let $Z = (Y, C, D, X) \in \mathcal{Z}$ be distributed according to P where P is determined by Q and (1). We assume the sample $S = \{Z_i\}_{i=1}^n \sim P^{\otimes n}$ is a size n i.i.d. sample[1]. We denote the sample space $S \in \mathcal{S} = \mathcal{Z}^n$.*

(ii) Unconfoundedness: $(Y_1, Y_0, C_1, C_0) \perp D | X$.

[1] To denote the probability of an event A under this sampling distribution, we will use the notation $P^n(A)$. To denote the probability of an event B under the distribution P, we write $P(B)$.

(iii) Bounded Outcomes and Costs: There exist positive $M_y, M_c < \infty$ such that the support of Y is contained in $[-M_y/2, M_y/2]$ and the support of C is contained in $[-M_c/2, M_c/2]$.

(iv) Strict overlap: Define $e(X) = E_P[D|X]$, where $E_P(\cdot)$ is the expectation with respect to P.[2] It is assumed that there exists $\kappa \in (0, 1/2)$ such that $e(x) \in [\kappa, 1-\kappa]$ for all $x \in \mathcal{X}$.

Assumption 3.1 mirrors treatment assumptions in Kitagawa and Tetenov (2018) and Mbakop and Tabord-Meehan (2021) and also includes similarly-formulated conditions for cost-related variables. Unconfoundedness states that, conditional on the covariates, the potential outcomes and costs are independent of the treatments assigned to the observed data. This and strict overlap will hold in randomized controlled trials (RCTs) which is our primary setting of interest. As such, we assume $e(x)$ is known. It is possible to adjust our procedures to a setting where $e(x)$ is estimated similarly to the e-hybrid rules utilized in Kitagawa and Tetenov (2018) and Mbakop and Tabord-Meehan (2021) while maintaining some of the theoretical motivations considered in Section 4. We leave a complete exploration of this topic to future research and work under the presumption that $e(x)$ is known.

Define the conditional average treatment effect (CATE) and the conditional average treatment cost (CATC), respectively, by

$$\delta_y(x) \equiv E_Q[Y_1 - Y_0|X = x], \quad \delta_c(x) \equiv E_Q[C_1 - C_0|X = x]. \tag{2}$$

Assumption 3.1 (iii) implies that $|\delta_y(X)|$ and $|\delta_c(X)|$ are bounded almost surely by M_y and M_c, respectively. Our procedures can be implemented without knowledge of M_y or M_c and several of the motivating regret bounds in Section 4 could be derived in slightly altered forms if instead we required that objects related to $|\delta_y(X)|$ and $|\delta_c(X)|$ are sub-Guassian or even sub-exponential with additional constraints on a hyper-parameter. Assumption 3.1 (iii) is typically a mild requirement that is often adopted in the treatment and classification literature; here it simplifies our exposition and path to generalization bounds. Note that Y and C may belong to any interval. The upper and lower bounds are taken to be symmetric around zero for convenience and without loss of generality.

In section 3.2 we propose treatment assignment rules that aim to balance two prevailing objectives. We seek rules that will maximize the expected outcome Y while also accounting for a potential budget constraint when we anticipate that resource, policy, or other limitations may preclude treating everyone with a positive CATE. Our proposed rules contain a parameter u, which can be chosen in a data-dependent manner, that modulates how much the second (budgetary) objective is prioritized. In particular, any choice of u corresponds to a different maximum expected cost in a budget-constrained welfare optimization problem. Before describing the treatment model and empirical approach, we first state the policy maker's problem at the population level under a given maximum budget B if the distribution Q were known.

The policy maker's goal is to obtain a treatment rule that maximizes welfare subject to a budget or quantity constraint. The treatment rule is intended for application to a target population wherein the joint distribution of (Y_0, Y_1, C_0, C_1, X) follows that associated with Q. We will consider stochastic treatment assignment rules, defining such a rule as a measurable map $f : \mathcal{X} \to [0, 1]$ from the covariate space to a treatment assignment probability. If $f(x) \in \{0, 1\}$, the treatment assignment for x is non-random. If $0 < f(x) < 1$, treatment is assigned randomly with treatment probability $f(x)$.

[2]Similarly, we denote expectation with respect to Q by $E_Q(\cdot)$. Expectation with respect to the distribution of the sample, $P^{\otimes n}$, will be denoted $E_{P^n}(\cdot)$.

The utilitarian welfare associated with f is given by

$$E_Q[Y_1 f(X) + Y_0(1 - f(X))]. \tag{3}$$

This is the expected value of Y when treatment is administered according to $f(X)$. Dropping terms that do not vary with f, the policy maker's objective function evaluated at f is defined by

$$W(f) \equiv E_Q[(Y_1 - Y_0)f(X)]. \tag{4}$$

Choosing f that maximizes $W(f)$ is equivalent to choosing f that maximizes utilitarian welfare. Thus we will refer to $W(f)$ as the welfare associated with f. Note that by the law of iterated expectations, $W(f) = E_Q[\delta_y(X)f(X)]$. Next, define the expected cost of f by

$$K(f) \equiv E_Q\left[(C_1 - C_0)f(X)\right], \tag{5}$$

which can similarly be written $K(f) = E_Q[\delta_c(X)f(X)]$. Given a budget constraint B, the policy maker's problem is to identify

$$f_B^* \in \arg\max_f \left\{W(f) : K(f) \leq B\right\}, \tag{6}$$

where the maximization is taken over all measurable functions from $\mathcal{X}$ to $[0,1]$.

Note that $K(f) = E_Q[C_1 f(X) + C_0(1 - f(X))] - E_Q[C_0]$. The budget constraint states that the expected additional cost due to implementing treatment policy f, that beyond what would be expected if treatment were never assigned, cannot exceed B. This is flexible, as it allows for cost savings (i.e. when $C_1 < C_0$ with positive probability) to be factored into the budget. Provided such savings are possible, a policy maker could be interested in, for example, $B = 0$. In this scenario the policy maker is looking for treatment policies that may improve welfare without increasing the expected cost beyond the setting were no treatments are administered. On the other hand, if the policy maker has a fixed budget allocated to treatments and cost savings do not feed back into the budget, one can simply define $C_0 = 0$, so that the observed C is equal to the cost of treatment when treatment is provided and is zero otherwise. If there is a a fixed quantity constraint consisting of a set number of treatments and no other budgetary concerns, one can set $C_0 = 0$ and $C_1 = 1$ so that the observed C is the treatment indicator. In this case B denotes the maximum proportion of the target population for which treatments are available.

If there is no budget constraint and the policy maker is able to choose any measurable $f : \mathcal{X} \to [0,1]$, it is straightforward to verify that an optimal treatment allocation rule is given by

$$f^*(x) = 1\{\delta_y(x) > 0\}. \tag{7}$$

f^* assigns treatment to any unit with a positive CATE. Here, and throughout the paper, the indicator function $1\{A\}$ takes the value 1 if event A occurs and is zero otherwise. Given a particular budget constraint B, a solution to the policy maker's problem is characterized in the following theorem.

Theorem 3.1 *Let (Y_0, Y_1, C_0, C_1, X) be distributed according to Q. Assume that $E_Q|\delta_y(X)| < \infty$, $E_Q|\delta_c(X)| < \infty$, and that $B > E_Q[\delta_c(X)1\{\delta_c(X) < 0\}]$. Then there exist constants $\eta_B \geq 0$ and $a_1, a_2 \in [0,1]$ such that*

$$f_B^*(x) = \begin{cases} 0 & \text{if } \delta_y(x) < \eta_B \delta_c(x), \\ a_1 1\{\delta_c(x) > 0\} + a_2 1\{\delta_c(x) < 0\} & \text{if } \delta_y(x) = \eta_B \delta_c(x), \\ 1 & \text{if } \delta_y(x) > \eta_B \delta_c(x), \end{cases} \tag{8}$$

satisfies (6). In particular, if $K(f^*) \leq B$, then one can take $\eta_B = a_1 = a_2 = 0$ and $f_B^* = f^*$; if $K(f^*) > B$ then (η_B, a_1, a_2) are chosen such that $K(f_B^*) = B$. If $E_Q[1\{\delta_y(X) = \eta_B \delta_c(X)\}] = 0$, f_B^* is deterministic and is the unique budget-constrained, welfare-optimizing policy in the sense that for any f' satisfying (6) it holds that $f'(X) = f_B^*(X)$ a.s.*

The choice of η_B in Theorem 3.1 is unique, however in general there may be different choices of a_1, a_2 that produce optimal rules when $E_Q[1\{\delta_y(X) = \eta_B \delta_c(X)\}] \neq 0$. Apart from this difference, Theorem 3.1 is a generalization of a result in Sun et al. (2021) which restricts itself to the setting where $C_1 \geq C_0$ almost surely. In practice, of course, Q is unknown to the researcher who must estimate a suitable model f empirically. Section 3.2 introduces the PAC-Bayesian setting for the empirical strategy we employ.

When $E_Q[1\{\delta_y(X) = \eta_B \delta_c(X)\}] = 0$, for example when $\delta_y(X)$ and $\delta_c(X)$ have bounded densities, Theorem 3.1 says the optimal treatment rule is deterministic and unique in terms of the resulting treatment decisions. However, the function $\delta_y(x) - \eta_B \delta_c(x)$ in the optimal rule in this setting, given by

$$f_B^*(x) = 1\{\delta_y(x) - \eta_B \delta_c(x) > 0\},$$

is not unique. Any measurable function $m(x) : \mathcal{X} \to \mathbb{R}$ that satisfies

$$\operatorname{sign}[m(x)] = \operatorname{sign}[\delta_y(x) - \eta_B \delta_c(x)],$$

yields an optimal treatment rule via $f_m(x) = 1\{m(x) > 0\}$. This situation is similar to that in the binary forecasting problem (cf. Elliott and Lieli (2013)) and is illustrated in Figure 1.

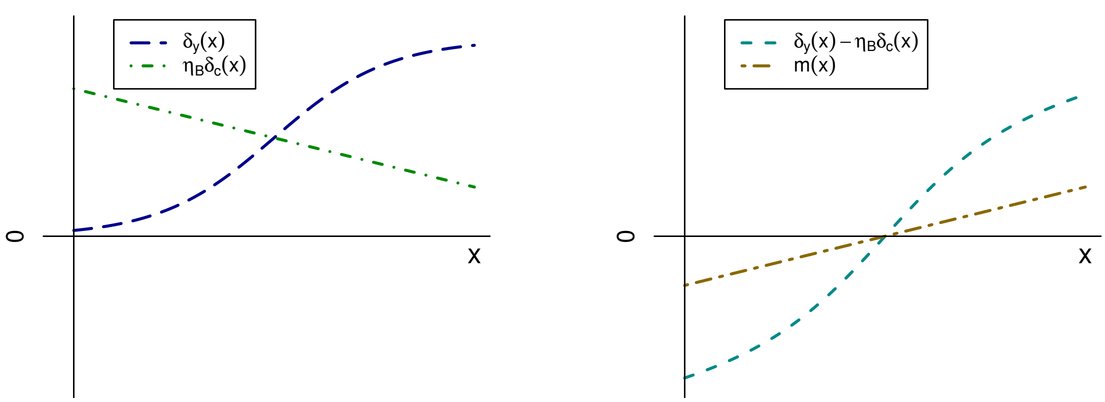

Figure 1: On the left, a plot of $\delta_y(x)$ and $\eta_B \delta_c(x)$ in a simple setting with a single crossing point and single explanatory variable. On the right, the corresponding $\delta_y(x) - \eta_B \delta_c(x)$ is plotted along with a second function, $m(x)$. Here, $m(x)$ differs from $\delta_y(x) - \eta_B \delta_c(x)$ everywhere except at the crossing point yet $1\{m(x) > 0\}$ and $1\{\delta_y(x) - \eta_B \delta_c(x)\} = f_B^*(x)$ yield identical treatment decisions.

In Section 3.2, we propose treatment rules that aggregate over a user-specified family of treatment rules in a way that is weighted towards models with high empirical budget-penalized welfare. There, we introduce Gibbs treatment rules, which aggregate over the rule family to derive a treatment probability, and related majority vote rules which aggregate over the rule family to assign treatment directly. Aside from the desirable theoretical properties derived in Section 4, some intuition behind such an approach is as follows. Two functions $\hat{m}(x)$ and $\hat{m}^*(x)$, with corresponding treatment rules $1\{\hat{m}(x) > 0\}$ and $1\{\hat{m}^*(x) > 0\}$, respectively, could yield identical or very similar treatment decisions over the sample covariate values. In a setting where different rules may have the same or very similar observable properties, it is reasonable to aggregate or average over rules with high empirical welfare. Rather than trying to select a single solution, we take the identification issue above as motivation for an ensemble approach.

3.2 Empirical Approach and PAC-Bayesian Setting

Underpinning the treatment rules we will consider is a family of non-stochastic treatment rules, indexed by $\theta \in \Theta$, denoted

$$\mathcal{F}_\Theta = \{f_\theta(x) : \mathcal{X} \to \{0, 1\}; \theta \in \Theta\}. \tag{9}$$

For a concrete example, we could let $\{\phi_1(x), \dots, \phi_q(x)\}$ be a set of feature transformations where $\phi_j(x) : \mathcal{X} \to \mathbb{R}$ for $j = 1, \dots, q$. Denoting $\phi(x) = (\phi_1(x), \dots, \phi_q(x))^\intercal$, we could then have

$$f_\theta(x) = 1\{\phi(x)^\intercal\theta > 0\} \text{ for } \theta \in \Theta = \mathbb{R}^q, \tag{10}$$

where $q \in \mathbb{N}$ need not be equal to d_x, the dimension of $\mathcal{X}$.

For any treatment assignment rule f, we define the welfare regret relative to the first-best prediction rule f^* in (7) by

$$R(f) \equiv W(f^*) - W(f).$$

Note that $R(f)$ is defined relative to the first-best treatment assignment without a budget constraint. We can also define

$$R_B(f) \equiv W(f_B^*) - W(f), \tag{11}$$

the welfare-regret under a maximum expected budget of B where f_B^* is defined in Theorem 3.1. With simple manipulations, the oracle-type inequalities involving $R(f)$ in Sections 4.1 and 4.2 apply to $R_B(f)$ rather than $R(f)$. For simplicity, we will mostly work with $R(f)$ which is non-negative. Note that $R_B(f)$ is only non-negative when attention is constrained to treatment rules with a maximal budget B. For particular models $f_\theta \in \mathcal{F}_\Theta$, with a slight abuse of notation, we will write

$$R(\theta) \equiv R(f_\theta), \; W(\theta) \equiv W(f_\theta), \text{ and } K(\theta) \equiv K(f_\theta).$$

Under the unconfoundedness and strict overlap conditions of Assumption 3.1, it holds that

$$W(f) = E_Q[(Y_1 - Y_0) f(X)] = E_P\left[\left(\frac{YD}{e(X)} - \frac{Y(1-D)}{1-e(X)}\right) f(X)\right].$$

A similar statement can be written for $K(f)$, now with C in place of Y. Defining

$$\delta_{y,i} = \left(\frac{Y_iD_i}{e(X_i)} - \frac{Y_i(1-D_i)}{1-e(X_i)}\right) \text{ and } \delta_{c,i} = \left(\frac{C_iD_i}{e(X_i)} - \frac{C_i(1-D_i)}{1-e(X_i)}\right),$$

the (ubiased) empirical counterparts of $W(f)$, $R(f)$, and $K(f)$, along with their notation for $f_\theta \in \mathcal{F}_\Theta$, are given by

$$W_n(f) \equiv \frac{1}{n}\sum_{i=1}^{n} \delta_{y,i} f(X_i), \qquad W_n(\theta) \equiv \frac{1}{n}\sum_{i=1}^{n} \delta_{y,i} f_\theta(X_i),$$

$$R_n(f) \equiv \frac{1}{n}\sum_{i=1}^{n} \delta_{y,i}\left(f^*(X_i) - f(X_i)\right), \qquad R_n(\theta) \equiv \frac{1}{n}\sum_{i=1}^{n} \delta_{y,i}\left(f^*(X_i) - f_\theta(X_i)\right),$$

$$K_n(f) \equiv \frac{1}{n}\sum_{i=1}^{n} \delta_{c,i} f(X_i), \qquad K_n(\theta) \equiv \frac{1}{n}\sum_{i=1}^{n} \delta_{c,i} f_\theta(X_i).$$

As f^* is unknown, the empirical regret $R_n(f) = W_n(f^*) - W_n(f)$ or $R_n(\theta)$ for $\theta \in \Theta$ cannot be evaluated in practice. $R_n(\theta)$ will arise in our analysis only as a theoretical object in relation to $R(\theta)$. We stress that the treatment assignment rules we consider can be expressed solely in terms of $W_n(\theta)$.

$\mathcal{F}_\Theta$ consisting of treatment rules of the form in (10) will be considered in Sections 4.2 and 5. In general, to accommodate broader treatment rule model families, we make the following technical assumptions.

Assumption 3.2 *(i) We assume that $(\Theta, \mathcal{B}_\theta)$ is a standard Borel space. (ii) We assume that $\mathcal{F}_\Theta$ is such that the maps $(S,\theta) \mapsto R_n(\theta) : \mathcal{S} \times \Theta \to \mathbb{R}$ and $(S,\theta) \mapsto K_n(\theta) : \mathcal{S} \times \Theta \to \mathbb{R}$ are measurable.*

We now introduce the stochastic treatment rules of interest. Let $\mathcal{P}(\Theta)$ be the set of probability measures on $(\Theta, \mathcal{B}_\theta)$ and, for any $\pi \in \mathcal{P}(\Theta)$, let $\mathcal{P}_\pi(\Theta) = \{\rho \in \mathcal{P}(\Theta) : \rho \ll \pi\}$. That is, $\mathcal{P}_\pi(\Theta)$ is the set of probability measures on $(\Theta, \mathcal{B}_\theta)$ that are absolutely continuous with respect to π. Rather than selecting a single value $\hat{\theta} \in \Theta$, for example that which maximizes $W_n(\theta)$, and then assigning treatment via $f_{\hat{\theta}}$, we seek probability measures $\rho \in \mathcal{P}(\Theta)$ from which we form stochastic treatment rules. Borrowing nomenclature from the classification literature, we work with Gibbs treatment rules. For $\rho \in \mathcal{P}(\Theta)$, the Gibbs treatment rule or method associated ρ, denoted $f_{G,\rho} : \mathcal{X} \to [0,1]$, is defined by

$$f_{G,\rho}(x) = \int_\Theta f_\theta(x) d\rho(\theta), \; x \in \mathcal{X}.$$

Assigning treatments via the Gibbs method is equivalent to assigning treatments as follows. For an individual with covariates X, a parameter value $\theta_\circ$ is drawn randomly according to ρ, i.e. $\theta_\circ \sim \rho$. Then, $f_{\theta_\circ}(X) \in \{0,1\}$ determines the treatment assignment. This process, with an independent draw from ρ, is repeated each time treatment is to be assigned. Note that, exchanging the order of integration, we can write

$$R(f_{G,\rho}) = \int_\Theta R(\theta) d\rho(\theta) \;\text{ and }\; R_n(f_{G,\rho}) = \int_\Theta R_n(\theta) d\rho(\theta),$$

which is called the Gibbs risk associated with ρ. Similarly, the expected cost of $f_{G,\rho}$ and its empirical counterpart can be written

$$K(f_{G,\rho}) = \int_\Theta K(\theta) d\rho(\theta) \;\text{ and }\; K_n(f_{G,\rho}) = \int_\Theta K_n(\theta) d\rho(\theta).$$

We will frequently be concerned with the cost or empirical cost associated with a Gibbs treatment rule utilizing some $\rho \in \mathcal{P}_\pi(\Theta)$. To simplify the exposition, we denote

$$B(\rho) = K\left(f_{G,\rho}\right), \quad \text{and} \quad \widehat{B}(\rho) = K_n\left(f_{G,\rho}\right). \tag{12}$$

A non-stochastic treatment rule that is closely related to the Gibbs rule is the so-called majority vote or Bayes method associated with $\rho \in \mathcal{P}(\Theta)$. This is given by

$$f_{\mathrm{mv},\rho}(x) = 1\left\{\int_\Theta f_\theta(x) d\rho(\theta) > \frac{1}{2}\right\}, \; x \in \mathcal{X}. \tag{13}$$

In practice, majority vote rules can deliver treatment rules that are numerically more stable than their Gibbs counterpart. If $\rho = \alpha\rho_1 + (1-\alpha)\rho_2$ for some $\rho_1, \rho_2 \in \mathcal{P}(\Theta)$ and constant α, then $R(f_{G,\rho}) = \alpha R(f_{G,\rho_1}) + (1-\alpha)R(f_{G,\rho_2})$. That is, the Gibbs risk is a linear functional of ρ. This linearity makes the Gibbs risk and Gibbs treatment rules more amenable to theoretical analysis. Our analysis will therefore focus on a family of Gibbs treatment rules. However, in Section 4.3, we show that the majority vote treatment rule associated with our Gibbs rules of interest inherit desirable properties from their Gibbs counterparts. In practice, either method is an acceptable choice and we consider both in our simulation study in Section 5.

In particular, we propose to utilize Gibbs treatment rules constructed from data-dependent[3] probability measures of the form $\hat{\rho}_{\lambda,u}$ defined below.

Definition 3.2 *For $\lambda > 0$, $u \geq 0$, and a reference measure $\pi \in \mathcal{P}(\Theta)$, define $\hat{\rho}_{\lambda,u}$ to be the (random) probability measure on Θ with the following Radon-Nikodym (RN) derivative with respect to π:*

$$\begin{aligned}\frac{d\hat{\rho}_{\lambda,u}}{d\pi}(\theta) &= \frac{\exp\left[-\lambda\left(R_n(\theta) + uK_n(\theta)\right)\right]}{\int_\Theta \exp\left[-\lambda\left(R_n\left(\tilde{\theta}\right) + uK_n\left(\tilde{\theta}\right)\right)\right] d\pi\left(\tilde{\theta}\right)} \\ &= \frac{\exp\left[-\lambda\left(uK_n(\theta) - W_n(\theta)\right)\right]}{\int_\Theta \exp\left[-\lambda\left(uK_n\left(\tilde{\theta}\right) - W_n\left(\tilde{\theta}\right)\right)\right] d\pi\left(\tilde{\theta}\right)}.\end{aligned}$$

Define $\rho^*_{\lambda,u}$ to be the probability measure on Θ with the following RN derivative with respect to π:

$$\frac{d\rho^*_{\lambda,u}}{d\pi}(\theta) = \frac{\exp\left[-\lambda\left(R(\theta) + uK(\theta)\right)\right]}{\int_\Theta \exp\left[-\lambda\left(R\left(\tilde{\theta}\right) + uK\left(\tilde{\theta}\right)\right)\right] d\pi\left(\tilde{\theta}\right)}.$$

$\hat{\rho}_{\lambda,u}$ is sometimes called a Gibbs posterior distribution or a Boltzmann distribution. As $\lambda \to \infty$, $\hat{\rho}_{\lambda,u}$ concentrates around the value of θ such that f_θ minimizes the budget-penalized empirical regret criterion $R_n(f_\theta) + uK_n(f_\theta)$. Equivalently, it concentrates around the value of θ the maximizes $W_n(\theta) - uK_n(\theta)$ over Θ. This reduces to the empirical welfare maximizer when $u = 0$. In general, $\hat{\rho}_{\lambda,u}$ assigns higher probability to regions of the parameter or model space with low budget-penalized empirical regret. u modulates the trade off between emphasis on

[3]In general, by data-dependent probability measures on $(\Theta, \mathcal{B}_\theta)$ we mean regular conditional probability measures (RCPMs): letting $\mathcal{B}_s$ denote the σ-algebra associated with the sample space $\mathcal{S}$, $\rho(S, \cdot)$ is an RCPM on $(\Theta, \mathcal{B}_\theta)$ if (i) for any fixed $A \in \mathcal{B}_\theta$, the map $S \mapsto \rho(S, A) : (\mathcal{S}, \mathcal{B}_s) \to \mathbb{R}_+$ is measurable; and (ii) for any $S \in \mathcal{S}$, the map $A \mapsto \rho(S, A) : \mathcal{B}_\theta \to [0, 1]$ is a probability measure. For additional measure-theoretic details, for example the decomposition and measurability of the Kullback-Leibler divergence (utilized throughout the paper) between RCPMs, we refer the reader to Catoni (2004), in particular Proposition 1.7.1 and its proof on pages 50-54.

low regret vs expected cost. As subsequent analysis will show, different choices of u correspond in a one-to-one manner with different budget constraints. We will consider the setting where u is cross validated and the setting where it is determined by a particular choice of a budget constraint parameter B. $\rho^*_{\lambda,u}$ is a theoretical counterpart to $\hat{\rho}_{\lambda,u}$ that will be useful when we analyze statistical properties related to $\hat{\rho}_{\lambda,u}$. λ is typically chosen via cross validation while choices where $\lambda = \mathcal{O}(\sqrt{n})$ will yield optimal or near-optimal rates of convergence in Section 4.

In the PAC-Bayesian literature, probability measures over the model or parameter space that are traditionally chosen independently of the sample are often called prior probability measures. In our setting, the choice of π utilized in Definition 3.2 will fall into this category. Probability measures utilized for treatment or prediction, such as $\hat{\rho}_{\lambda,u}$, are called posterior distributions. However, this nomenclature does not have the same connotation as in traditional Bayesian methodology. While knowledge of the DGP could allow for a prior to be chosen that improves the performance of rules suggested from PAC-Bayesian analysis, often the prior is taken to be uniform or normal centered at the origin. Additionally, the posterior, for example, does not need to be proportional to a likeilihood function. The statistical analysis itself is frequentest in nature. The role and choice of π will be discussed further later in the paper. For now we make the following assumption.

Assumption 3.3 $\pi \in \mathcal{P}(\Theta)$ *is a (deterministic) probability measure that does not depend on the sample.*

3.3 Initial properties of the Gibbs posterior

Here we derive initial properties of $\hat{\rho}_{\lambda,u}$ that link the choice of u to a particular budget constraint. These provide intuition behind Definition 3.2 and are utilized in proving the results of Section 4.

Let $D_{\text{KL}}(\rho, \pi)$ denote the Kullback–Leibler (KL) divergence between $\rho, \pi \in \mathcal{P}(\Theta)$,

$$D_{\text{KL}}(\rho, \pi) = \begin{cases} \int_\Theta \log\left[\frac{d\rho}{d\pi}(\theta)\right] d\rho(\theta), & \text{if } \rho \ll \pi \\ \infty, & \text{else.} \end{cases}$$

Suppose the policy maker has a maximum expected budget of $B \in \mathbb{R} \cup \{\infty\}$, where $B = \infty$ is the unconstrained setting. If the data generating process were known, among Gibbs treatment rules we would be interested in a solution to

$$\min_{\rho \in \mathcal{P}(\Theta)} \int_\Theta R(\theta) d\rho(\theta), \text{ subject to } \int_\Theta K(\theta) d\rho(\theta) \le B. \tag{14}$$

In practice, we will instead focus on a subset $\mathcal{P}_\pi(\Theta) \subset \mathcal{P}(\Theta)$ and solve the following empirical problem:

$$\min_{\rho \in \mathcal{P}_\pi(\Theta)} \left[\int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\text{KL}}(\rho, \pi)\right], \text{ subject to } \int_\Theta K_n(\theta) d\rho(\theta) \le B. \tag{15}$$

(15) includes a regularization term in the form of $D_{\text{KL}}(\rho, \pi)$, discouraging any choice for ρ that has a large KL divergence from the reference measure π. In practice, $\mathcal{P}_\pi(\Theta)$ is flexible and optimal choices for λ will entail $\lambda \to \infty$ as $n \to \infty$. When adapted to our setting, Lemma 3.1 below shows that, provided a feasibility or Slater condition holds, for some value $\hat{u} \ge 0$, $\hat{\rho}_{\lambda,\hat{u}}$ is the solution to (15). Of course, appearing to be a reasonable empirical counterpart of (14) is not, in and of itself, justification for $f_{G,\hat{\rho}_{\lambda,u}}$. In Section 4 we provide additional theoretical motivation for $f_{G,\hat{\rho}_{\lambda,u}}$, comparing it to alternative Gibbs rules and optimal (non-stochastic) models in $\mathcal{F}_\theta$.

Lemma 3.1 *Let $\pi \in \mathcal{P}(\Theta)$, $\lambda > 0$, and let $A(\theta)$ and $H(\theta)$ be bounded, measurable functions defined on $(\Theta, \mathcal{B}_\theta)$. For $u \geq 0$, define $\tilde{\rho}_{A,H,\lambda,u} \in \mathcal{P}_\pi(\Theta)$ to be the probability measure with RN derivative with respect to π given by*

$$\frac{d\tilde{\rho}_{A,H,\lambda,u}}{d\pi}(\theta) = \frac{\exp\left[-\lambda\left(A(\theta) + uH(\theta)\right)\right]}{\int_\Theta \exp\left[-\lambda\left(A\left(\tilde{\theta}\right) + uH\left(\tilde{\theta}\right)\right)\right] d\pi\left(\tilde{\theta}\right)},\ u \geq 0.$$

Define the expected cost associated with $\tilde{\rho}_{A,H,\lambda,u}$ by

$$\Lambda(u) = \int_\Theta H(\theta) d\tilde{\rho}_{A,H,\lambda,u}(\theta),\ u \geq 0.$$

Lastly, for budgets $B \in \mathbb{R} \cup \{\infty\}$, we denote

$$\mathcal{E}_B = \left\{\rho \in \mathcal{P}_\pi(\Theta) : \int_\Theta H(\theta) d\rho(\theta) \leq B\right\},\ B \in \mathbb{R} \cup \{\infty\}.$$

(a) Let $B \in \mathbb{R} \cup \{\infty\}$. If

$$\pi\left(\{\theta : H(\theta) < B\}\right) > 0, \tag{16}$$

then,

$$\tilde{\rho}_{A,H,\lambda,\overline{u}_B} = \arg\min_{\mathcal{E}_B}\left[\int_\Theta A(\theta)\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi)\right],$$

where $\overline{u}_B = 0$ if $\Lambda(0) \leq B$ and otherwise, when $\Lambda(0) > B$, $\overline{u}_B > 0$ is the unique positive real number satisfying $\Lambda(\overline{u}_B) = B$.

(b) Assume condition (16) holds for a budget B. Then, for any $B' \geq B$,

$$\begin{aligned}
&\min_{\mathcal{E}_{B'}}\left[\int_\Theta A(\theta)\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi)\right] \\
&= \int_\Theta A(\theta) d\tilde{\rho}_{A,H,\lambda,\overline{u}_{B'}}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\tilde{\rho}_{A,H,\lambda,\overline{u}_{B'}}, \pi\right) \\
&= \sup_{u \geq 0}\left[\int_\Theta A(\theta) d\tilde{\rho}_{A,H,\lambda,u}(\theta) + u\left(\int_\Theta H(\theta) d\tilde{\rho}_{A,H,\lambda,u}(\theta) - B'\right) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\tilde{\rho}_{A,H,\lambda,u}, \pi\right)\right],
\end{aligned} \tag{17}$$

where $\overline{u}_{B'}$ is defined as $\overline{u}_B$ is in part (a), now with B' in place of B.

(c) For any $\tilde{u} \geq 0$, denote $B_{\tilde{u}} = \Lambda(\tilde{u})$. In place of (16), assume that

$$\mathbb{V}_{\theta \sim \pi}\left[H(\theta)\right] > 0, \tag{18}$$

where $\mathbb{V}_{\theta \sim \pi}[H(\theta)]$ denotes the variance of $H(\theta)$ when $\theta \sim \pi$. Then the property in part (b) holds for $B = B_{\tilde{u}}$. In particular the equalities in (17) holds for any $B' \geq B_{\tilde{u}}$.

The property in part (a) when $B = \infty$ (so $\overline{u}_B = 0$) is often utilized in the PAC-Bayesian literature with $A(\theta)$ taken as some loss or regret function; see Catoni (2007) and Alquier et al. (2016) among many possible examples. Lemma 3.1 (a) extends this setting to accommodate a secondary constraint objective associated with $H(\theta)$. $\Lambda(u)$, the cost associated with the choice of u, is decreasing in u. Intuitively, as the exponential re-weighting of π depends more heavily on $H(\theta)$ via a higher u, regions of the parameter or model space with a higher cost receive a less weighting and the overal cost is reduced. Convex optimization problems where the objective or constraint set involves the Kullback-Liebler divergence have been considered in earlier work, for example in Csiszár (1975). Rather than establishing (a) from the more abstract setting there, the proof in the Appendix utilizes well known properties of the KL divergence, stated as Lemma A.1 and Corollary A.1 in the Appendix. We note that Corollary A.1 (b) is a well known change-of-measure inequality (c.f. Csiszár (1975) and Donsker and Varadhan (1975)) that is widely utilized in deriving PAC-Bayesian generalization bounds.

Parts (b) and (c) will be useful for deriving the oracle-type inequalities in Section 4. When $B = B'$ in part (b), the result states that the duality gap between the primal and dual of the minimization problem in part (a) is zero. That is,

$$\min_{\rho\in\mathcal{P}_\pi(\Theta)} \sup_{u\geq 0} \left[\int_\Theta A(\theta)d\rho(\theta) + \frac{1}{\lambda}D_{\mathrm{KL}}(\rho,\pi) + u\left(\int_\Theta H(\theta)d\rho(\theta) - B\right)\right]$$
$$= \sup_{u\geq 0} \min_{\rho\in\mathcal{P}_\pi(\Theta)} \left[\int_\Theta A(\theta)d\rho(\theta) + \frac{1}{\lambda}D_{\mathrm{KL}}(\rho,\pi) + u\left(\int_\Theta H(\theta)d\rho(\theta) - B\right)\right].$$

Note that the left-hand side of the above equality, the primal problem, is equivalent to the optimization porblem in part (a). The right-hand side is the dual of this problem. That the right-hand side above is equivalent to the expression after the second equality in (17) can be seen from Part (a) or from Corollary A.1 (a) in the Appendix. The assumptions involving (16) and (18) in parts (b) and (c) constitute constraint qualifications.

We will apply Lemma 3.1 with $A(\theta) = R_n(\theta)$ or $R(\theta)$ and $H(\theta) = K_n(\theta)$ or $K(\theta)$. We consider two scenarios or perspectives. In the first, we have a (nonrandom) predetermined budget B and utilize a corresponding, sample dependent, choice of $\hat{u}$. In the second scenario, we start from a predetermined, non-random choice of u (or multiple values of u), which then corresponds to a sample dependent budget (or budgets) associated with $f_{G,\hat{\rho}_{\lambda,u}}$. We require the following assumptions. The first, corresponding to the condition in (16), will be required in the former scenario while the second condition, corresponding to condition (18), is required in the latter.

Assumption 3.4 *(i) Let $B \in \mathbb{R}\cup\{\infty\}$ be a desired budget. It is assumed that*

$$\pi\left(\theta\in\Theta : K(\theta) < B\right) > 0 \quad \text{and} \quad \pi\left(\theta\in\Theta : K_n(\theta) < B\right) > 0 \quad P^n \text{ almost surely.}$$

(ii) For $u \geq 0$, with $\hat{\rho}_{\lambda,u}$ defined with $\pi \in \mathcal{P}(\Theta)$ as in Definition 3.2, with $B(\hat{\rho}_{\lambda,u})$ as in (12), is assumed that, P^n almost surely,

$$\pi\left(\theta\in\Theta : K(\theta) < B\left(\hat{\rho}_{\lambda,u}\right)\right) > 0 \quad \text{and} \quad \mathbb{V}_{\theta\sim\pi}\left[K_n(\theta)\right] > 0,$$

where, for a fixed sample $S \in \mathcal{S}$, $\mathbb{V}_{\theta\sim\pi}[K_n(\theta)]$ denotes the variance of $K_n(\theta)$ when $\theta\sim\pi$.

Assumption 3.4 involves $\mathcal{F}_\Theta$, π and the sampling distribution P. Condition (i) requires that the budget of interest is not ruled out under the prior or reference measure π and is not exactly at the boundary of theoretical or empirical feasibility. With additional exposition, the condition that $\pi\left(\theta\in\Theta : K_n(\theta) < B\right) > 0$ holds P^n a.s. could be replaced by the condition that $\pi\left(\theta\in\Theta : K_n(\theta) < B\right) > 0$ holds with high probability. For example, with probability at least $1-\xi$, for some $\xi\in[0,1)$. In this case the theorems in Section 4 will remain valid except that the high probability bounds there, that hold with probability at least $1-\epsilon$ for $\epsilon\in(0,1]$, will now hold with probability at least $1-\epsilon-\xi$. Condition (ii) requires that there is always variation in empirical costs within models in $\mathcal{F}_\Theta$ and that the budget associated with the policy $f_{G,\hat{\rho}_{\lambda,u}}$, i.e. $B(\hat{\rho}_\lambda)$, is not at the boundary of feasible budgets among models selected by π. Adjustments to require the conditions in (ii) to hold with high probability rather than almost surely could be made similarly to the adjustment for (i).

Given Lemma 3.1 and the assumption above, the following definition will be relevant when the analysis starts from a fixed budget parameter that will correspond to a particular choice of u.

Definition 3.3 *Let $\hat{\rho}_{\lambda,u}$ and $\rho^*_{\lambda,u}$ be defined with $\pi \in \mathcal{P}(\Theta)$ as in Definition 3.2. For $B \in \mathbb{R}$, define $\hat{u}(B,\lambda)$ by*

$$\hat{u}(B,\lambda) = \arg\max_{u\geq 0} \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + u\left(\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) - B\right) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \pi\right),$$

$$u^*(B,\lambda) = \arg\max_{u\geq 0} \int_\Theta R(\theta) d\rho^*_{\lambda,u}(\theta) + u\left(\int_\Theta K(\theta) d\rho^*_{\lambda,u}(\theta) - B\right) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\rho^*_{\lambda,u}, \pi\right).$$

For $B = \infty$, define $\hat{u}(\infty,\lambda) = 0$ and $u^*(\infty,\lambda) = 0$.

To conclude the section, we point out immediate corollaries of Lemma 3.1 and Assumption 3.4 relevant to our setting. Define the sets

$$\mathcal{E}_B = \left\{\rho \in \mathcal{P}_\pi(\Theta) : \int_\Theta K(\theta) d\rho(\theta) \leq B\right\},\ B \in \mathbb{R} \cup \{\infty\} \tag{19}$$

and

$$\widehat{\mathcal{E}}_B = \left\{\rho \in \mathcal{P}_\pi(\Theta) : \int_\Theta K_n(\theta) d\rho(\theta) \leq B\right\},\ B \in \mathbb{R} \cup \{\infty\}. \tag{20}$$

In the scenario where we start from a pre-selected B, $\mathcal{E}_B$ is the (non-random) subset of $\mathcal{P}_\pi(\Theta)$ corresponding to Gibbs treatment rules with expected cost within the budget. $\widehat{\mathcal{E}}_B$ a random set that serves as an empirical counterpart, denoting the $\rho \in \mathcal{P}_\pi(\Theta)$ with Gibbs rules that meet the budget constraint empirically.

When analysis begins with a pre-determined value of u, $B(\hat{\rho}_{\lambda,u})$ as in Assumption 3.4 and its empirical counterpart,

$$\widehat{B}\left(\hat{\rho}_{\lambda,u}\right) = \int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta), \tag{21}$$

are both random. $B(\hat{\rho}_{\lambda,u})$ is the expected cost of $f_{G,\hat{\rho}_{\lambda,u}}$ in the target population given the sample-dependent $\hat{\rho}_{\lambda,u}$. This is not observed. However, it is a key object of interest, as it tells the researcher the expected cost of the estimated policy $f_{G,\hat{\rho}_{\lambda,u}}$ associated with u. For a predetermined u, both $\mathcal{E}_{B(\hat{\rho}_{\lambda,u})}$ and $\widehat{\mathcal{E}}_{\widehat{B}(\hat{\rho}_{\lambda,u})}$ are random sets. The former corresponds to all Gibbs treatment policies with an expected budget in the target population that is less than or equal to that of $f_{G,\hat{\rho}_{\lambda,u}}$. The latter serves as an empirical counterpart for which membership can be evaluated from the sample.

For the scenario where B is predetermined we have the following properties that follow from 3.1 and Assumption 3.4 (i).

Corollary 3.1 *Let Assumptions 3.2 and 3.4 (i) hold for $B \in \mathbb{R} \cup \{\infty\}$.*

(i) The following properties hold P^n almost surely. For any $\lambda > 0$ and $B' \geq B$, $\hat{u}(B', \lambda)$ exists, is unique, and satisfies that $\hat{u}(B', \lambda) = 0$ when $\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,0}(\theta) \leq B'$ and $\hat{u}(B', \lambda)$ is positive and satisfies $\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) = B'$ when $\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,0}(\theta) > B'$. Additionally,

$$\hat{\rho}_{\lambda,\hat{u}(B',\lambda)} = \underset{\widehat{\mathcal{E}}_{B'}}{\arg\min} \left[\int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) \right],$$

and

$$\begin{aligned} &\min_{\widehat{\mathcal{E}}_{B'}} \left[\int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) \right] \\ &= \sup_{u \geq 0} \left[\int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + u \left(\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) - B' \right) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \pi\right) \right]. \end{aligned} \tag{22}$$

(ii) For any $\lambda > 0$ and $B' \geq B$, $u^*(B', \lambda)$ exist, is unique, and satisfies that $u^*(B', \lambda) = 0$ when $\int_\Theta K(\theta) d\rho^*_{\lambda,0}(\theta) \leq B'$ whereas, when $\int_\Theta K(\theta) d\rho^*_{\lambda,0}(\theta) > B'$, $u^*(B', \lambda)$ is positive and satisfies $\int_\Theta K(\theta) d\rho^*_{\lambda,u^*(B',\lambda)}(\theta) = B'$. Additionally,

$$\rho^*_{\lambda,u^*(B',\lambda)} = \underset{\mathcal{E}_{B'}}{\arg\min} \left[\int_\Theta R(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) \right],$$

and

$$\begin{aligned} &\min_{\mathcal{E}_{B'}} \left[\int_\Theta R(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) \right] \\ &= \sup_{u \geq 0} \left[\int_\Theta R(\theta) d\rho^*_{\lambda,u}(\theta) + u \left(\int_\Theta K(\theta) d\rho^*_{\lambda,u}(\theta) - B' \right) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\rho^*_{\lambda,u}, \pi\right) \right]. \end{aligned} \tag{23}$$

Lastly, when u is predetermined, we have the following properties that follow from Lemma 3.1 and Assumption 3.4 (ii).

Corollary 3.2 *Let Assumptions 3.2 and 3.4 (ii) hold for $u \geq 0$. For any $\lambda > 0$, we have the following properties.*

(i) With probability one,

$$\hat{\rho}_{\lambda,u} = \underset{\widehat{\mathcal{E}}_{\widehat{B}(\hat{\rho}_{\lambda,u})}}{\arg\min} \left[\int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) \right],$$

and equation (22) holds for any $B' \geq \widehat{B}(\hat{\rho}_{\lambda,u})$.

(ii) For any $B' \geq B(\hat{\rho}_{\lambda,u})$, with probability one,

$$\begin{aligned} &\min_{\rho \in \mathcal{E}_{B'}} \left[\int_\Theta R(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\rho, \pi\right) \right] \\ &= \sup_{a \geq 0} \left[\int_\Theta R(\theta) d\rho^*_{\lambda,a}(\theta) + a \left(\int_\Theta K(\theta) d\rho^*_{\lambda,a}(\theta) - B' \right) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\rho^*_{\lambda,a}, \pi\right) \right]. \end{aligned}$$

4 PAC-Bayesian Analysis

Here we provide theoretical motivation for decision rules utilizing $\hat{\rho}_{\lambda,u}$ or $\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}$. In Section 4.1, we first construct PAC-Bayesian generalization bounds that are similar to counterparts in earlier literature. Then we derive oracle-type inequalities that compare the proposed treatment rules to alternatives in terms of regret in the target population for a given budget. The results in Section 4.1 allow for a general choice of the prior or reference measure π utilized in the definition of $\hat{\rho}_{\lambda,u}$ and $\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}$. As a result, several bounds there contain KL divergence terms related to the complexity of the learning problem and the model class $\mathcal{F}_\Theta$. In Section 4.2, we specify $\mathcal{F}_\Theta$ to consist of rules of the form in (10) and take π to be an uninformative multivariate normal distribution. In this setting, we obtain oracle-type inequalities that compare the regret of our proposed treatment assignment rules directly to that of the rules in $\mathcal{F}_\Theta$ with the lowest welfare regret that are in budget. In section 4.3, we show that desirable properties for the majority vote rules associated with $\hat{\rho}_{\lambda,u}$ can be inherited by their majority vote counterparts.

Our analysis builds from results and techniques in the PAC-Bayesian literature that are not always stated in ways that are directly applicable to our setting. Results from earlier literature are adapted to our setting in Appendix Section A.1, which also contains additional properties of interest. For the most part, proofs are included there for completeness even when the adjustments are fairly minor. This spares the reader from visiting multiple references requiring concerted adjustments at certain steps of our analysis. Proofs specific to Section 4 are contained in Appendix Section A.3.

4.1 Regret Bounds and Oracle-Type Inequalities

The first step in our analysis, Theorem 4.1, obtains alterations of earlier PAC-Bayesian generalization bounds for the treatment assignment setting. A variant of part (a) appears in Catoni (2007) which considers classification in the 0/1-loss setting. In our setting, it can be derived as a special case of a bound appearing in Alquier et al. (2016) or via a general approach to PAC-Bayesian bounds outlined, for example, in Germain et al. (2015). We utilize the latter approach which is useful during additional steps of our analysis. The proofs of parts (b) and (c) utilize the approach of Lever et al. (2010), with part (b) being an alteration of Theorem 3 in that work.

Theorem 4.1 *Let $\pi \in \mathcal{P}(\Theta)$ and let Assumptions 3.1, 3.2, and 3.3 hold. Set*

$$\{V_n(\theta), V(\theta), M_\ell\} = \{R_n(\theta), R(\theta), M_y\} \text{ or else } \{V_n(\theta), V(\theta), M_\ell\} = \{K_n(\theta), K(\theta), M_c\}.$$

We have the following properties.

(a) Let $\epsilon \in (0, 1]$, $\lambda > 0$ and $s \in \{-1, 1\}$. With probability at least $1 - \epsilon$, for all $\rho \in \mathcal{P}_\pi(\Theta)$ simultaneously it holds that

$$\int_\Theta s\,[V_n(\theta) - V(\theta)]\, d\rho(\theta) \leq \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_\ell^2}{8n\kappa^2} + \log\frac{1}{\epsilon}\right].$$

(b) Let $\lambda > 0$, $u \geq 0$, and $\epsilon \in (0, 1]$. With probability at least $1 - \epsilon$, it holds that

$$\left(\int_\Theta V(\theta) d\hat{\rho}_{\lambda,u}(\theta) - \int_\Theta V_n(\theta) d\hat{\rho}_{\lambda,u}(\theta)\right)^2$$
$$\leq \frac{M_\ell^2}{2n\kappa^2}\left[\frac{\lambda\sqrt{2}\,(M_y + uM_c)}{\kappa\sqrt{n}}\sqrt{\log\left(2\sqrt{n}\right) + \log\frac{2}{\epsilon}} + \frac{\lambda^2\,(M_y + uM_c)^2}{2n\kappa^2} + \log\left(2\sqrt{n}\right) + \log\frac{2}{\epsilon}\right].$$

(c) *Let $\lambda > 0$, $u \geq 0$, and $\epsilon \in (0, 1]$. With probability at least $1 - \epsilon$, it holds that*

$$\begin{aligned}
&\int_\Theta V(\theta) d\hat{\rho}_{\lambda,u}(\theta) - \int_\Theta V_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) \\
&\leq \frac{\sqrt{2}\,(M_y + uM_c)}{\kappa\sqrt{n}} \sqrt{\log\left(2\sqrt{n}\right) + \log\frac{2}{\epsilon}} + \frac{\lambda\,(M_y + uM_c)^2}{2n\kappa^2} + \frac{1}{\lambda}\left[\frac{\lambda^2 M_\ell^2}{8n\kappa^2} + \log\frac{2}{\epsilon}\right].
\end{aligned}$$

Theorem 4.1 contains high probability bounds for notions of the generalization error between the target population regret (or alternatively, expected cost) and its empirical counterpart for Gibbs treatment rules. For example, one notion of generalization error for the cost of policy $f_{G,\hat{\rho}_{\lambda,u}}$ could be the absolute difference,

$$\left|K\left(f_{G,\hat{\rho}_{\lambda,u}}\right) - K_n\left(f_{G,\hat{\rho}_{\lambda,u}}\right)\right|.$$

Suppose we take $\lambda = a\kappa\sqrt{n}/(M_y + uM_c)$ for some constant $a > 0$. Then Part (b) says that with probability at least $1 - \epsilon$, this absolute difference is less than or equal to

$$\frac{M_c}{\kappa\sqrt{2n}}\left[a\sqrt{\log\left(4n\right) + 2\log\frac{2}{\epsilon}} + \frac{a^2}{2} + \log(2\sqrt{n}) + \log\frac{2}{\epsilon}\right]^{1/2} = \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right).$$

When M_c and M_y are known, this upper bound can be evaluated for a given choice of a. We say that $K_n(f_{G,\hat{\rho}_{\lambda,u}})$ is Probably (with probability at least $1-\epsilon$) and Approximately (the $\mathcal{O}(\sqrt{\log(n)/n})$ upper bound on the absolute difference) Correct for $K(f_{G,\hat{\rho}_{\lambda,u}})$. This suggest that for a predetermined choice of u, $K_n(f_{G,\hat{\rho}_{\lambda,u}})$ will give a reasonable estimate of the expected cost in the target population, $K(f_{G,\hat{\rho}_{\lambda,u}})$, provided that λ is not too large. Part (c) is a variation of the style of bound in (b) that is useful in deriving subsequent results. We note that the above choice for λ may not best in practice, or even feasible if the upper bound M_c is not known. In practice λ is chosen via cross validation, which can be accommodated by Theorem 4.1 similarly to the choice of u as discussed below.

The bounds in Theorem 4.1 can be adjusted to accommodate the setting where λ, u, or pairs (λ, u) are selected from a finite set of values $\mathcal{W}$. With $|\mathcal{W}|$ denoting the number of elements in $\mathcal{W}$, one can apply a union bound argument similar to that in the proof of part (b). The theorem is applied once for each element of $\mathcal{W}$ with size $\epsilon/|\mathcal{W}|$ for each repetition. Then, applying the union bound argument, the bounds as stated in Theorem 4.1 remain valid for any element of $\mathcal{W}$ with the alteration that the term $\log\frac{1}{\epsilon}$ in part (a) is replaced by $(\log\frac{1}{\epsilon} + \log|\mathcal{W}|)$ and the terms $\log\frac{2}{\epsilon}$ in parts (b) and (c) are replaced by $(\log\frac{2}{\epsilon} + \log|\mathcal{W}|)$. For example, when $\lambda = \mathcal{O}(\sqrt{n})$, this adds a term that is $\mathcal{O}(\log|\mathcal{W}|/\sqrt{n})$ to the right hand side of the high probability bound in part (a). This observation is applicable to the remaining theorems in the paper, with minor adjustments. Therefore, it is not unreasonable to start with multiple values for u. Then one may choose u in $\hat{\rho}_{\lambda,u}$ for the final policy based on the empirical estimates of the associated budgets, $K_n(f_{G,\hat{\rho}_{\lambda,u}})$ for $u \in \mathcal{W}$, or via cross validation.

Before comparing our suggested treatment policies to alternative choices, we discuss a final insight from Theorem 4.1. Part (a) yields that, with probability at least $1 - \epsilon$,

$$R(f_{G,\rho}) \leq \int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{1}{\epsilon}\right], \tag{24}$$

for all $\rho \in \mathcal{P}_\pi$ simultaneously. Given a budget B such that Assumption 3.4 (i) holds, Corollary 3.1 (i) states that, almost surely, $\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}$ produces the smallest upper bound for the target population regret in (24) among all $\rho \in \mathcal{P}_\pi(\Theta)$ such that $K_n(f_{G,\rho}) \leq B$. Similarly, starting from a given value of u, Corollary 3.2 with Assumption 3.4 (ii) shows that $\hat{\rho}_{\lambda,u}$ results in the smallest upper bound for the target population regret among Gibbs rules with an empirical budget less than or equal to $\widehat{B}(\hat{\rho}_{\lambda,u})$ defined in (21).

Although Theorem 4.1 (a) is most useful for our analysis, in the PAC-Bayesian literature there are alternative generalization bounds to (24) that apply for all $\rho \in \mathcal{P}_\pi(\Theta)$ and could be adapted to our setting. Most notably, variants of the bounds in Seeger (2002) and Catoni (2007) are fairly ubiquitous in the literature. Either directly or via a slight relaxation, these bounds also suggest choosing ρ to minimize

$$\int_\Theta R_n(\theta)d\rho(\theta) + \frac{1}{\lambda}D_{\mathrm{KL}}(\rho, \pi), \tag{25}$$

for some $\lambda > 0$. Hence, if we impose an empirical budget constraint these would again lead back to $\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}$ and $\hat{\rho}_{\lambda,u}$. We note that Seeger's bound is utilized in our analysis to derive parts (b) and (c) of Theorem 4.1 and appears as Theorem A.2 in Appendix Section A.1. While this bound does not yield a closed form solution $\tilde{\rho}$ that minimizes an upper bound on the regret, we refer to the discussion in Thiemann et al. (2017) regarding a relaxation that suggests minimizing (25) with λ replaced by λn, an equivalent problem when λ is cross validated. The style of bound in Catoni (2007), in particular Theorem 1.2.6 there, can be adapted to our setting via the approach in Germain et al. (2015) and again suggests choosing ρ to minimize (25).

Next we derive oracle-type inequalities that compare the target population regret associated with $\hat{\rho}_{\lambda,u}$ or $\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}$ to that of alternative choices of ρ among Gibbs treatment rules within a relevant budget. It may be helpful to recall the definitions of $\mathcal{E}_B$ and $\mathcal{E}_{B(\hat{\rho}_{\lambda,u})}$ from (19) and (12),

$$\mathcal{E}_B = \{\rho \in \mathcal{P}_\pi(\Theta) : K\left(f_{G,\rho}\right) \leq B\} \text{ and } \mathcal{E}_{B(\hat{\rho}_{\lambda,u})} = \left\{\rho \in \mathcal{P}_\pi(\Theta) : K\left(f_{G,\rho}\right) \leq K\left(f_{G,\hat{\rho}_{\lambda,u}}\right)\right\}.$$

We have the following result.

Theorem 4.2 *Let $\pi \in \mathcal{P}(\Theta)$, $\lambda > 0$, and $\epsilon \in (0,1]$. Under Assumptions 3.1, 3.2, and 3.3, we have the following properties.*

(a) Let $B \in \mathbb{R} \cup \{\infty\}$, denote $\hat{u} = \hat{u}(B,\lambda)$ and let Assumption 3.4 (i) hold. With probability at least $1-\epsilon$, it holds that

$$R\left(f_{G,\hat{\rho}_{\lambda,\hat{u}}}\right) \leq \min_{\rho\in\mathcal{E}_B}\left\{R\left(f_{G,\rho}\right) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho,\pi) + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}}\right\}.$$

(b) Fix $u \geq 0$ and let Assumption 3.4 (ii) hold. With probability at least $1-\epsilon$, it holds that

$$R\left(f_{G,\hat{\rho}_{\lambda,u}}\right) \leq \min_{\rho\in\mathcal{E}_{B(\hat{\rho}_{\lambda,u})}}\left\{R\left(f_{G,\rho}\right) + \frac{1}{\lambda}D_{\mathrm{KL}}(\rho,\pi) + uU_1\left(\epsilon;\lambda,u,n\right) + U_2\left(\epsilon;\lambda,u,n\right)\right\}.$$

where

$$U_1\left(\epsilon;\lambda,u,n\right) = \frac{\sqrt{2}\left(M_y+uM_c\right)}{\kappa\sqrt{n}}\sqrt{\log\left(2\sqrt{n}\right)+\log\frac{4}{\epsilon}} + \frac{\lambda\left(M_y+uM_c\right)^2}{2n\kappa^2},$$

and

$$U_2\left(\epsilon;\lambda,u,n\right) = \sqrt{\frac{(M_y+uM_c)^2\log(4/\epsilon)}{2n\kappa^2}} + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2+uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right].$$

Note that if $\lambda = \mathcal{O}(n^{1/2})$, then for any $u \geq 0$ and $\epsilon \in (0, 1]$,

$$U_1(\epsilon; \lambda, u, n) = \mathcal{O}\left(\sqrt{\frac{\log(n)}{n}}\right) \text{ and } U_2(\epsilon; \lambda, u, n) = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

Theorem 4.2 contains sharp oracle-type inequalities that hold with high probability. They differ slightly from traditional oracle inequalities in that the right-hand sides contain objects that are random.

Consider part (b) first. In this case, the randomness on the right-hand side of the inequality stems from $\mathcal{E}_{B(\hat{\rho}_{\lambda,u})}$ which depends on the sample through $B(\hat{\rho}_{\lambda,u}) = K(f_{G,\hat{\rho}_{\lambda,u}})$, the un-observable expected target population cost of $\hat{\rho}_{\lambda,u}$. For a predetermined u, it is natural to ask if there are alternatives in $\mathcal{P}_\pi(\Theta)$ that would yield lower regret for the same or lower expected cost. $\mathcal{E}_{B(\hat{\rho}_{\lambda,u})}$ is therefore the natural set of interest for comparison with $\hat{\rho}_{\lambda,u}$ as it is the subset of $\mathcal{P}_\pi(\Theta)$ with Gibbs rules that have target population costs no greater than $B(\hat{\rho}_{\lambda,u})$. Given a budget $B(\hat{\rho}_{\lambda,u})$, an oracle with knowledge of $R(\theta)$ could solve for $\arg\min_{\rho \in \mathcal{E}_{B(\hat{\rho}_{\lambda,u})}} R(f_{G,\rho})$. For $\lambda \to \infty$, we may consider $\arg\min_{\rho \in \mathcal{E}_{B(\hat{\rho}_{\lambda,u})}} R(f_{G,\rho}) + \lambda^{-1} D_{\mathrm{KL}}(\rho, \pi)$ as a second-best oracle solution. When $\lambda = \mathcal{O}(n^{1/2})$, for example, part (b) indicates that with high probability $\hat{\rho}_{\lambda,u}$ is close to the second best oracle solution. In Section 4.2 we consider oracle-type inequalities without the KL penalty term appearing.

In part (a), the interpretation is similar to that in part (b), except that now the set of alternative Gibbs estimators for comparison are those that satisfy the predetermined budget B. This set is non-random, however now the right-hand side contains a term involving the random $\hat{u} = \hat{u}(B, \lambda)$ as defined in Definition 3.3 and Corollary 3.1. Note that $\hat{u}$ is the value taken by the Lagrange multiplier u in the problem

$$\min_{\rho \in \mathcal{E}_B} \sup_{u \geq 0} \left\{ \int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) + u \left(\int_\Theta K_n(\theta) d\rho(\theta) - B \right) \right\}.$$

$\hat{u}$ measures the marginal decrease in empirical penalized regret (alternatively, the increase in empirical penalized welfare) resulting from a marginal relaxation of the budget. Recall the welfare and budget are measured per treatment. For example, when benefits and costs are measured in dollars, how many dollars of penalized welfare are obtained (empirically) by increasing the maximum empirical cost by a dollar. In more extreme scenarios where a small increase in the budget produces a large increase in empirical welfare, the bound becomes less meaningful as the right-hand side approaches the maximum possible regret (if this level is exceeded, the bound becomes trivial). An example of an extreme setting would be when $\hat{u} = \mathcal{O}_p(n^\alpha)$ for some $\alpha \geq 1/2$. When $\hat{u}n^{-1/2}$ is large relative to typical or maximal values of the regret (which ranges from zero to twice the maximal welfare), this situation is visible to the analyst. For a fixed λ, a statement similar to part (a) can be obtained where $\hat{u}$ is replaced by a non-random constant if we make additional assumptions on the data generating distribution P. For example, if we instead assume the marginal increase in population penalized regret associated with a small relaxation of the empirical budget is $\mathcal{O}_p(1)$. As it stands, the bound produces a robustness check for the method's motivation. Intuitively, if it is easy to dramatically change the empirical welfare by relatively small budget changes, so that $\hat{u}n^{-1/2}$ is large, we may be in a situation where it is difficult to learn policies well for the given B and the proposed rules should be treated cautiously.

If regions of the model space with desirable regret and budget are assigned lower probability by π, the distributions $\rho \in \mathcal{P}_\pi(\Theta)$ with the best trade-off between $R(f_{G,\rho})$ and $D_{\mathrm{KL}}(\rho, \pi)$ in

Theorem 4.2 will tend to have larger $D_{\mathrm{KL}}(\rho, \pi)$ terms. As a result, the upper bounds will be larger and less informative. Similarly, applying Theorem 4.1 part (a) with $\rho = \hat{\rho}_{\lambda,c}$ for either $c = u \geq 0$ or $c = \hat{u}(B, \lambda)$, and noting Corollaries 3.1 and 3.2, the regret and budget bounds there are influenced by the trade-off between empirical regret (or cost) and $D_{\mathrm{KL}}(\hat{\rho}_{\lambda,c}, \pi)$. $D_{\mathrm{KL}}(\hat{\rho}_{\lambda,c}, \pi)$ increases when $\hat{\rho}_{\lambda,c}$ involves a greater re-weighting of π in definition 3.2. The impact of the KL terms in the bounds of this subsection are therefore related to the learning problem and model space complexity. It is influenced by how large the model space is, how narrow the subset of the model space with low regret/budget is, the relative difference in between lower and higher regret regions and the noisiness of the data. In parts (b) and (c) of Theorem 4.1, where the KL term is absent, this role falls more to the λ parameter: if the problem is more complex, larger (relative to n) values of λ are needed to achieve lower regret or cost. If λ is too large, remainder terms in the generalization error bounds increase. See Lever et al. (2010) for further discussion of complexity in the setting of bounds of the form in (b) and (c).

Conversely, when the policy maker has (sample independent) knowledge of the data generating process, they may be able to select or alter a given choice of π to focus on the regions of the model space that best balance regret and cost. Then $D_{\mathrm{KL}}(\rho, \pi)$ can be smaller for ρ that put the greatest weight on the most desirable regions of the parameter space. The result is smaller upper bounds in Theorem 4.2 and Theorem 4.1 (a). A benefit of the Gibbs rules associated with $\hat{\rho}_{\lambda,u}$ and $\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}$ is that economic theory or situation-specific knowledge can be factored into the treatment rule via π. Compatibility with expert knowledge may be a valuable advantage in settings where resource limitations imply that some individuals with a positive CATE will not be treated. As we will see in Section 4.2, such knowledge is not required for the procedures to have desirable properties.

4.2 Normal Prior

As noted at the end of Section 4.1, perhaps unsurprisingly, knowledge about the data generating process can confer estimation benefits through the choice of π. While it is a positive attribute that the proposed treatment rules can utilize this information when available, it is important to emphasize that such knowledge is not a requirement. Learning procedures based on PAC-Bayesian analysis often utilize uninformative or less informative choices for π, such as normal distributions, uniform distributions when Θ is compatible, or sparsity inducing distributions.

Here we take π to be a multivariate normal distribution centered at the origin and utilize the models of the form in (10). We show that the proposed treatment rules maintain desirable properties. In doing so, the KL divergence term is removed from the oralce inequalities, resulting in a clearer comparison to alternative treatment rules. We leave an exploration of alternative prior choices and the settings where they may be desirable to future research.

We satisfy Assumptions 3.2 and 3.3 with the following, more specific, condition. Note that in the assumption below we are treating q as fixed; it does not grow with the sample size.

Assumption 4.1 *It is assumed that $\mathcal{F}_\Theta$ consists of treatment rules f_θ as described by* (10)*, with $\Theta = \mathbb{R}^q$ equipped with the standard Borel σ-algebra. Let*

$$\Phi_{\mu,\sigma^2} \in \mathcal{P}(\mathbb{R}^q)$$

denote a multivariate normal distribution with mean vector μ and covariance matrix $\sigma^2 I_q$ for some $\sigma > 0$. We assume that $\pi = \Phi_{0,\sigma_\pi^2}$ for some $\sigma_\pi > 0$ that does not depend on the sample.

Next, we define

$$\Theta_B = \{\theta \in \mathbb{R}^q : K(\theta) \leq B\} \text{ and } \Theta_{B(\hat{\rho}_{\lambda,u})} = \{\theta \in \mathbb{R}^q : K(\theta) \leq B(\hat{\rho}_{\lambda,u})\},$$

and denote

$$\overline{\theta} \in \arg\min_{\Theta_B} [R(\theta)] \text{ and } \overline{\theta}_u \in \arg\min_{\Theta_{B(\hat{\rho}_{\lambda,u})}} [R(\theta)]. \tag{26}$$

Note that $\Theta_{B(\hat{\rho}_{\lambda,u})}$ and $\overline{\theta}_u$ are random as they vary with $B(\hat{\rho}_{\lambda,u})$. $\Theta_{B(\hat{\rho}_{\lambda,u})}$ is the set of parameters such that the corresponding models in $\mathcal{F}_\Theta$ have lower expected target population cost than $f_{G,\hat{\rho}_{\lambda,u}}$. $\overline{\theta}_u$ is the minimizer of the population regret among this set. With regard to $\overline{\theta}$ and $\overline{\theta}_u$, we assume the following condition.

Assumption 4.2 *With probability one, $\overline{\theta}$ and $\overline{\theta}_u$ as defined in (26) exist and are nonzero.*

This type of condition is implicitly assumed in, for example, Kitagawa and Tetenov (2018) and in Sun (2021). It simplifies the exposition rather than allowing that the models associated with these parameters have regret that is arbitrarily close to an infimum. The requirement that $\overline{\theta}$ and $\overline{\theta}_u$ are nonzero simply specifies that the covariates are relevant to the welfare. Lastly, our analysis will also require the following technical condition.

Assumption 4.3 *There exists a constant $\nu > 0$ such that*

$$P\left[(\phi(X)^\intercal\theta)\left(\phi(X)^\intercal\theta'\right) < 0\right] \leq \nu\|\theta - \theta'\|$$

for any θ and $\theta' \in \mathbb{R}^q$ such that $\|\theta\| = \|\theta'\| = 1$.

Assumption 4.3 or a direct analog is applied in several classification and bipartite ranking applications utilizing PAC-Bayesian approaches. For examples see Ridgway et al. (2014), Alquier et al. (2016), and Guedj and Robbiano (2018). It is a fairly mild requirement and, as is shown in Alquier et al. (2016) (c.f. p. 10 there), it is satisfied whenever $\phi(X)/\|\phi(X)\|$ has a bounded density on the unit sphere.

We have the following result.

Theorem 4.3 *Let Assumptions 3.1, 4.1, 4.2, and 4.3 hold. Let $\sigma_\pi = 1/\sqrt{q}$. Then we have the following properties for any $\epsilon \in (0, 1]$.*

(a) Let Assumption 3.4 (i) hold for a given $B \in \mathbb{R} \cup \{\infty\}$. Let $\lambda = \kappa\sqrt{nq}/M_y$, $\hat{u} = \hat{u}(B, \lambda)$ and $u^* = u^*(B, \lambda/2)$. With probability at least $1 - \epsilon$, it holds that

$$R\left(f_{G,\hat{\rho}_{\lambda,\hat{u}}}\right) \leq R\left(\overline{\theta}\right) + \sqrt{\frac{q}{n}}\log(4n)\frac{M_y}{\kappa} + \frac{2M_y\log\frac{3}{\epsilon}}{\kappa\sqrt{nq}} + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} + \frac{u^*\nu M_c}{\sqrt{n}} + \overline{U}_1(n; q),$$

where $\overline{U}_1(n; q) = \mathcal{O}(n^{-1/2})$ with the explicit formulation given in the proof.

(b) Fix $u \geq 0$ and set $\lambda = \kappa\sqrt{nq}/(M_y + uM_c)$. Let Assumption Assumption 3.4 (ii) hold. With probability at least $1 - \epsilon$,

$$R\left(f_{G,\hat{\rho}_{\lambda,u}}\right) \leq R\left(\overline{\theta}_u\right) + \frac{M_y + uM_c}{\kappa}\left[\overline{U}_2(n; q, u, \epsilon) + \overline{U}_3(n; q, u, \epsilon) + \overline{U}_4(n; q, u)\right],$$

where $\overline{U}_2(n; q, u, \epsilon) = \mathcal{O}(\log(n)n^{-1/2})$, $\overline{U}_3(n; q, u, \epsilon) = \mathcal{O}(n^{-1/2})$, and $\overline{U}_4(n; q, u) = \mathcal{O}(n^{-1/2})$, with the explicit forms given in the proof.

Note that the values for λ in parts (a) and (b) are chosen to produce the nearly optimal rate of convergence in part (b). In practice there may be better choices and we will typically choose λ via cross validation. As noted in the discussion following Theorem 4.1, we may choose λ, u, or pairs (λ, u) from a finite set of values $\mathcal{W}$. In this case the theorem above can be adjusted to hold simultaneously for all elements of $\mathcal{W}$ by replacing the terms $\log(\epsilon/3)$ on the right-hand side of the inequality in (a) by $\log(\epsilon/3) + \log|\mathcal{W}|$ and the terms on the right-hand side of (b) that involve $\log(\epsilon/4)$, which appear in the $\overline{U}_j$ terms defined in the proof, are replaced by $\log(\epsilon/4) + \log|\mathcal{W}|$. For example, for fixed $\epsilon \in (0, 1]$ and $u \geq 0$, this adds a term that is $\mathcal{O}(\log(|\mathcal{W}|)n^{-1/2})$ to the right hand side of (b).

In Theorem 4.3, $f_{G,\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}}$ and $f_{G,\hat{\rho}_{\lambda,u}}$ are compared to the best (non-stochastic) models in $\mathcal{F}_\Theta$ with an expected cost no greater than B or $B(\hat{\rho}_{\lambda,u})$, respectively. Additionally, the absence of KL terms in the inequalities allows for a more salient comparison to relevant alternatives. In part (b), for any $u \geq 0$ and $\epsilon \in (0, 1]$, the terms beside $R(\overline{\theta}_u)$ on the right hand-side are collectively $\mathcal{O}(\log(n)n^{-1/2})$. With high probability, the regret of $f_{G,\hat{\rho}_{\lambda,u}}$ gets close to the regret an oracle would obtain choosing the best rule from the subset of $\mathcal{F}_\Theta$ with a target population budget no greater than that of $f_{G,\hat{\rho}_{\lambda,u}}$. The rate $\log(n)n^{-1/2}$ is nearly optimal. For example, in the unconstrained case with $B = \infty$, which corresponds to $u = 0$ or $\hat{u} = u^* = 0$, Kitagawa and Tetenov (2018) show that $n^{-1/2}$ is the optimal rate for bounds on the expected regret of the empirical welfare maximizer over $\mathcal{F}_\Theta$, provided $\mathcal{F}_\Theta$ has a finite VC-dimension (see the discussion there for more details).

Part (a) has the complication of involving $\hat{u} = \hat{u}(B, \lambda)$ and $u^* = u^*(B, \lambda/2)$ as λ grows with n. The effect of $\hat{u}$ is related to the marginal decrease in $R_n(f_{G,\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}}) + \lambda^{-1}D_{\mathrm{KL}}(\hat{\rho}_{\lambda,\hat{u}(B,\lambda)}, \pi)$ associated with marginal increases in B as the penalty diminishes (λ increases). The behavior of u^* is related to the marginal decrease in the penalized regret of $f_{G,\rho^*_{\lambda,u^*(B,\lambda/2)}}$ associated with marginal increases in B. Suppose we are unlikely to have large marginal gains in empirical or theoretical penalized regret associated with a marginal increase in B at all or small penalty levels (i.e. as $\lambda \to \infty$). Then (a) implies that, with high probability and for large enough sample sizes, the regret of $f_{G,\hat{\rho}_{\lambda,\hat{u}}}$ is close to the regret that would be obtained by an oracle choosing the best policy from the subset of $\mathcal{F}_\Theta$ with an expected cost in the target population that is less than or equal to B. For example, if $u^* = \mathcal{O}(1)$ and $\hat{u} = \mathcal{O}_p(1)$ as n and λ increase, then the terms on the right-hand side of the inequality in (a) other than $R(\overline{\theta})$ are $\mathcal{O}_p(\log(n)n^{-1/2})$.

We conclude this subsection with remarks regarding implications for the proposed treatment assignment rules. One drawback of starting from a fixed B and utilizing $\hat{u} = \hat{u}(B, \lambda)$ is the absence of a counterpart to Theorem 4.1 (b) for the cost $K(f_{G,\hat{\rho}_{\lambda,\hat{u}}})$ when $\hat{u}$ is random. Even when $\hat{u}$ and u^* are well behaved so that Theorem 4.3 (a) implies it is likely that $f_{G,\hat{\rho}_{\lambda,\hat{u}}}$ will have regret comparable to the best rules in $\mathcal{F}_\Theta$ with expected cost less than B, this may be achieved with an expected cost greater than B. On the other hand Theorem 4.1 (a) with $\rho = \hat{\rho}_{\lambda,\hat{u}}$ yields that with probability at least $1 - \epsilon$,

$$K\left(f_{G,\hat{\rho}_{\lambda,\hat{u}}}\right) \leq B + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,\hat{u}}, \pi\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_c^2}{8n\kappa^2} + \log\frac{1}{\epsilon}\right],$$

where we have used the fact that $K_n(f_{G,\hat{\rho}_{\lambda,\hat{u}}}) \leq B$ a.s. under the assumptions of the theorem. When $\lambda = \mathcal{O}(n^{1/2})$, for example, whether or not we have an upper bound that approaches B depends on the behavior of this KL term. Unfortunately, $\hat{u}$ and the KL term above are difficult to analyze in this scenario as $\hat{u}$ is essentially defined implicitly to ensure $K_n(f_{G,\hat{\rho}_{\lambda,\hat{u}}}) \leq B$. It is possible to cross validate B, for example examining values less than B to try and ensure the expected budget is not violated. The comments regarding extending the high probability bounds to apply simultaneously for multiple values of u can be applied to choices for B as well.

On the whole, the procedure starting with a set of values for u may be more compelling. By Theorem 4.1 (b) and the surrounding discussion, for values u in a reasonably sized set $\mathcal{W}$, the values of $K_n(f_{G,\hat{\rho}_{\lambda,u}})$ provide reasonable estimates of $K(f_{G,\hat{\rho}_{\lambda,u}})$, the expected costs of these policies conditional on the rules estimated from the sample. These can be utilized to select u. Alternatively, u can be chosen from $\mathcal{W}$ via cross validation or by some other method. For example, in the case of pure quantity constraints, it may be possible use data from the target population to select u to achieve the correct (or nearly correct) proportion of treatments assigned in the target population. Theorem 4.3 (b) and its extension to hold for all $u \in \mathcal{W}$ simultaneously, then indicate it is likely $R(f_{G,\hat{\rho}_{\lambda,u}})$ for the selected u will be comparable to the best treatment rules in $\mathcal{F}_\Theta$ among those whose target population cost does not exceed that of $f_{G,\hat{\rho}_{\lambda,u}}$. Hence by starting from a set of u values, the policy maker can trace out reasonable estimates of the target population budget horizon. At the same time, the policy selected according to these budget estimates is likely to be the best bang for the buck in that the associated regret gets close to that which an oracle would choose for the same target population cost.

4.3 The Majority Vote Treatment Rule

Let $\rho \in \mathcal{P}_\pi(\Theta)$. As mentioned in Section 3.2, the non-stochastic majority vote treatment rule $f_{\mathrm{mv},\rho}$ in (13) is a close relative of the Gibbs rule $f_{G,\rho}$ that can prove numerically more stable in practice. In the classification literature, it is well known that the risk associated with the majority vote rule, where risk is defined for a zero-one loss function, is upper bounded by twice the risk associated with the Gibbs classification method (e.g., Langford and Shawe-Taylor (2003), McAllester (2003a)). Hence analysis of the Gibbs treatment rule is often used to justify use of the majority vote. Additionally, the "2×" upper bound can be loose and it is not uncommon for majority vote rules to outperform Gibbs rules. We refer to Germain et al. (2015) for further discussion regarding the majority vote versus the Gibbs method for classification settings. Here, we show that, as in the classification setting, the majority vote treatment rule can inherit desirable qualities from the Gibbs treatment rule in the budget constrained treatment rule setting.

While the majority vote rule $f_{\mathrm{mv},\rho}$ is not guaranteed to satisfy the same budget as its Gibbs counterpart $f_{G,\rho}(x)$, we can still show that when $f_{G,\rho}(x)$ is close to $f^*_{B(\rho)}(x)$, the optimal rule for its budget,

$$B(\rho) = K(f_{G,\rho}),$$

then $f_{\mathrm{mv},\rho}$ will also be close to $f^*_{B(\rho)}$. The measurement of closeness, defined shortly, depends on both the welfare achieved and deviations from the budget $B(\rho)$. We will suppose that

$$B(\rho) > E_Q[\delta_c(X)1\{\delta_c(X) < 0\}]. \tag{27}$$

That is, $f_{G,\rho}$ does not achieve the exact cost of the cost-minimizing rule $1\{\delta_c(x) < 0\}$ for $x \in \mathcal{X}$. If (27) were an equality, the budget of $f_{G,\rho}$ would be such that a policy maker faced with this budget would need to ignore welfare and seek the lowest cost rule. Hence, when we are interested in maximizing welfare with a budget constraint, it is reasonable to rule out the case where the solution to the policy maker's problem is to ignore welfare and seek the lowest cost. In addition to (27), we will assume that $\delta_y(X)$ and $\delta_c(X)$ have bounded densities so that optimal solution to the decision makers in Theorem 3.1 is deterministic.

Under (27), Assumption 3.1, and the condition that $\delta_y(X)$ and $\delta_c(X)$ have bounded densities, Theorem 3.1 yields that the optimal budget-constrained policy for the budget $B(\rho)$ of the Gibbs rule $f_{G,\rho}$ is of the form

$$f^*_{B(\rho)}(x) = 1\{\delta_y(x) - \eta_{B(\rho)}\delta_c(x) > 0\}, \quad x \in \mathcal{X}, \tag{28}$$

for a constant $\eta_{B(\rho)}$. It also follows from Theorem 3.1 that either $\eta_{B(\rho)} = 0$ and $K(f^*_{B(\rho)}) < B(\rho)$ or else $\eta_{B(\rho)} > 0$ and $K(f^*_{B(\rho)}) = B(\rho)$. Recalling the definition of the welfare-regret under a budget constraint in (11),

$$R_{B(\rho)}(f) \equiv W\left(f^*_{B(\rho)}\right) - W\left(f\right),$$

it is clear that $R_{B(\rho)}(f_{G,\rho})$ is non-negative. It is small only when $f_{G,\rho}$ attains a welfare that is close to the budget optimal rule in its own budget class. We will show that when $R_{B(\rho)}(f_{G,\rho})$ is small, $f_{\mathrm{mv},\rho}$ has similar welfare to the optimal policy $f^*_{B(\rho)}$ and is unlikely to violate the budget $B(\rho)$ by a large amount.

First note that if a decision maker faced a budget of $B(\rho)$, it would be reasonable to seek a rule $f : \mathcal{X} \to [0, 1]$ that minimizes

$$L_{B(\rho)}(f) \equiv E_Q\left[\left(\delta_y(X) - \eta_{B(\rho)}\delta_c(X)\right)\left(f^*_{B(\rho)}(X) - f(X)\right)\right],$$

with the associated loss function

$$\begin{aligned}\ell_{B(\rho)}(f, x) &= \left(\delta_y(x) - \eta_{B(\rho)}\delta_c(x)\right)\left(f^*_{B(\rho)}(x) - f(x)\right) \\ &= \begin{cases} 0 & \text{if } f^*_{B(\rho)}(x) = f(x), \\ \left|\delta_y(x) - \eta_{B(\rho)}\delta_c(x)\right| & \text{if } f^*_{B(\rho)}(x) \neq f(x). \end{cases}\end{aligned}$$

By the form of $f^*_{B(\rho)}$ in (28), $L_{B(\rho)}(f)$ is non-negative and attains the value zero only when $f(X) = f^*_{B(\rho)}(X)$ almost surely. Of course, such a loss function cannot yield an estimation strategy directly because δ_y, δ_x, and $\eta_{B(\rho)}$ are unknown. However, when $L_{B(\rho)}(f)$ is small, this means we are unlikely to encounter a set of co-variates X for which f assigns treatment and $\eta_{B(\rho)}\delta_c(X)$ exceeds $\delta_y(X)$ by a large amount. We have the following result

Theorem 4.4 *Let $\rho \in \mathcal{P}_\pi(\Theta)$. Let Assumptions 3.1 and 3.2 hold and also assume that (27) holds and $\delta_c(X)$ and $\delta_y(X)$ have bounded densities so that $E_Q[1\{\delta_y(X) = \eta_{B(\rho)}\delta_c(X)\}] = 0$. Then*

$$L_{B(\rho)}\left(f_{\mathrm{mv},\rho}\right) \leq 2R_{B(\rho)}\left(f_{G,\rho}\right).$$

We note that the expectation in the definition of $L_{B(\rho)}(f)$ is taken with respect to a draw from the target population. When ρ is dependent on the sample data, the result and proof still hold, conditional on the estimated rule or sample, provided that (27) can be assumed to hold almost surely for ρ or with high probability if considering probabilistic bounds such as those in Sections 4.1 and 4.2. This is reasonable to assume for $\hat{\rho}_{\lambda,u}$, particularly when u is not so large that no treatments will be assigned. The notion that, for appropriately chosen values of λ, $R_{B(\hat{\rho}_{\lambda,u})}(f_{G,\hat{\rho}_{\lambda,u}})$ is small is exactly the implication of Theorems 4.2 (b) and 4.3 (b).

For example, assume that the conditions of Theorem 4.3 hold, take $\lambda = \kappa\sqrt{nq}/(M_y + uM_c)$ (althought we continue to write λ to reduce clutter in the notation), and suppose that (27) holds almost surely for $\rho = \hat{\rho}_{\lambda,u}$ and that $\delta_c(X)$ and $\delta_y(X)$ have bounded densities. Then by Theorem 4.3 (b), with probability at least $1-\epsilon$ it holds that

$$\begin{aligned} & R_{B(f_{G,\hat{\rho}_{\lambda,u}})}\left(f_{G,\hat{\rho}_{\lambda,u}}\right) \\ & \leq \operatorname*{arg\,min}_{\theta\in\Theta_{B(\hat{\rho}_{\lambda,u})}} \left[R_{B(f_{G,\hat{\rho}_{\lambda,u}})}(\theta)\right] + \frac{M_y + uM_c}{\kappa}\left[\overline{U}_2(n;q,u,\epsilon) + \overline{U}_3(n;q,u,\epsilon) + \overline{U}_4(n;q,u)\right] \end{aligned}$$

where we have done some simple algebra on the inequality of part (b) of Theorem 4.3 utilizing the definitions of regret and regret under a budget constraint. The above also uses the notation

$$R_{B(f_{G,\hat{\rho}_{\lambda,u}})}(\theta) = W\left(f^*_{B(f_{G,\hat{\rho}_{\lambda,u}})}\right) - W\left(f_\theta\right).$$

Recall that the terms outside of the arg min on the right-hand side of the above inequality are at most $\mathcal{O}(\log(n)n^{-1/2})$ for fixed $u \geq 0$, $q \in \mathbb{N}$ and $\epsilon \in (0,1]$. If, for example,

$$\delta_y(X) = \phi(X)^\intercal\theta_y, \quad \text{and} \quad \delta_c(X) = \phi(X)^\intercal\theta_c,$$

for some $\theta_y, \theta_c \in \mathbb{R}^q$, then we would have

$$\operatorname*{arg\,min}_{\theta\in\Theta_{B(\hat{\rho}_{\lambda,u})}} \left[R_{B(f_{G,\hat{\rho}_{\lambda,u}})}(\theta)\right] = 0.$$

In this case, the above combined with Theorem 4.4 produce that, with probability at least $1-\epsilon$, $L_{B(\hat{\rho}_{\lambda,u})}(f_{\mathrm{mv},\hat{\rho}_{\lambda,u}})$ is bounded above by terms that are $\mathcal{O}(\log(n)n^{-1/2})$.

5 Simulation Study and Implementation Details

In this section we evaluate the proposed treatment assignment methodology in a simulation environment. We also discuss model estimation and implementation. Section 5.1 describes the simulation environment and findings. Section 5.2 describes a model estimation strategy using the Sequential Monte Carlo (SMC) approach and discusses the implementation choices utilized in the simulation.

5.1 Simulation Study

Here we consider assigning treatments utilizing $\hat{\rho}_{\lambda,u}$ in two simulation environments[4]. The main takeaway is as follows. In the statistical settings considered where a strong alternative methodology is available, our method does as well as the state-of-the-art alternative. However, while we obtain competitive performance in settings with overlapping applicability, our methodology also extends to settings where the alternative method is not applicable.

The simulation environments for $(Y_0, Y_1, C_0, C_1, D, X)$ are described below. For each DGP considered, the treatment probability for sample observations is $e(x) = 1/2$ for all $x \in \mathcal{X}$, so that $D \sim \mathrm{Bern}(1/2)$ and is independent of the other variables.

[4]This section may be updated with additional simulation settings to better contrast strengths and weaknesses of the methodologies considered.

- DGP1: We let $X = (X_1, X_2, X_3)$ where $X_j \sim \text{Unif}(0,1)$ for $j = 1,2,3$ are i.i.d. uniform random variables. Potential outcomes are determined via

$$Y_d = 3 - 2X_1 + X_2 - X_3 + d\left(1 - X_1^2 + X_2 + X_3\right) + \epsilon.$$

Here, ϵ is a truncated standard normal (taking values in $[-2,2]$) that is independent of all other variables considered. Potential costs are determined via

$$C_0 = 0, \quad C_1 \sim \text{Binom}\left(5, \frac{1 - X_3^2 + 2X_2}{5}\right).$$

- DGP2: We let $X = (X_1, X_2, X_3)$ where $X_j \sim \text{Unif}(-1,1)$ for $j = 1,2,3$ are i.i.d. uniform random variables. Letting $\Lambda(v) = (1 + \exp(-v))^{-1}$ denote the logistic function, potential outcomes are determined via

$$Y_d = 1 + \max\{X_1 + X_2\} + X_3 + 2d\Lambda\left(\frac{2\left(X_1 + X_2\right)}{3}\right) + \epsilon, \quad d \in \{0,1\}.$$

Again ϵ is a truncated standard normal (taking values in $[-2,2]$) that is independent of all other variables considered. Potential costs are now determined via

$$C_0 = 0, \quad C_1 \sim \text{Binom}\left(5, \frac{2\Lambda\left(2X_2 + X_3\right)}{5}\right).$$

For each DGP, we run simulations on training sets with sample size $n = 1,000$. A testing sample of size $n_{\text{test}} = 10,000$ is utilized to approximate the true costs and benefits from of any considered policy. The testing sample is re-used across training sample iterations and we consider 100 training simulation replicates. Using knowledge of the DGP, for each testing set observation we can calculate the expected value of treating that sample, $E_Q[Y_{1,i} - Y_{0,i}|X_i]$, and the expected cost $E_Q[C_{1,i}|X_i]$ for all $i = 1, \dots, n_{\text{test}}$. Then, for any policy $f(x) : \mathcal{X} \to [0,1]$, we then use the testing set to obtain the (approximate) true gain and cost associated with the policy,

$$\text{Gain of } f = E_Q\left[\left(Y_1 - Y_0\right) f(X)\right],$$

and

$$\text{Cost of } f = E_Q\left[C_1 f(X)\right].$$

We consider both Gibbs (stochastic) treatment rules and majority vote treatment rules utilizing the Gibbs posterior $\hat{\rho}_{\lambda,u}$. In the our figures the Gibbs treatment rule is denoted PB-SA while the majority vote rule is denoted PB-MV. We consider values of u increasing from 0.2 to 4 over 40 evenly spaced steps. We also include the value $u = 0$ (corresponding to the unconstrained setting). For each choice of u, λ is chosen by 2-fold cross validation to maximize the average of $W_n(f) - uK_n(f)$ across hold-out folds, where $f = f_{G,\hat{\rho}_{\lambda,u}}$ for PB-SA and $f = f_{\text{mv},\hat{\rho}_{\lambda,u}}$ for PB-MV. The estimation procedure for objects involving Gibbs posteriors $\hat{\rho}_{\lambda,u}$ is described in Section 5.2. For each choice of u, we obtain a treatment rule with a different gain-cost pair. During the estimation stage, we also obtain cross-validation-based estimates of the gain and cost associated with each such rule, i.e. for each choice of u. We take $\mathcal{F}_\Theta$ with elements $f_\theta(x)$ as described in (9) and (10). In particular, we utilize polynomial transformations of the co-variate space of order at most 2. Let

$$\mathcal{F}_\Theta^{\text{poly}} = \left\{ f_\theta(x) = \sum_{j=1}^{10} \theta_j \phi_j(x), \theta \in \mathbb{R}^{10} \right\},$$

where the summation is over all monomials $\phi_j(x) = \prod_{\ell=1}^{3} x_\ell^{p_{j\ell}}$ with $\sum_{\ell=1}^{d} p_{j\ell} \leq 2$, $p_{j\ell} \in \{0,1,2\}$, and we are denoting elements of $\mathcal{X} = \mathbb{R}^3$ by (x_1, x_2, x_3). As a final step we take $\mathcal{F}_\Theta$ to be $\mathcal{F}_\Theta^{\text{poly}}$ where the monomials are normalized by their sample mean and standard deviation calculated from training data. We set π to be the standard multivariate normal distribution over $\mathbb{R}^{10}$.

As an alternative treatment rule, we consider the approach of Sun et al. (2021). When $C_0 \leq C_1$ almost surely, under mild conditions f_B^* in Theorem 3.1 takes the form

$$f_B^*(x) = 1\left\{\frac{\delta_y(x)}{\delta_c(x)} > \eta_B\right\},$$

for some (unknown) constant η_B. Sun et al. (2021) derive this property and show the conditional gain-cost ratio function $\delta_y(x)/\delta_c(x)$ can be estimated by re-purposing the instrumental random forest method of Athey et al. (2019). We denote their resulting treatment rule by IRF. Their approach estimates $\delta_y(X_i)/\delta_c(X_i)$ for each target population observation, ranks observations by their estimated ratios, and then assign treatments in decreasing order of the ranking until the budget is exhausted. We call this method of assignment, where a ranking or scoring is derived for members of the target population who are then treated in order of this score until the budget is reached, a "batch implementation." Note this requires the $C_0 \leq C_1$ assumption. Lastly, as a baseline rule, we estimate $\delta_y(x) = E_Q[Y_1 - Y_0 | X = x]$ using the causal random forest of Athey et al. (2019) and then use the resulting target population scores for a batch implementation similarly to the IRF method. This approach, which does not factor costs into the treatment decisions, is denoted "Ignore Cost" in our figures and IC in our discussion here.

We utilize "cost curves" to compare the gain-cost trade-off of the considered rules at different budget levels. These are constructed as follows. For a single training sample iteration, for each u we estimate a Gibbs rule (PB-SA) and a majority vote rule (PB-MV). We then evaluate the true cost and gain associated with these treatment rules (for different choices of u) using the test data. Once we have the true costs associated with these rules, we estimate the IRF ratios and IC CATE scores from the training sample and implement these rules via batch implementation in the testing data. For each u choice and for each PB-MV and PB-SA rule, the IRF and IC rules are implemented to stop assigning treatment when they reach the same target population cost as the PB-MV or PB-SA rule of interest. In this way we are comparing models with the same true costs.

For each training sample, the various true gain-cost points associated with different u choices for the PB-MV and PB-SA methods are plotted in gain-cost space along with the associated points for the IRF and IC models. The gain-cost curve for the iteration is then estimated by interpolating between these points. For a single training sample iteration, this process is illustrated in Figures 4 and 5 for DGP1. Then, the gain-cost curves for all training sample iterations are averaged (vertically) to produce the final (approximately) true gain-cost curves. This produces Figures 3 and 7 for DGP1 and DGP2, respectively. The black lines in these figures give the gain-cost pairs that would result from randomly assigning treatment in the target population until the particular cost level is achieved.

Above, the IRF and IC rules feature batch implementations while the PB-MV and PB-SA approaches do not. To compare like-for-like, the majority vote models associated with $\hat{\rho}_{\lambda,u}$ for a range of u values are amenable to a batch implementation method. This is notable as, when batch implementation is feasible, it controls costs accurately. We propose the following batch implementation method which we refer to as PB-Batch (and "PB, Batch" in Figures 2 and 6). From a minimal cost point (for example, zero), create equally spaced cost bins leading up to a desired budget level. First, one finds the PB-MV model (among different u choices) with an estimated cost nearest to the first cost bin cutoff. Rank the individuals in the target population according to the majority vote scores $\int_{\Theta} f_\theta(X_i) d\hat{\rho}_{\tilde{\lambda}_u,u}(\theta)$ for the given choice of u, where $\tilde{\lambda}_u$ reflects that this value is cross validated for u. Then assign treatments in descending order of score until the overall cost reaches that of the first bin's cost cutoff. Next, move to the next cost bin, find the PB-MV model with estimated cost closest to the new cost bin cutoff, rank the remaining untreated observations by the new majority vote scores and treat until the total cost reaches new cost bin's cutoff. This is continued until the desired budget level is reached. In our simulations, we created 20 equally spaced cost bins starting from zero. We treated each end point as a desired budget level and applied the PB-Batch implementation. For each budget level we also applied the IRF and IC rules. The associated gains were computed for each training iteration and then averaged over all training sample iterations to produce Figures 2 and 6 for DGP1 and DGP2, respectively.

The simulation results are presented in Figures 2-7 after Section 6. For DGP1, the batch implementation associated with the majority vote models performs quite closely to the IRF method. In particular, the performance of PB-Batch is the same as that for IRF for lower cost levels and decreases very slightly at higher budget levels. For example, PB-Batch and IRF each achieve a gain of 0.47 at the cost level of 0.25, whereas the IC approach that ignores costs achieves a gain of 0.32 at this cost level. At a cost of 0.75, PB-Batch achieves a gain of 1.01, the IRF attains a gain of 1.02, and the IC method attains a gain of 0.87. The overall simulation appears to have a margin of error around 0.005. Turning to the non-batch implemented rules PB-MV and PB-SA, we see that the majority vote rule, PB-MV, performs similarly to its batch counterpart and better than the Gibbs rule in these simulation environments. However, it has a slightly higher deterioration in performance at higher budget levels than PB-Batch. PB-SA, the stochastic Gibbs assignment method, performs slightly below the alternative PAC-Bayesian and IRF alternatives.

For DGP2, the simulation results for treatment rules other than the baseline IC method largely mirror the results for DGP1 except that here the underperformance of PB-SA is reduced at all budget levels. Now, it nearly matches the alternatives with slightly decreasing performances at higher cost levels. For DGP2, the IC rule that doesn't account for cost does poorly, matching or slightly underperforming a rule that randomly assigns treatment (the black lines in Figures 6 and 7). For example, at a budget of 0.5, the PB-Batch and IRF method have gains of 0.63, randomly assigning treatment has a gain of 0.5, and the IC model that ignores cost has a gain of only 0.48. The PB-MV and PB-SA rules yield gains of 0.61 and 0.60 at this cost level.

In the two considered simulation environments, the PB-Batch implementation of the PAC-Bayesian approach largely matches the performance of the causal forest method. This is followed closely by the majority vote method, which had gains within 0.00-0.02 of the higher performing batch methods at all budget levels in the two simulation environments. It is important to note that the PAC-Bayesian based rules can be applied in scenarios where the IRF cannot, which we discuss below. Given that the PAC-Bayesian methods match or nearly match the performance of the state-of-the art causal forest methods, extrapolating from our simulation environments we may hope that this performance is then extended to additional settings via the PAC-Bayesian approaches.

There are a number of settings where the causal forest based IRF method is not viable whereas the PAC-Bayesian approaches considered here remain applicable. Batch implementations are not always viable. The cost of a treatment may not be realized until sometime after treatment assignment and one may not always have the full target population available when the rule must be set. Batch implementations, where treatment is assigned until the budget is hit, could also be unacceptable to policy makers in settings where the "budget" is something with a negative connotation like a complication rate in a medical setting.

Lastly, the IRF rule can only be applied when $C_0 \leq C_1$ a.s., which rules out certain circumstances relevant to policy makers. For example, as noted in Sun (2021), Hendren and Sprung-Keyser (2020) identify fourteen welfare programs out of 133 considered that are estimated to have negative or zero net cost to the government. Sun (2021), for example, seeks to estimate a medicaid expansion policy under the budget constraint $B = 0$. She aims to find a policy that expands access without increasing overall cost. However, the constrained maximum welfare estimator employed by Sun (2021) may be difficult to implement when allowing for more flexible decision rule classes (she considers income threshold rules that vary with the number of children in a household). While one could estimate $\delta_y(x)$ and $\delta_c(x)$ separately and try to build a workaround via Theorem 3.1 when $C_0 > C_1$ is possible, the resulting ratio estimates will again require batch implementation, which adds a complication in this setting, and may have increased variance when the region of interest is an interval around zero.

5.2 Implementation and Estimation via Sequential Monte Carlo

To implement the Gibbs treatment rule associated with $\hat{\rho}_{\lambda,u}(\theta)$, we must evaluate treatment assignment probabilities or majority vote scores of the form

$$\int_{\Theta} f_{\theta}(x) d\hat{\rho}_{\lambda,u}(\theta), \quad x \in \mathcal{X}. \tag{29}$$

To do so, we utilize the Sequental Montel Carlo (SMC) procedure considered, for example, in Del Moral et al. (2006). While a Markov Chain Monte Carlo (MCMC) approach also could be derived, recently Ridgway et al. (2014) and Alquier et al. (2016) have highlighted the usefulness of the SMC procedure in PAC-Bayesian applications. One benefit is the ability to sample from a sequence of Gibbs posterior distributions for a range of λ and u values. This can ease the computational burden for cross validation. Here we discuss key elements of the approach, provide an estimation algorithm for our setting, and discuss implementation. We also discuss the choices utilized in implementing the procedure for Section 5.1.

Throughout, we make the following computational adjustment to the definition of $\hat{\rho}_{\lambda,u}$ in order to make the implementation choices for Section 5.1 applicable to more general settings. We define $\hat{\rho}_{\lambda,u}$ to be the distribution over Θ with RN derivative with respect to π given by

$$\frac{d\hat{\rho}_{\lambda,u}}{d\pi}(\theta) = \frac{\exp\left[-\lambda\left(u\overline{K}_n(\theta) - \overline{W}_n(\theta)\right)\right]}{Z(\lambda,u)}, \tag{30}$$

where

$$Z(\lambda,u) = \int_{\Theta} \exp\left[-\lambda\left(u\overline{K}_n(\theta) - \overline{W}_n(\theta)\right)\right] d\pi(\theta),$$

and

$$\overline{W}_n(\theta) = \frac{W_n(\theta)}{\frac{1}{n}\sum_{i=1}^n \delta_{y,i}} \text{ and } \overline{K}_n(\theta) = \frac{K_n(\theta)}{\frac{1}{n}\sum_{i=1}^n \delta_{y,i}}.$$

This adjustment is relevant when the average treatment effect is expected to be is positive. Clearly, if we denote $\hat{\delta}_y = n^{-1}\sum_{i=1}^n \delta_{y,i}$, the distribution $\hat{\rho}_{\lambda,u}$ in (30) is equivalent to $\hat{\rho}_{(\hat{\delta}_y\lambda),u}$ in Definition 3.3. In practice we choose λ via cross validation from a wide range of values.

For given choices of $\lambda > 0$ and $u \geq 0$, the SMC algorithm we adopt samples from $\hat{\rho}_{\lambda,u}$ to evaluate (29) by simulating a set of parameter draws from each of a sequence of distributions $\{\hat{\rho}_{\lambda_t,u_t}\}_{t=0}^{T}$. Here,

$$(0,0) = (\lambda_0, u_0) < (\lambda_1, u_1) < \cdots < (\lambda_T, u_T) = (\lambda, u)$$

is an increasing temperature ladder[5] that must be specified. By $(a,b) < (c,d)$, we mean that $a \leq c$ and $b \leq d$ while at least one of these inequalities is strict. Note that $\hat{\rho}_{\lambda_0,u_0} = \pi$, which the user may specify and we assume can be sampled from. The temperature ladder $\{(\lambda_t, u_t)\}_{t=0}^{T}$ is intended to be such that the corresponding distributions $\hat{\rho}_{\lambda_t,u_t}$ progress gradually from π to the target distribution $\hat{\rho}_{\lambda,u}$.

For each $t = 0, \ldots, T$, the SMC algorithm produces a set of N weighted samples $\{\Psi_t^{(i)}, \theta_t^{(i)}\}_{i=1}^{N}$ with $\Psi_t^{(i)} > 0$ and $\sum_{i=1}^{N} \Psi_t^{(i)} = 1$ where $\theta_t^{(i)} \in \Theta$ for all t and i in our setting. The set of parameter draws $\{\theta_t^{(i)}\}_{i=1}^{N}$ are referred to as particles (there are N weighted particles for each t). SMC combines MCMC moves with sequential importance sampling; we refer to Del Moral et al. (2006) for additional details and discussion. This produces weighted particles that emulate, in terms of computing expectations, samples from the distributions $\hat{\rho}_{\lambda_t,u_t}$ associated with

$$\frac{d\hat{\rho}_{\lambda_t,u_t}}{d\pi}(\theta) = \frac{\exp\left[-\lambda_t\left(u_t\overline{K}_n(\theta) - \overline{W}_n(\theta)\right)\right]}{Z_t}, \quad Z_t = \int_\Theta \exp\left[-\lambda_t\left(u_t\overline{K}_n(\theta) - \overline{W}_n(\theta)\right)\right] d\pi(\theta).$$

Under general conditions, for a $\hat{\rho}_{\lambda_T,u_T}$-integrable function $\varphi : \Theta \to \mathbb{R}$,

$$\sum_{i=1}^{N} \Psi_T^{(i)} \varphi\left(\theta_T^{(i)}\right) \overset{a.s.}{\to} \int_\Theta \varphi(\theta)\, d\hat{\rho}_{\lambda_T,u_T}(\theta) \quad \text{as } N \to \infty.$$

In our setting, we are interested in $\varphi(\theta) = f_\theta(x)$ to approximate (29) via

$$\sum_{i=1}^{N} \Psi_T^{(i)} f_{\theta_T^{(i)}}(x), \quad x \in \mathcal{X}.$$

Once we have run the SMC algorithm to yield $\{\Psi_T^{(i)}, \theta_T^{(i)}\}_{i=1}^{N}$ for a given pair $(\lambda, u) = (\lambda_T, u_T)$, the treatment probability or majority vote score for any value x in the covariate space can be computed as above. Alternatively, for example, we may be interested in $\varphi(\theta) = K_n(\theta)$, to approximate $K_n(f_{G,\hat{\rho}_{\lambda,u}})$.

The SMC algorithm utilized to estimate the treatment rules in Section 5.1 is detailed in the algorithm tables below. We set the input parameters τ_{ESS} and N there equal to $1/2$ and $1,000$, respectively. τ_{ESS} is an Effective Sample Size threshold criterion. When the variance of the weights at a given step t is too high, the SMC procedure utilizes a re-sampling step. This is referred to in Step 2 of the algorithm below. In our application we utilize systematic resampling, which is also outlined below. The choice of temperature ladder, additional algorithm details, and cross-validation points are detailed below the algorithm descriptions.

[5] The $\{\lambda_t\}_{t=0}^{\infty}$ sequence is traditionally called a temperature ladder in the SMC literature and in literature dealing with Gibbs or Boltzmann distributions. For the convenience of reference, we refer to the (λ_t, u_t) pairs as a temperature ladder.

Tempering SMC Algorithm

Input N (number of particles), $\tau_{\text{ESS}} \in (0,1)$ (ESS threshold), $\{(\lambda_t, u_t)\}_{t=1}^T$ (temperature ladder).

Output $\{\Psi_t^{(i)}, \theta_t^{(i)}\}_{i=1}^N$ for $t = 0, \dots, T$.

Step 1: initialization

- Set $t \leftarrow 0$. For $i = 1, \dots, N$, draw $\theta_0^{(i)} \sim \pi$ and set $\Psi_0^{(i)} \leftarrow 1/N$.

Iterate steps 2 and 3
Step 2: Resampling

- If

$$\left\{\sum_{i=1}^N \left(\Psi_t^{(i)}\right)^2\right\}^{-1} < \tau_{\text{ESS}} N,$$

  resample $\left\{\Psi_t^{(i)}, \theta_t^{(i)}\right\}_{i=1}^N$ yielding equally weighted resampled particles $\left\{\frac{1}{N}, \overline{\theta}_t^{(i)}\right\}_{i=1}^N$ and set $\left\{\Psi_t^{(i)}, \theta_t^{(i)}\right\}_{i=1}^N \leftarrow \left\{\frac{1}{N}, \overline{\theta}_t^{(i)}\right\}_{i=1}^N$. Otherwise, leave $\left\{\Psi_t^{(i)}, \theta_t^{(i)}\right\}_{i=1}^N$ unaltered.

Step 3: Sampling

- Set $t \leftarrow t + 1$; if $t = T + 1$, stop.

- For $i = 1, \dots, N$, draw $\theta_t^{(i)} \sim K_t(\theta_{t-1}^{(i)}, \cdot)$, where K_t is an MCMC kernel with invariant distribution $\hat{\rho}_{\lambda_t, u_t}$, and evaluate the unnormalized importance weights

$$\omega_t^{(i)}\left(\theta_{t-1}^{(i)}\right) = \exp\left[\lambda_{t-1}\left(u_{t-1}\overline{K}_n\left(\theta_{t-1}^{(i)}\right) - \overline{W}_n\left(\theta_{t-1}^{(i)}\right)\right) - \lambda_t\left(u_t\overline{K}_n\left(\theta_{t-1}^{(i)}\right) - \overline{W}_n\left(\theta_{t-1}^{(i)}\right)\right)\right].$$

- For $i = 1, \dots, N$, set

$$\Psi_t^{(i)} \leftarrow \frac{\Psi_{t-1}^{(i)}\omega_t\left(\theta_{t-1}^{(i)}\right)}{\sum_{j=1}^N \Psi_{t-1}^{(j)}\omega_t\left(\theta_{t-1}^{(j)}\right)}.$$

Resampling Algorithm (systematic resampling):

Input A set of (normalized) weights and associated particles, $\left\{\Psi_t^{(i)}, \theta_t^{(i)}\right\}_{i=1}^N$ for some $t \in \{0, \dots, T\}$.

Output Resampled particles for equal weighting, $\left\{\overline{\theta}_t^{(i)}\right\}_{i=1}^N$

- Draw $u \sim U\left[0, \frac{1}{N}\right]$.

- Compute cumulative weights $C^{(i)} = \sum_{m=1}^i \Psi_t^{(m)}$ for $i = 1, ..., N$.

- Set $m \leftarrow 1$.

- **For** $i = 1 : N$

  While $u < C^{(i)}$ **do** $\overline{\theta}_t^{(m)} \leftarrow \theta_t^{(i)}$.

  $m \leftarrow m + 1$, and $u \leftarrow u + 1/N$.

  End For

Some additional implementation details utilized in Section 5.1 are as follows. For the MCMC kernel in the sampling step of the SMC algorithm, we use a Gaussian random-walk Metropolis kernel with covariance matrix proportional to the empirical covariance matrix of the current set of particles. We scale the empirical covariance of the step t particles by $t^{-0.9}$. In practice, the scaling factor can be adjusted, possibly dynamically rather than with a general rule like $t^{-0.9}$, to ensure reasonable acceptance rates in the MCMC steps of the SMC algorithm. The prior was taken to be the standard multivariate normal distribution.

For a temperature ladder, we set $T = 800$ and considered values of $u_T \in \mathcal{W}_u$ where $\mathcal{W}_u$ contains zero and numbers increasing from 0.2 to 4 over 40 evenly spaced steps. For each value of u_T, $\{u_t\}_{t=0}^{200}$ increases from 0 to the given value of u_T in evenly spaced steps. For the remaining steps, $\{u_t\}_{t=201}^{800}$, u_t remains constant at the value of u_T under consideration. For each $u_T \in \mathcal{W}_u$ choice, we utilize the same $\{\lambda_t\}_{t=0}^{800}$, which follows a piece-wise linear structure: $\{\lambda_t\}_{t=0}^{200}$ increases from 0 to 4 in equally spaced steps. Then, $\{\lambda_t\}_{t=201}^{320}$ increases from 4 to $32 = 2^5$ in equally spaced increments, $\{\lambda_t\}_{t=321}^{470}$ increases from 2^5 to 2^8 in equally spaced increments, and $\{\lambda_t\}_{t=471}^{800}$ increases from 2^8 to 2^{10} in equally spaced increments. This produces a temperature complete ladder $\{(\lambda_t, u_t)\}_{t=0}^{800}$ for each choice of u_T. For λ values of interest less than 2^{10}, the temperature ladder is cut short (at fewer than 800 step) to end once λ_t reaches the desired value. We only cross validate for $\lambda \geq 4$ so that u_T is always the final value of u_t in the temperature ladder pairs. We let $\mathcal{W}_\lambda$ denote the elements of the full $\{\lambda_t\}_{t=0}^{800}$ sequence that were closest to elements in $\{2^2, (2^2+2^3)/2, 2^3, (2^3+2^4)/2, 2^4, \ldots, 2^{10}\}$. Rather than considering each value in $\{\lambda_t\}_{t=0}^{800}$ as a potential choice for a decision model via cross validation, for each $u \in \mathcal{W}_u$, λ was cross validated from values in $\mathcal{W}_\lambda$. The total set of potential (u, λ) pairs considered for final decision rules was then given by $\mathcal{W} = \mathcal{W}_u \times \mathcal{W}_\lambda$.

6 Conclusion

In this paper, we proposed a new approach to estimating treatment rules in a budget constrained setting. Utilizing the PAC-Bayesian framework, theoretical properties of interest were derived, including generalization bounds and oracle-type inequalities. The proposed rules minimize an upper bound on the welfare regret under an empirical budget constraint. Additionally, the procedures can yield budget-efficient rules. By this we mean that, under reasonable conditions, the regret is likely to approach that of the models in the underlying model class with the lowest possible regret for the same target population budget. Another benefit is that the proposed rules can take advantage of well developed Bayesian estimation machinery. This is in contrast to adaptations of the EWM approach, such as that in Sun (2021), which may become challenging to estimate when one is interested in models with greater flexibility than, for example, cut-off rules.

There are a number of considerations for future work. The bounds in Section 4.2, when a particular normal prior is utilized, feature a term that is $\sqrt{q}/\sqrt{n}$ where q is the dimension of the feature space. It would be of interest to determine if different prior choices, for example a sparsity-inducing prior or a normal prior with a different form for the covariance matrix, would allow for settings with higher dimensional feature spaces while maintaining the above notion of budget efficiency. Rather than using the Gibbs posterior to form treatment rules, it could also prove fruitful to approximate the Gibbs posterior with alternative distribution such as a normal distribution. So-called variational approximations of Gibbs posteriors for general PAC-Bayesian approaches are considered in Alquier et al. (2016). This could yield greater flexibility in terms of functional form constraints, beyond control over the variables included in the treatment rules featured here. It would also be of interest to incorporate estimated propensity scores into the PAC-Bayesian framework here and to explore how this impacts rates of convergence. Lastly, the analysis for balancing the primary welfare or regret objective against that of a secondary cost objective can be generalized to settings beyond the welfare-based potential outcomes framework here. Balancing a secondary objective of concern could also be of interest in classification or regression settings.

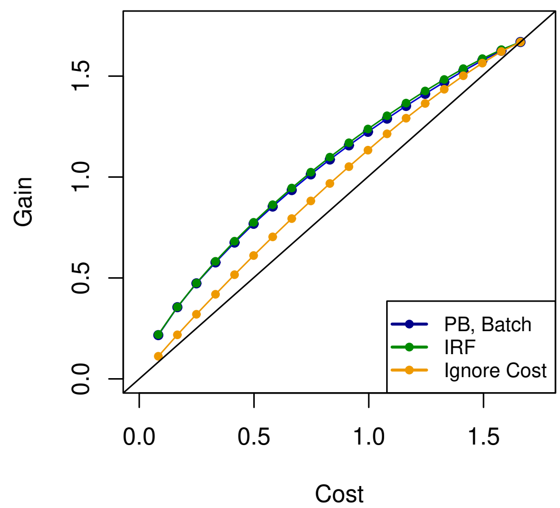

Figure 2: Cost curves when all methods utilize batch implementation for DGP1.

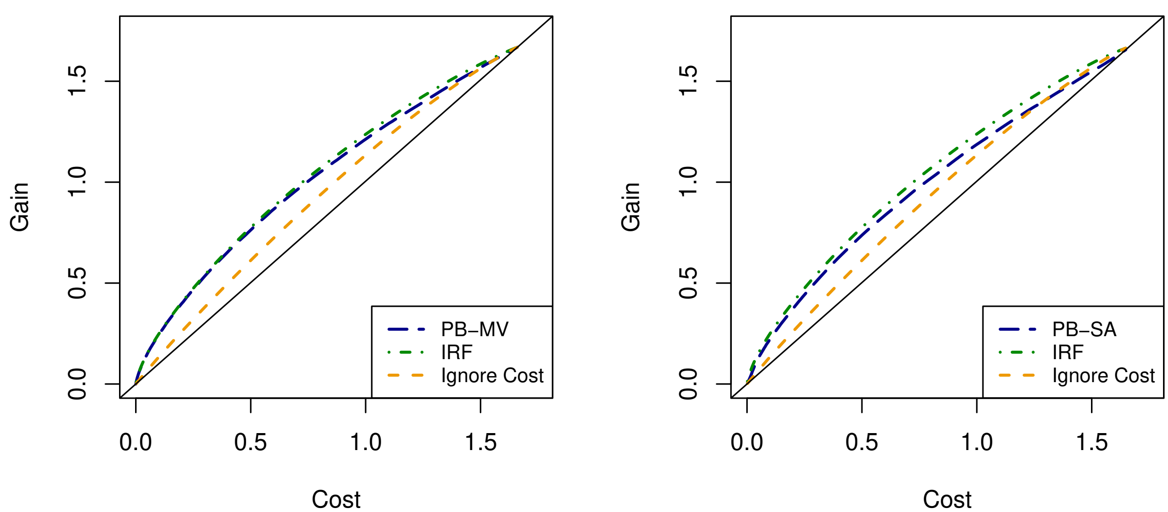

Figure 3: Under DGP1, cost curves for the PB-MV and PB-SA methods, which do not feature batch implementation, compared with the batch-implemented IRF and IC methods.

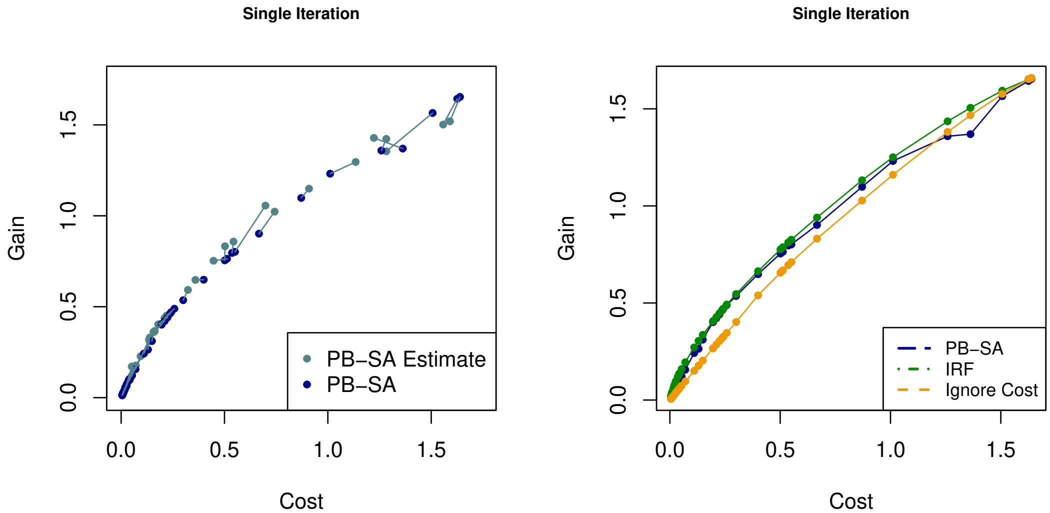

Figure 4: For DGP1, the left-hand side plots the estimated and actual cost-gain pairs (one point for each u) for a single training sample iteration for the PB-SA model. On the right-hand side the actual cost-gain pairs for PB-SA for various u values are plotted again, now compared with the actual cost-gain pairs associated with the IRF and IC rules that produce the same target population cost. The points on the right are then interpolated to produce cost-gain curve estimates for a single iteration. These curves are averaged (vertically) over all simulation iterations to produce the right-hand side of Figure 3.

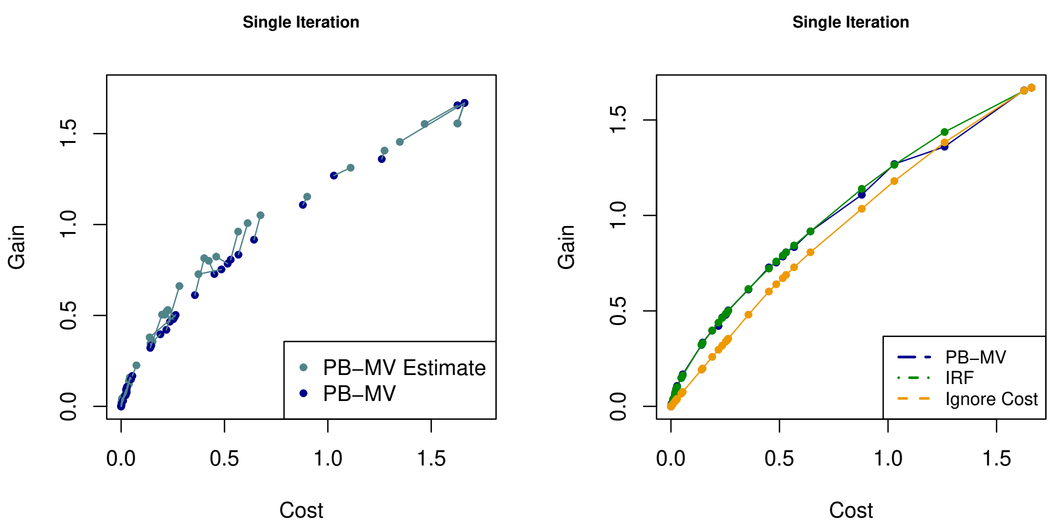

Figure 5: Illustrates a single training sample iteration for DGP1 when considering the PB-MV treatment model.

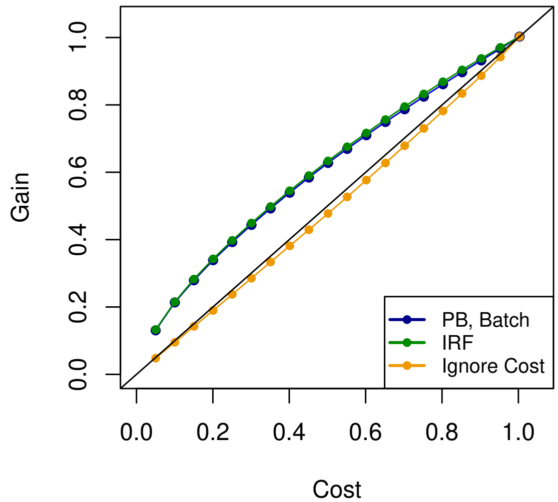

Figure 6: Cost curves when all methods utilize batch implementation for DGP2.

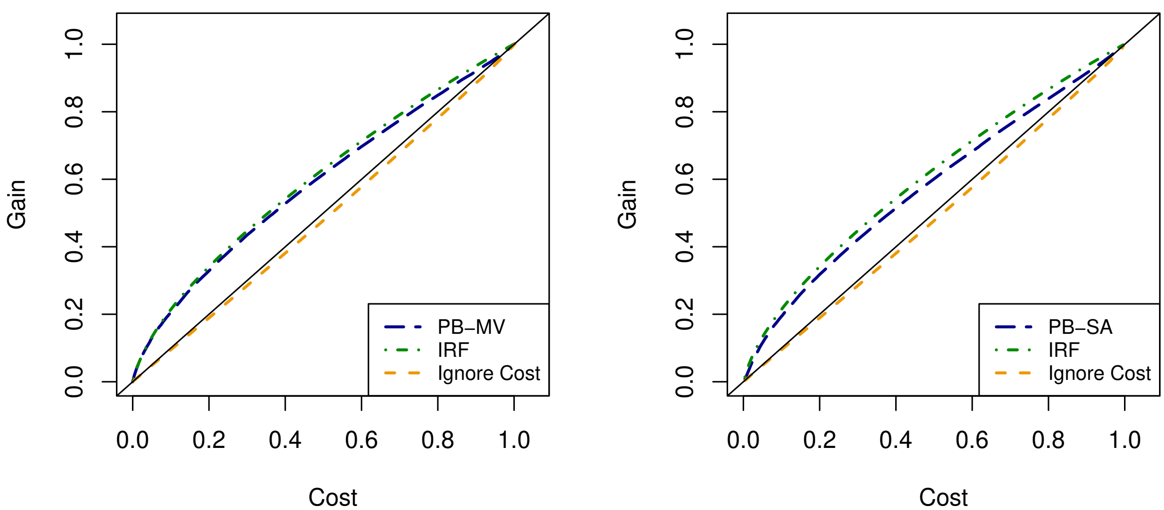

Figure 7: Under DGP2, cost curves for the PB-MV and PB-SA methods, which do not feature batch implementation, compared with the batch-implemented IRF and IC methods.

Appendix A: Proofs

A.1 Preliminaries and Adaptations From Earlier Literature to Our Setting

Here we consider preliminary properties to be utilized in subsequent analysis and recall results from the PAC-Bayesian literature that are also needed, sometimes with minor modifications. For the most part, proofs (and citations) are included for completeness even when a result is a fairly straightforward adaption.

Let $\mathcal{M}(\Theta)$ be the set of measurable functions on $(\Theta, \mathcal{B}_\theta)$ and let

$$\mathcal{M}_b^\pi(\Theta) = \left\{ A : A \in \mathcal{M}(\Theta) \text{ and } \int_\Theta \exp(A(\theta))\, d\pi(\theta) < \infty \right\},$$

which is a subset of $\mathcal{M}(\Theta)$ that has a finite exponential moment under π. We have the following lemma and corollary that will be utilized repeatedly in subsequent analysis. In particular they serve as a base in deriving Lemma 3.1 in Section 3.3.

Lemma A.1 *For $\pi \in \mathcal{P}(\Theta)$ and $A \in \mathcal{M}(\Theta)$ such that $-A \in \mathcal{M}_b^\pi(\Theta)$, let $\rho_{A,\pi} \in \mathcal{P}_\pi(\Theta)$ be the probability measure on Θ with the Radon–Nikodym (RN) derivative with respect to π given by*

$$\frac{d\rho_{A,\pi}}{d\pi}(\theta) = \frac{\exp(-A(\theta))}{\int_\Theta \exp\left(-A\left(\tilde{\theta}\right)\right) d\pi\left(\tilde{\theta}\right)}.$$

Then for any probability measure $\rho \in \mathcal{P}_\pi(\Theta)$ we have

$$\log\left[\int_\Theta \exp(-A(\theta))\, d\pi(\theta)\right] = -\left[\int_\Theta A(\theta)\, d\rho(\theta) + D_{\mathrm{KL}}(\rho, \pi)\right] + D_{\mathrm{KL}}(\rho, \rho_{A,\pi}). \qquad (31)$$

Proof of Lemma A.1. By definition,

$$\begin{aligned}
&D_{\mathrm{KL}}(\rho, \rho_{A,\pi}) \\
&= \int_\Theta \log\left[\frac{d\rho}{d\rho_{A,\pi}}(\theta)\right] d\rho(\theta) \\
&= \int_\Theta \log\left\{\frac{d\rho}{d\pi}(\theta)\left[\frac{d\rho_{A,\pi}}{d\pi}(\theta)\right]^{-1}\right\} d\rho(\theta) \\
&= \int_\Theta \left[\log\frac{d\rho}{d\pi}(\theta) - \log\frac{\exp(-A(\theta))}{\int_\Theta \exp\left(-A\left(\tilde{\theta}\right)\right) d\pi\left(\tilde{\theta}\right)}\right] d\rho(\theta) \\
&= \int_\Theta A(\theta)\, d\rho(\theta) + \int_\Theta \log\left[\int_\Theta \exp\left(-A\left(\tilde{\theta}\right)\right) d\pi\left(\tilde{\theta}\right)\right] d\rho(\theta) + \int_\Theta \left[\log\frac{d\rho}{d\pi}(\theta)\right] d\rho(\theta) \\
&= \int_\Theta A(\theta)\, d\rho(\theta) + \log\left[\int_\Theta \exp(-A(\theta))\, d\pi(\theta)\right] + \int_\Theta \left[\log\frac{d\rho}{d\pi}(\theta)\right] d\rho(\theta) \\
&= \int_\Theta A(\theta)\, d\rho(\theta) + \log\left[\int_\Theta \exp(-A(\theta))\, d\pi(\theta)\right] + D_{\mathrm{KL}}(\rho, \pi).
\end{aligned}$$

Hence,

$$\log\left[\int_\Theta \exp(-A(\theta))\, d\pi(\theta)\right] = -\left[\int_\Theta A(\theta)\, d\rho(\theta) + D_{\mathrm{KL}}(\rho, \pi)\right] + D_{\mathrm{KL}}(\rho, \rho_{A,\pi}).$$

■

Corollary A.1 *(a) Let $\lambda > 0$, $\pi \in \mathcal{P}(\Theta)$, and let $A \in \mathcal{M}(\Theta)$ be such that $-\lambda A \in \mathcal{M}_b^\pi(\Theta)$. Then*

$$\rho_{\lambda A,\pi} = \underset{\rho \in \mathcal{P}_\pi(\Theta)}{\arg\min} \left[\int_\Theta A(\theta)\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) \right],$$

and

$$\min_{\rho \in \mathcal{P}_\pi(\Theta)} \left[\int_\Theta A(\theta)\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) \right] = -\frac{1}{\lambda} \log \left[\int_\Theta \exp(-\lambda A(\theta))\, d\pi(\theta) \right].$$

(b) For any $\mathcal{A}(\cdot) \in \mathcal{M}_b^\pi(\Theta)$, $\pi \in \mathcal{P}(\Theta)$, $\rho \in \mathcal{P}_\pi(\Theta)$,

$$\int_\Theta \mathcal{A}(\theta)\, d\rho(\theta) \leq \log \left[\int_\Theta \exp(\mathcal{A}(\theta))\, d\pi(\theta) \right] + D_{\mathrm{KL}}(\rho, \pi).$$

Proof of Corollary A.1. Part (a). Note $\rho_{\lambda A,\pi} = \arg\min_{\rho \in \mathcal{P}_\pi(\Theta)} D_{\mathrm{KL}}(\rho, \rho_{\lambda A,\pi})$ as $D_{\mathrm{KL}}(\rho, \pi) \geq 0$ with equality if and only if $\rho = \pi$ π-almost surely. Replacing A with λA in Lemma A.1 and noting that the left-hand-side of (31) does not vary with ρ we have

$$\begin{aligned}
\rho_{\lambda A,\pi} &= \underset{\rho \in \mathcal{P}_\pi(\Theta)}{\arg\min} \left[D_{\mathrm{KL}}(\rho, \rho_{\lambda A,\pi}) \right] \\
&= \underset{\rho \in \mathcal{P}_\pi(\Theta)}{\arg\min} \left[\int_\Theta \lambda A(\theta) d\rho(\theta) + D_{\mathrm{KL}}(\rho, \pi) \right] \\
&= \underset{\rho \in \mathcal{P}_\pi(\Theta)}{\arg\min} \left[\int_\Theta A(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) \right].
\end{aligned}$$

By equation (31) we then have

$$\begin{aligned}
\min_{\rho \in \mathcal{P}(\Theta)} \left[\int_\Theta \lambda A(\theta)\, d\rho(\theta) + D_{\mathrm{KL}}(\rho, \pi) \right] &= \int_\Theta \lambda A(\theta)\, d\rho_{\lambda A,\pi}(\theta) + D_{\mathrm{KL}}(\rho_{\lambda A,\pi}, \pi) \\
&= -\log \left[\int_\Theta \exp(-\lambda A(\theta))\, d\pi(\theta) \right].
\end{aligned}$$

This is equivalent to the second statement in part (a).

Part (b) Taking $A = -\mathcal{A}$ in Lemma A.1, we obtain that for any probability measure $\rho \in \mathcal{P}_\pi(\Theta)$,

$$\log \left[\int_\Theta \exp(\mathcal{A}(\theta))\, d\pi(\theta) \right] = \left[\int_\Theta \mathcal{A}(\theta)\, d\rho(\theta) - D_{\mathrm{KL}}(\rho, \pi) \right] + D_{\mathrm{KL}}(\rho, \rho_{-A,\pi}). \tag{32}$$

Note that $D_{\mathrm{KL}}(\rho, \rho_{-A,\pi}) \geq 0$. It follows that

$$\begin{aligned}
\log \left[\int_\Theta \exp(\mathcal{A}(\theta))\, d\pi(\theta) \right] &= \left[\int_\Theta \mathcal{A}(\theta)\, d\rho(\theta) - D_{\mathrm{KL}}(\rho, \pi) \right] + D_{\mathrm{KL}}(\rho, \rho_{-A,\pi}) \\
&\geq \left[\int_\Theta \mathcal{A}(\theta)\, d\rho(\theta) - D_{\mathrm{KL}}(\rho, \pi) \right].
\end{aligned}$$

This implies that

$$\int_\Theta \mathcal{A}(\theta)\, d\rho(\theta) \leq D_{\mathrm{KL}}(\rho, \pi) + \log \left[\int_\Theta \exp(\mathcal{A}(\theta))\, d\pi(\theta) \right].$$

■

The following Theorem helps to produce PAC-Bayesian generalization bounds in our setting similar to counterparts in the classification literature. In particular, it essentially the same as Theorem 18 in Germain et al. (2015) with the loss function altered to the structure our setting; it is also similar to Theorem 4.1 in Alquier et al. (2016). The proof follows similar steps to those in Germain et al. (2015) and Alquier et al. (2016). We note that the proof applies to more general sample spaces, not just those following Assumption 3.1. We follow the current formulation to avoid additional exposition/notation.

Theorem A.1 *Let Assumptions 3.1 and 3.2 (i) hold and let $\pi \in \mathcal{P}(\Theta)$. Let $\ell(Z,\theta) : \mathcal{Z} \times \Theta \to \mathcal{R}$ denote a measurable loss function with range $\mathcal{R} \subseteq \mathbb{R}$. Define*

$$L(\theta) = E_P\left[\ell(Z,\theta)\right], \quad L_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \ell(Z_i,\theta),$$

and, for $\rho \in \mathcal{P}_\pi(\Theta)$,

$$L\left(f_{G,\rho}\right) = \int_\Theta L(\theta) d\rho(\theta), \quad L_n\left(f_{G,\rho}\right) = \int_\Theta L_n(\theta) d\rho(\theta).$$

Let $D : \mathcal{R} \times \mathcal{R} \to \mathbb{R}$ be any convex function and let $\lambda > 0$. Suppose

$$E_{P^n}\left[\int_\Theta \exp\left(\lambda D\left[L_n(\theta), L(\theta)\right]\right) d\pi(\theta)\right] \leq \exp\left(f(\lambda,n)\right), \tag{33}$$

where $f(\lambda,n) < \infty$ and may depend on λ and n. Then for any $\epsilon \in (0,1]$ it holds with probability at least $1-\epsilon$ that, simultaneously for all $\rho \in \mathcal{P}_\pi(\Theta)$,

$$D\left[L_n\left(f_{G,\rho}\right), L\left(f_{G,\rho}\right)\right] \leq \frac{f(\lambda,n) + \log\left(\frac{1}{\epsilon}\right) + D_{\mathrm{KL}}(\rho,\pi)}{\lambda}.$$

Proof of Theorem A.1. (33) implies that

$$\int_\Theta \exp\left(\lambda D\left[L_n(\theta), L(\theta)\right]\right) d\pi(\theta) < \infty,$$

holds almost surely. Therefore, applying Corollary A.1 (b) with $\mathcal{A}(\theta) = \lambda D[L_n(\theta), L(\theta)]$, the event

$$\left\{\int_\Theta \lambda D[L_n(\theta), L(\theta)] d\rho(\theta) \right.$$
$$\left. \leq \log\left[\int_\Theta \exp\left(\lambda D\left[L_n(\theta), L(\theta)\right]\right) d\pi(\theta)\right] + D_{\mathrm{KL}}(\rho,\pi) \text{ for all } \rho \in \mathcal{P}_\pi\left(\Theta\right) \text{ simultaneously}\right\},$$

occurs with probability one. Applying Jensen's inequality to the object on the left-hand-side of the inequality in this event, we have

$$P^n\left\{\lambda D[L_n\left(f_{G,\rho}\right), L\left(f_{G,\rho}\right)] \right.$$
$$\left. \leq \log\left[\int_\Theta \exp\left(\lambda D\left[L_n(\theta), L(\theta)\right]\right) d\pi(\theta)\right] + D_{\mathrm{KL}}(\rho,\pi) \text{ for all } \rho \in \mathcal{P}_\pi\left(\Theta\right) \text{ simultaneously}\right\},$$
$$= 1 \tag{34}$$

By Markov's inequality and then applying (33),

$$
\begin{aligned}
&P^n\left\{\int_\Theta \exp\left(\lambda D\left[L_n(\theta), L(\theta)\right]\right) d\pi(\theta) > \exp\left[f(\lambda, n) + \log\left(\frac{1}{\epsilon}\right)\right]\right\} \\
&\leq \frac{E_{P^n}\left[\int_\Theta \exp\left(\lambda D\left[L_n(\theta), L(\theta)\right]\right) d\pi(\theta)\right]}{\exp\left[f(\lambda, n) + \log\left(\frac{1}{\epsilon}\right)\right]} \\
&\leq \epsilon.
\end{aligned}
$$

Therefore,

$$
P^n\left\{\log\left[\int_\Theta \exp\left(\lambda D\left[L_n(\theta), L(\theta)\right]\right) d\pi(\theta)\right] \leq f(\lambda, n) + \log\left(\frac{1}{\epsilon}\right)\right\} \geq 1 - \epsilon
$$

Note that this high probability bound does not involve ρ. Combining it with (34), we have

$$
\begin{aligned}
&P^n\left\{D[L_n\left(f_{G,\rho}\right), L\left(f_{G,\rho}\right)] \leq \frac{f(\lambda, n) + \log\left(\frac{1}{\epsilon}\right) + D_{\mathrm{KL}}(\rho, \pi)}{\lambda} \text{ for all } \rho \in \mathcal{P}_\pi\left(\Theta\right) \text{ simultaneously}\right\} \\
&\geq 1 - \epsilon
\end{aligned}
$$

■

The following lemma will be combined with Theorem A.1 to produce Theorem A.2 below. The lemma yields a key step in adapting PAC-Bayesian bounds from the 0/1-loss classification literature to more general settings, a procedure utilized in Maurer (2004) and Germain et al. (2015). For us it will allow us to follow those author's adaption of a well known PAC-Bayesian bound, appearing, for example, in Seeger (2002), to more general settings. This then serves as a key input for producing Lemma A.3 following the analysis of Lever et al. (2010).

Lemma A.2 *Let X be any random variable taking values in $[0,1]$ with $EX = \mu$. Denote $\mathbf{X} = (X_1, \ldots, X_n)$ where $X_1, \ldots, X_n$ are iid realizations of X. Let $\mathbf{X}' = (X_1', \ldots, X_n')$ where $X_1', \ldots, X_n'$ are iid realizations of a Bernoulli random variable X' with probability of success μ. If $f : [0,1]^n \to \mathbb{R}$ is convex, then*

$$
E\left[f\left(\mathbf{X}\right)\right] \leq E\left[f\left(\mathbf{X}'\right)\right]
$$

Proof of Lemma A.2. This lemma is due to Maurer (2004). Another proof with more details is given in Germain et al. (2015); see Lemmas 51 and 52 there. For intuition, we can regard $\mathbf{X}'$ as a mean-preserving spread of $\mathbf{X}$ and $-f$ as the utility function. Then the lemma says that $\mathbf{X}$ is preferred by an expected utility maximizer having concave utility $-f\left(\cdot\right)$. ■

Now we use Lemma A.2 combined with Theorem A.1 to produce Theorem A.2 below, which is a variant of a well known bound appearing in Seeger (2002). To do this, we follow the analysis in Germain et al. (2015) to verify the bound for our setting. The proof closely follows that in Germain et al. (2015). Theorem 20 in Germain et al. (2015), for example, is a very similar and can apply to a variety of settings. The only difference here is that the structure of what plays the role of a loss function is stated differently in Theorem A.1.

The following notation is used in the next theorem. We let

$$
\mathrm{kl}(a, b) = a \log\frac{a}{b} + (1-a)\log\frac{1-a}{1-b}, \tag{35}
$$

and adopt the convention that $0 \log 0 = 0$, $a \log \frac{a}{0} = \infty$ if $a > 0$ and $0 \log \frac{0}{0} = 0$. Note that $\mathrm{kl}(a,b)$ is the KL-divergence between two Bernoulli random variables with success probabilities a and b.

Theorem A.2 *Set any prior $\pi \in \mathcal{P}(\theta)$ and $\epsilon \in (0,1]$. Let Assumption 3.1, 3.2, and 3.3 hold. Let $\ell(Z,\theta) : \mathcal{Z} \times \Theta \to [0,1]$ denote a measurable loss function with range $[0,1]$ (equipped with the standard Borel sigma field). Define*

$$L(\theta) = E_P\left[\ell(Z,\theta)\right], \quad L_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \ell(Z_i,\theta),$$

and, for $\rho \in \mathcal{P}_\pi(\Theta)$,

$$L\left(f_{G,\rho}\right) = \int_\Theta L(\theta) d\rho(\theta), \quad L_n\left(f_{G,\rho}\right) = \int_\Theta L_n(\theta) d\rho(\theta).$$

(a). With probability at least $1-\epsilon$, for all posteriors $\rho \in \mathcal{P}_\pi(\Theta)$ simultaneously it holds that

$$\mathrm{kl}\left(L_n\left(f_{G,\rho}\right), L\left(f_{G,\rho}\right)\right) \le \frac{1}{n}\left[D_{\mathrm{KL}}(\rho,\pi) + \log\left(2\sqrt{n}\right) + \log\frac{1}{\epsilon}\right].$$

(b). With probability at least $1-\epsilon$, for all posteriors $\rho \in \mathcal{P}_\pi(\Theta)$ simultaneously it holds that

$$\left(L_n\left(f_{G,\rho}\right) - L\left(f_{G,\rho}\right)\right)^2 \le \frac{1}{2n}\left[D_{\mathrm{KL}}(\rho,\pi) + \log\left(2\sqrt{n}\right) + \log\frac{1}{\epsilon}\right].$$

Proof of Theorem A.2. Part (a) Given the adaptation of Theorem A.1 to our setting, the proof follows that of Lemma 19 in Germain et al. (2015) or Theorem 1 in Maurer (2004). We will apply Theorem A.1 with

$$D\left(a,b\right) = \frac{n}{\lambda}\mathrm{kl}\left(a,b\right).$$

That $\mathrm{kl}(\cdot,\cdot)$ is convex follows from Theorem 2.7.2 of Cover and Thomas (2006). We must verify that the condition in (33) holds with $f(\lambda,n) = \log(2\sqrt{n})$. We will show that for any $\theta \in \Theta$,

$$E_{P^n}\left\{\exp\left[n\mathrm{kl}\left(L_n(\theta), L(\theta)\right)\right]\right\} \le \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1-\frac{k}{n}\right)^{n-k} \equiv \xi(n). \tag{36}$$

It can be shown (c.f. Lemma 19 in Germain et al. (2015) and the references therein) that $\sqrt{n} \le \xi(n) \le 2\sqrt{n}$. Then, by Assumption 3.3, we can reverse the order of integration on the object on the left hand side of condition 33, so that (36) yields that (33) holds with $f(\lambda,n) = \log(2\sqrt{n})$. All that remains is to prove (36).

Let $\theta \in \Theta$. First note that in edge cases where $L(\theta) = 0$ or $L(\theta) = 1$, we then have with probability one that $L_n(\theta) = 0$ or $L_n(\theta) = 1$, respectively, in which case $\mathrm{kl}(L_n(\theta), L(\theta)) = 0$ and (36) holds. Now consider any θ such that $L(\theta) \in (0,1)$. Note that

$$\exp\left\{\lambda D\left(L_n(\theta), L(\theta)\right)\right\} = \exp\left\{n \cdot \mathrm{kl}\left(\frac{1}{n}\sum_{i=1}^{n}\ell(Z_i,\theta), L(\theta)\right)\right\}$$

is a convex function of $\mathbf{X} = (\ell(Z_1,\theta), \ldots, \ell(Z_n,\theta))$. Then, by Lemma A.2,

$$E_{P^n}\left\{\exp\left\{\lambda D\left(L_n(\theta), L(\theta)\right)\right\}\right\} \le E\exp\left\{n \cdot \mathrm{kl}\left(\frac{1}{n}\sum_{i=1}^{n} X_i', L(\theta)\right)\right\}, \tag{37}$$

where $X_1', \ldots, X_n'$ are iid Bernoulli random variables with success probability $L(\theta)$ and the expectation on the right is taken with respect to their joint distribution. Denoting $X' = \sum_{i=1}^n X_i'$, we have

$$
\begin{aligned}
&E \exp\left\{ n \cdot \mathrm{kl}\left(\frac{1}{n}X', L(\theta)\right)\right\} \\
&\quad = E\left(\frac{\frac{1}{n}X'}{L(\theta)}\right)^{X'} \left(\frac{1-\frac{1}{n}X'}{1-L(\theta)}\right)^{n-X'} \\
&\quad = \sum_{k=0}^{n} \Pr\left(X' = k\right) \left(\frac{\frac{k}{n}}{L(\theta)}\right)^{k} \left(\frac{1-\frac{k}{n}}{1-L(\theta)}\right)^{n-k} \\
&\quad = \sum_{k=0}^{n} \binom{n}{k} \left(L(\theta)\right)^k \left(1-L(\theta)\right)^{n-k} \left(\frac{\frac{k}{n}}{L(\theta)}\right)^{k} \left(\frac{1-\frac{k}{n}}{1-L(\theta)}\right)^{n-k} \\
&\quad = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{k}{n}\right)^{k} \left(1-\frac{k}{n}\right)^{n-k} = \xi(n)
\end{aligned}
\tag{38}
$$

Therefore (36) holds for any $\theta \in \Theta$, completing the proof.

Part (b). Part (b) follows from part (a) with an application of Pinsker's inequality,

$$
2(a-b)^2 \leq \mathrm{kl}(a,b) \tag{39}
$$

■

The following lemma adapts Lemma 2 of Lever et al. (2010) to our setting, it will aid in removing a D_{KL} term from several bounds in Section 4.

Lemma A.3 *Let $\hat{\rho}_{\lambda,u}$ and $\rho^*_{\lambda,u}$ be as in Definition 3.2 with $\pi \in \mathcal{P}(\Theta)$, $\lambda > 0$, and $u \geq 0$. Let Assumptions 3.1, 3.2, and 3.3 hold and let $\epsilon \in (0,1]$. With probability at least $1-\epsilon$ it holds that*

$$
D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) \leq \frac{\lambda\sqrt{2}\left(M_y + uM_c\right)}{\kappa\sqrt{n}} \sqrt{\log\left(\frac{2\sqrt{n}}{\epsilon}\right)} + \frac{\lambda^2\left(M_y + uM_c\right)^2}{2n\kappa^2}.
$$

Proof of Lemma A.3. The proof follows that of Lemma 2 in Lever et al. (2010), with some minor adjustments, which are straightforward with Theorem A.2 taking the place of Seeger's (c.f. Seeger (2002)) bound in the setting of Lever et al. (2010). To lighten the exposition, we will write

$$
M(\theta; u) = R(\theta) + uK(\theta) \text{ and } M_n(\theta; u) = R_n(\theta) + uK_n(\theta)
$$

when writing the RN deriviatives of $\hat{\rho}_{\lambda,u}$ and $\rho^*_{\lambda,u}$ with respect to π and related objects. Note that for any $\theta \in \Theta$ we have $M_n(\theta; u) \in [-(M_y + uM_c)/2\kappa, (M_y + uM_c)/2\kappa]$ by Assumption 3.1 (iii) and (iv).

First, observe that

$$
\begin{aligned}
&D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) \\
&= \int_\Theta \log\left[\left(\frac{d\hat{\rho}_{\lambda,u}}{d\pi}(\theta)\right)\left(\frac{d\pi}{d\rho^*_{\lambda,u}}(\theta)\right)\right] d\hat{\rho}_{\lambda,u}(\theta) \\
&= \int_\Theta \left(\log\left[\frac{\exp\left(-\lambda M_n(\theta;u)\right)}{\exp\left(-\lambda M(\theta;u)\right)}\right] - \log\left[\frac{\int_\Theta \exp\left(-\lambda M_n(\theta,u)\right) d\pi(\theta)}{\int_\Theta \exp\left(-\lambda M(\theta;u)\right) d\pi(\theta)}\right]\right) d\hat{\rho}_{\lambda,u}(\theta) \\
&= \int_\Theta \log\left[\frac{\exp\left(-\lambda M_n(\theta;u)\right)}{\exp\left(-\lambda M(\theta;u)\right)}\right] d\hat{\rho}_{\lambda,u}(\theta) \\
&\qquad\qquad - \log\left[\frac{\int_\Theta \exp\left(-\lambda\left[M_n(\theta;u) + M(\theta;u) - M(\theta;u)\right]\right) d\pi(\theta)}{\int_\Theta \exp\left(-\lambda M(\theta;u)\right) d\pi(\theta)}\right] \\
&= \lambda \int_\Theta M(\theta;u) - M_n(\theta;u) d\hat{\rho}_{\lambda,u}(\theta) - \log\left[\int_\Theta \exp\left(\lambda\left[M(\theta;u) - M_n(\theta;u)\right]\right) d\rho^*_{\lambda,u}\right] \\
&\leq \lambda\left[\int_\Theta M(\theta;u) - M_n(\theta;u) d\hat{\rho}_{\lambda,u}(\theta) - \int_\Theta M(\theta;u) - M_n(\theta;u) d\rho^*_{\lambda,u}\right],
\end{aligned} \tag{40}
$$

where the last inequality follows from Jensen's inequality.

Next we utilize an Theorem A.2 (b). For the setting there, let

$$
\ell(Z,\theta) = \left(\ell_y(Z,\theta) + u\ell_c(Z,\theta) + \frac{M_y + uM_c}{2\kappa}\right)\left(\frac{\kappa}{M_y + uM_c}\right)
$$

where

$$
\ell_y(Z,\theta) = \left(\frac{YD}{e(X)} - \frac{Y(1-D)}{1-e(X)}\right)\left(f^*(X) - f_\theta(X)\right), \tag{41}
$$

$$
\ell_c(Z,\theta) = \left(\frac{CD}{e(X)} - \frac{C(1-D)}{1-e(X)}\right) f_\theta(X), \tag{42}
$$

and f^* is as in (7).

Note then that, by Assumption (3.1) (iii) and (iv), for all $\theta \in \Theta$, we have $\ell(Z,\theta) \in [0,1]$ almost surely. Additionally, we have

$$
\begin{aligned}
L(\theta) = E_P[\ell(Z,\theta)] &= \left(R(\theta) + uK(\theta) + \frac{M_y + uM_c}{2\kappa}\right)\left(\frac{\kappa}{M_y + uM_c}\right) \\
&= \left(M(\theta;u) + \frac{M_y + uM_c}{2\kappa}\right)\left(\frac{\kappa}{M_y + uM_c}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
L_n(\theta) &= \left(R_n(\theta) + uK_n(\theta) + \frac{M_y + uM_c}{2\kappa}\right)\left(\frac{\kappa}{M_y + uM_c}\right) \\
&= \left(M_n(\theta;u) + \frac{M_y + uM_c}{2\kappa}\right)\left(\frac{\kappa}{M_y + uM_c}\right).
\end{aligned}
$$

Given the above setting, we will apply Theorem A.2 (b). Note that in Theorem A.2, the prior π does not have to be the same as that used in the definition of $\hat{\rho}_{\lambda,u}$ and $\rho^*_{\lambda,u}$, provided that the posteriors of interest are still absolutely continuous with respect to the prior. Rather that utilizing the theorem with the π associated with $\hat{\rho}_{\lambda,u}$ and $\rho^*_{\lambda,u}$, we instead use $\rho^*_{\lambda,u}$ as the prior. Note this prior choice satisfies Assumption 3.3, i.e. it does not depend on the sample. Applying Theorem A.2 (b) and taking the square root of each side in the high probability bound there, utilizing posteriors $\rho = \hat{\rho}_{\lambda,u}$ and $\rho = \rho^*_{\lambda,u}$, with probability at least $1-\epsilon$ it holds simultaneously that

$$\int_\Theta L(\theta) - L_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) \le \frac{1}{\sqrt{2n}} \sqrt{D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) + \log\left(\frac{2\sqrt{n}}{\epsilon}\right)},$$

$$-\left(\int_\Theta L(\theta) - L_n(\theta) d\rho^*_{\lambda,u}(\theta)\right) \le \frac{1}{\sqrt{2n}} \sqrt{\log\left(\frac{2\sqrt{n}}{\epsilon}\right)}.$$

In terms of $M(\theta; u)$ and $M_n(\theta; u)$, this reads: with probability at least $1-\epsilon$, the following events holds simultaneously

$$\int_\Theta M(\theta) - M_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) \le \frac{M_y + uM_c}{\kappa\sqrt{2n}} \sqrt{D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) + \log\left(\frac{2\sqrt{n}}{\epsilon}\right)},$$

$$-\left(\int_\Theta M(\theta) - M_n(\theta) d\rho^*_{\lambda,u}(\theta)\right) \le \frac{M_y + uM_c}{\kappa\sqrt{2n}} \sqrt{\log\left(\frac{2\sqrt{n}}{\epsilon}\right)}.$$

Applying the above two inequalities to (40), we obtain

$$\begin{aligned} & D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) \\ & \le \frac{\lambda\left(M_y + uM_c\right)}{\kappa\sqrt{2n}} \sqrt{D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) + \log\left(\frac{2\sqrt{n}}{\epsilon}\right)} + \frac{\lambda\left(M_y + uM_c\right)}{\kappa\sqrt{2n}} \sqrt{\log\left(\frac{2\sqrt{n}}{\epsilon}\right)} \end{aligned}$$

Straightforward algebraic manipulations of the above produce that

$$\begin{aligned} & \left(D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right)\right)^2 - \frac{2\lambda\left(M_y + uM_c\right)}{\kappa\sqrt{2n}} \sqrt{\log\left(\frac{2\sqrt{n}}{\epsilon}\right)} D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) + \frac{\lambda^2\left(M_y + uM_c\right)^2}{2n\kappa^2} \log\left(\frac{2\sqrt{n}}{\epsilon}\right) \\ & \le \frac{\lambda^2\left(M_y + uM_c\right)^2}{2n\kappa^2} D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) + \frac{\lambda^2\left(M_y + uM_c\right)^2}{2n\kappa^2} \log\left(\frac{2\sqrt{n}}{\epsilon}\right), \end{aligned} \tag{43}$$

If

$$D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}\right) \le \frac{2\lambda\left(M_y + uM_c\right)}{\kappa\sqrt{2n}} \sqrt{\log\left(\frac{2\sqrt{n}}{\epsilon}\right)},$$

the statement of the lemma holds. Otherwise, this and the fact that $D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}) \ge 0$ imply that $D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u}) > 0$. Then, canceling out terms on either side of the inequality in (43) and dividing each side by $D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \rho^*_{\lambda,u})$ produces the statement of the lemma. ∎

The remainder of the section contains straightforward lemmas that will be utilized in proofs for results in Section 4 and one more substantial result adapted from Freund et al. (2004) that will conclude this subsection.

Lemma A.4 *Let Assumptions 3.1 and 3.2 hold. Let $\rho' \in \mathcal{P}(\Theta)$ be a (deterministic) probability that does not depend on the sample. Then*

$$P^n\left(\int_\Theta K_n(\theta)d\rho'(\theta) \le \int_\Theta K(\theta)d\rho'(\theta) + \sqrt{\frac{M_c^2 \log(1/\epsilon)}{2n\kappa^2}}\right) \ge 1-\epsilon.$$

Proof of Lemma A.4. Define the mapping

$$K(Z_1,\ldots,Z_n) = \int_\Theta K_n(\theta)d\rho'(\theta).$$

It is straightforward to check that, under Assumption 3.1 (iii), K satisfies the bounded differences property in Section 6.1 of Boucheron et al. (2013) with (in their notation) $c_i = M_c/(n\kappa)$ for $i=1,\ldots,n$. It follows by McDiarmid's inequality (c.f. McDiarmid (1989)) that, for any $t \ge 0$,

$$\begin{aligned}
&P^n\left(\int_\Theta K_n(\theta)d\rho'(\theta) - E_{P^n}\left[\int_\Theta K_n(\theta)d\rho'(\theta)\right] > t\right)\\
&= P^n\left(\int_\Theta K_n(\theta)d\rho'(\theta) - \int_\Theta K(\theta)d\rho'(\theta) > t\right) \le \exp\left\{-\frac{2n\kappa^2t^2}{M_c^2}\right\}.
\end{aligned}$$

Substituting $t = \sqrt{M_c^2\log(1/\epsilon)/(2n\kappa^2)}$, for any $\epsilon \in (0,1]$, this says

$$P^n\left(\int_\Theta K_n(\theta)d\rho'(\theta) - \int_\Theta K(\theta)d\rho'(\theta) > \sqrt{\frac{M_c^2 \log(1/\epsilon)}{2n\kappa^2}}\right) \le \epsilon.$$

The result follows by taking the compliment and rearranging terms. ■

Lemma A.5 *The KL divergence between $\rho : N(\mu_\rho, \Sigma_\rho)$ and $\pi : N(\mu_\pi, \Sigma_\pi)$ on $\mathbb{R}^q$, where μ_θ and μ_ρ are mean vectors and Σ_π and Σ_ρ are covariance matrices, is*

$$D_{\mathrm{KL}}(\rho,\pi) = \frac{1}{2}(\mu_\rho-\mu_\pi)'\Sigma_\pi^{-1}(\mu_\rho-\mu_\pi) + \frac{1}{2}\left[\mathrm{tr}\left(\Sigma_\rho\Sigma_\pi^{-1}\right)-q\right] - \frac{1}{2}\log\frac{\det(\Sigma_\rho)}{\det(\Sigma_\pi)}.$$

Proof of Lemma A.5. By definition and via simple calculations, we have

$$\begin{aligned}
&D_{\mathrm{KL}}(\rho,\pi)\\
&= -\frac{1}{2}E_{\theta\sim\rho}\left[\log\frac{\det(\Sigma_\rho)}{\det(\Sigma_\pi)} + (\theta-\mu_\rho)'\Sigma_\rho^{-1}(\theta-\mu_\rho) - (\theta-\mu_\pi)'\Sigma_\pi^{-1}(\theta-\mu_\pi)\right]\\
&= -\frac{1}{2}\log\frac{\det(\Sigma_\rho)}{\det(\Sigma_\pi)} - \frac{1}{2}\left[q - E_{\theta\sim\rho}(\theta-\mu_\rho+\mu_\rho-\mu_\pi)'\Sigma_\pi^{-1}(\theta-\mu_\rho+\mu_\rho-\mu_\pi)\right]\\
&= -\frac{1}{2}\log\frac{\det(\Sigma_\rho)}{\det(\Sigma_\pi)} - \frac{1}{2}\left[q - tr\left(\Sigma_\rho\Sigma_\pi^{-1}\right) - (\mu_\rho-\mu_\pi)'\Sigma_\pi^{-1}(\mu_\rho-\mu_\pi)\right]\\
&= \frac{1}{2}(\mu_\rho-\mu_\pi)'\Sigma_\pi^{-1}(\mu_\rho-\mu_\pi) + \frac{1}{2}\left[tr\left(\Sigma_\rho\Sigma_\pi^{-1}\right)-q\right] - \frac{1}{2}\log\frac{\det(\Sigma_\rho)}{\det(\Sigma_\pi)}.
\end{aligned}$$

■

The last results needed for our analysis are stated in the two lemmas below. The first is a more elementary property used in proving the second, which is utilized during a step in the proof of Theorem 4.2 in Section 4. Both are close adaptions of analysis in Freund et al. (2004). After a translation of the problem via Corollary A.1, we follow the method of proof there, adapting the analysis there in the 0/1 loss setting to ours with fairly straightforward modifications.

Lemma A.6 *For $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, and with $\{a_i\}_{i=1}^m$ such that $a_i \geq 0$ for all $i = 1, \ldots, m$, the function*

$$x \mapsto -\log \left[\sum_{i=1}^m a_i \exp[x_i] \right]$$

is concave.

Proof of Lemma A.6. Let $\alpha \in (0,1)$ and $x, y \in \mathbb{R}^m$. We will show that

$$K(x) = \log \left[\sum_{i=1}^m a_i \exp[x_i] \right]$$

is convex. Let $p = 1/\alpha$, $q = 1/(1-\alpha)$ and define $r_i = a_i^{1/p} \exp[\alpha x_i]$ and $s_i = a_i^{1/q} \exp[(1-\alpha) y_i]$. As $1/p + 1/q = 1$, by Hölder's inequality,

$$\sum_{i=1}^m r_i s_i \leq \left(\sum_{i=1}^m r_i^p \right)^{1/p} \left(\sum_{i=1}^m s_i^q \right)^{1/q}.$$

Taking the logarithm of each side and plugging in the definitions of p, q, r_i and s_i, this is equivalent to

$$K(\alpha x + (1-\alpha) y) \leq \alpha K(x) + (1-\alpha) K(y),$$

completing the proof. ■

The following lemma combines pieces of Lemmas 1 and 2 of Freund et al. (2004) and translates those results for the 0/1-loss setting to a useful ingredient for ours.

Lemma A.7 *Let $\hat{\rho}_{\lambda,u}$ and $\rho^*_{\lambda,u}$ be as in Definition 3.2 with $\pi \in \mathcal{P}(\Theta)$, $\lambda > 0$, and $u \geq 0$. Let assumptions 3.1, 3.2, and 3.3 hold. Then, for any $\epsilon \in (0,1]$, it holds that*

$$P^n \Bigg\{ \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + u \int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \pi)$$
$$\leq \int_\Theta R(\theta) d\rho^*_{\lambda,u}(\theta) + u \int_\Theta K(\theta) d\rho^*_{\lambda,u}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,u}, \pi) + \sqrt{\frac{(M_y + uM_c)^2 \log(1/\epsilon)}{2n\kappa^2}} \Bigg\} \geq 1 - \epsilon.$$

Proof of Lemma A.7. Define the mapping

$$K_u(Z_1, \ldots, Z_n) = \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + u \int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \pi).$$

Note that by Corollary A.1 (a), replacing $A(\theta)$ in the Corollary with $R(\theta) + uK_n(\theta)$,

$$K_u(Z_1, \ldots, Z_n) = -\frac{1}{\lambda} \log \left[\int_\Theta \exp[-\lambda (R_n(\theta) + uK_n(\theta))] d\pi(\theta) \right]. \tag{44}$$

First we show that for any $\epsilon \in (0,1]$ it holds that

$$P^n \left(K_u(Z_1, \ldots, Z_n) > E_{P^n}[K_u(Z_1, \ldots, Z_n)] + \sqrt{\frac{(M_y + uM_c)^2 \log(1/\epsilon)}{2n\kappa^2}} \right) \leq \epsilon. \tag{45}$$

To show this, for any $i \in \{1, \ldots, n\}$, let $Z_i' \in \mathcal{Z}$ and let $(Z_1, \ldots, Z_n) \in \mathcal{Z}^n$. Let $K_n(\theta)$ and $R_n(\theta)$ be computed utilizing $(Z_1, \ldots, Z_{i-1}, Z_i, Z_{i+1}, \ldots, Z_n)$ and let $K_n'(\theta)$ and $R_n'(\theta)$ be computed as $K_n(\theta)$ and $R_n(\theta)$ are, respectively, except utilizing the sample $(Z_1, \ldots, Z_{i-1}, Z_i', Z_{i+1}, \ldots, Z_n)$ instead of $(Z_1, \ldots, Z_{i-1}, Z_i, Z_{i+1}, \ldots, Z_n)$. Also, let $K_{n-i}(\theta)$ and R_{n-i} denote the computation of $K_n(\theta)$ and $R_n(\theta)$, respectively, except with the sample of size $n-1$ that drops observation Z_i. Then by construction $K_{n-i}(\theta) = K_{n-1}'(\theta)$ and $R_{n-i}(\theta) = R_{n-1}'(\theta)$. Under Assumptions 3.1 (iii) and (iv),

$$-\frac{M_y + uM_c}{2\kappa} \leq \ell_y(Z_i) + u\ell_c(Z_i, \theta) \leq \frac{M_y + uM_c}{2\kappa}$$

almost surely where ℓ_y and ℓ_c are defined in (41) and (42) and are summed over i in $R_n(\theta)$ and $K_n(\theta)$, respectively. It follows from (44) that,

$$\begin{aligned}
&|K_u(Z_1, \ldots, Z_{i-1}, Z_i, Z_{i+1}, \ldots, Z_n) - K_u(Z_1, \ldots, Z_{i-1}, Z_i', Z_{i+1}, \ldots, Z_n)| \\
&= \left| -\frac{1}{\lambda} \log \left[\frac{\int_\Theta \exp\left[-\lambda\left(R_n(\theta) + uK_n(\theta)\right)\right] d\pi(\theta)}{\int_\Theta \exp\left[-\lambda\left(R_n'(\theta) + uK_n'(\theta)\right)\right] d\pi(\theta)} \right] \right| \\
&\leq -\frac{1}{\lambda} \log \left[\left(\frac{\exp\left[-\lambda(M_y + uM_c)/(2n\kappa)\right]}{\exp\left[\lambda(M_y + uM_c)/(2n\kappa)\right]} \right) \left(\frac{\int_\Theta \exp\left[-\lambda\left(R_{n-i}(\theta) + uK_{n-i}(\theta)\right)\right] d\pi(\theta)}{\int_\Theta \exp\left[-\lambda\left(R_{n-i}'(\theta) + uK_{n-i}'(\theta)\right)\right] d\pi(\theta)} \right) \right] \\
&= \frac{M_y + uM_c}{n\kappa},
\end{aligned}$$

Thus, K_u satisfies the bounded differences property in Section 6.1 of Boucheron et al. (2013) with (in their notation) $c_i = (M_y + uM_c)/(n\kappa)$. By McDiarmid's inequality, (see McDiarmid (1989)), it holds that for any $t \geq 0$,

$$P^n\left(K_u(Z_1, \ldots, Z_n) - E_{P^n}\left[K_u(Z_1, \ldots, Z_n)\right] > t\right) \leq \exp\left(-\frac{2nt^2\kappa^2}{(M_y + uM_c)^2}\right).$$

Substituting $t = \sqrt{(M_y + uM_c)^2 \log(1/\epsilon)/(2n\kappa^2)}$, we obtain that for for any $\epsilon \in (0, 1]$,

$$P^n\left(K_u(Z_1, \ldots, Z_n) > E_{P^n}\left[K_u(Z_1, \ldots, Z_n)\right] + \sqrt{\frac{(M_y + uM_c)^2 \log(1/\epsilon)}{2n\kappa^2}}\right) \leq \epsilon.$$

Therefore (45) holds.

Next will show that

$$E_{P^n}\left[K_u(Z_1, \ldots, Z_n)\right] \leq \int_\Theta R(\theta) d\rho_{\lambda,u}^*(\theta) + u \int_\Theta K(\theta) d\rho_{\lambda,u}^*(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho_{\lambda,u}^*, \pi). \tag{46}$$

To do so, we follow arguments in Section 7 of Freund et al. (2004) with adjustments to suit our setting.

First note that by Corollary A.1 (a),

$$\begin{aligned}
\int_\Theta R(\theta) d\rho_{\lambda,u}^*(\theta) + u \int_\Theta K(\theta) d\rho_{\lambda,u}^*(\theta) &+ \frac{1}{\lambda} D_{\mathrm{KL}}(\rho_{\lambda,u}^*, \pi) \\
&= -\frac{1}{\lambda} \log \left[\int_\Theta \exp\left[-\lambda\left(R(\theta) + uK(\theta)\right)\right] d\pi(\theta) \right].
\end{aligned} \tag{47}$$

Next, by Assumption 3.1 and the definitions of $R(\theta)$ and $K(\theta)$, it follows that

$$-M_y - uM_c \leq R(\theta) + uK(\theta) \leq M_y + uM_c,$$

for all $\theta \in \Theta$. For any $\delta > 0$, let

$$\mathcal{B}_i = \{\theta \in \Theta : -(M_y + uM_c) + i\delta \le R(\theta) + uK(\theta) < -(M_y + uM_c) + (i+1)\delta\},$$

Then $\mathcal{B}_0, \ldots, \mathcal{B}_k$ with $k = \lfloor 2(M_y + uM_c)/\delta \rfloor$, form a partition of Θ. For $i \in \{0, \ldots, k\}$ such that $\pi(\mathcal{B}_i) > 0$, define

$$\tilde{\varepsilon}_i \equiv \frac{\int_{\mathcal{B}_i} R_n(\theta) + uK_n(\theta) d\pi(\theta)}{\pi(\mathcal{B}_i)},$$

Then, as π is independent of the sample by Assumption 3.3 and $E_{P^n}[R_n(\theta) + uK_n(\theta)] = R(\theta) + uK(\theta)$,

$$E_{P^n}[\tilde{\varepsilon}_i] = \frac{\int_{\mathcal{B}_i} R(\theta) + uK(\theta) d\pi(\theta)}{\pi(\mathcal{B}_i)} \le -(M_y + uM_c) + (i+1)\delta.$$

Combining this with the fact that $R(\theta) + uK(\theta) > -(M_y + uM_c) + i\delta$ for $\theta \in \mathcal{B}_i$,

$$\begin{aligned}
\int_\Theta \exp[-\lambda(R(\theta) + uK(\theta))] d\pi(\theta) &\le \sum \pi(\mathcal{B}_i) \exp[-\lambda(-(M_y + uM_c) + i\delta)] \\
&\le \sum \pi(\mathcal{B}_i) \exp[-\lambda(E_{P^n}[\tilde{\varepsilon}_i - \delta])] \\
&= \exp[\lambda\delta] \sum \pi(\mathcal{B}_i) \exp[-\lambda(E_{P^n}[\tilde{\varepsilon}_i])],
\end{aligned}$$

where the sums above are to be understood as summing over all $i \in \{0, \ldots, k\}$ such that $\pi(\mathcal{B}_i) > 0$. Taking the logarithm of each side of this inequality and multiplying by $-1/\lambda$, we have

$$\begin{aligned}
&-\frac{1}{\lambda}\log\left[\int_\Theta \exp[-\lambda(R(\theta) + uK(\theta))] d\pi(\theta)\right] \\
&\ge -\delta - \frac{1}{\lambda}\log\left[\sum \pi(\mathcal{B}_i)\exp[-\lambda(E_{P^n}[\tilde{\varepsilon}_i])]\right] \\
&\ge -\delta - \frac{1}{\lambda}E_{P^n}\left[\log\left(\sum \pi(\mathcal{B}_i)\exp[-\lambda\tilde{\varepsilon}_i]\right)\right] \qquad (48) \\
&= -\delta - \frac{1}{\lambda}E_{P^n}\left[\log\left(\sum \pi(\mathcal{B}_i)\exp\left[-\lambda\frac{\int_{\mathcal{B}_i} R_n(\theta) + uK_n(\theta)d\pi(\theta)}{\pi(\mathcal{B}_i)}\right]\right)\right] \\
&\ge -\delta - \frac{1}{\lambda}E_{P^n}\left[\log\left(\sum \pi(\mathcal{B}_i)\frac{\int_{\mathcal{B}_i}\exp[-\lambda(R_n(\theta) + uK_n(\theta))] d\pi(\theta)}{\pi(\mathcal{B}_i)}\right)\right] \qquad (49) \\
&= -\delta - \frac{1}{\lambda}E_{P^n}\left[\log\left(\int_\Theta \exp[-\lambda(R_n(\theta) + uK_n(\theta))] d\pi(\theta)\right)\right] \\
&= -\delta + E_{P^n}[K_u(Z_1, \ldots, Z_n)] \qquad (50)
\end{aligned}$$

In the above, (48) follows from an application of Jensen's inequality applied to the concave function

$$x \mapsto -\log\left(\sum_i \pi(\mathcal{B}_i)\exp[x_i]\right),$$

where the concavity of this function follows from Lemma A.6. (49) follows from another application of Jensen's inequality now applied to the convex function $\exp(x)$. (50) follows from (44). δ was arbitrary, so this produces

$$-\frac{1}{\lambda}\log\left[\int_\Theta \exp[-\lambda(R(\theta) + uK(\theta))] d\pi(\theta)\right] \ge E_{P^n}[K_u(Z_1, \ldots, Z_n)],$$

which, in light of (47), shows that (46) holds.

(45) and (46) together yield that

$$P^n\Bigg(K_u(Z_1,\ldots,Z_n) > \int_\Theta R(\theta)d\rho^*_{\lambda,u}(\theta) + u\int_\Theta K(\theta)d\rho^*_{\lambda,u}(\theta) + \frac{1}{\lambda}D_{\mathrm{KL}}(\rho^*_{\lambda,u},\pi) + \sqrt{\frac{(M_y+uM_c)^2\log(1/\epsilon)}{2n\kappa^2}}\Bigg) \le \epsilon,$$

which produces the statement of the lemma upon taking the compliment. ■

A.2 Proofs for Section 3

Proofs for Subsection 3.1: Statistical setting and policy maker's problem

We will utilize the following lemma in the proof of Theorem 3.1.

Lemma A.8 *Under the assumptions and setting of Theorem 3.1, let* $\delta_c^+(x) = \max(\delta_c(x), 0)$ *and* $\delta_c^-(x) = \max(-\delta_c(x), 0)$ *denote the positive and negative parts of* $\delta_c(x)$*, respectively. Define*

$$\beta(b) = E_Q[\delta_c(X)1\{\delta_y(X) > b\delta_c(X)\}],\ b \in \mathbb{R},$$

which is the expected budget of the non-stochastic treatment assigmnet rule $1\{\delta_y(x) > b\delta_c(x)\}$*.*
(i) Let $\eta_B = \inf\{b \ge 0 : \beta(b) \le B\}$*.* $\beta(b)$ *is non-increasing in* b *and* $0 \le \eta_B < \infty$*.*

(ii) Let

$$a_1 = \begin{cases} \frac{B-\beta(\eta_B)}{E_Q\left[\delta_c^+(X)1\{\delta_y(X)=\eta_B\delta_c(X)\}\right]} & \text{if } \beta(\eta_B) < B \text{ and } \eta_B > 0, \\ 0 & \text{else,} \end{cases}$$

and

$$a_2 = \begin{cases} \frac{\beta(\eta_B)-B}{E_Q\left[\delta_c^-(X)1\{\delta_y(X)=\eta_B\delta_c(X)\}\right]} & \text{if } \beta(\eta_B) > B, \\ 0 & \text{else.} \end{cases}.$$

Then these are well defined probabilities in that $\beta(\eta_B) < B$ *and* $\eta_B > 0$ *implies* $E_Q[\delta_c^+(X)1\{\delta_y(X) = \eta_B\delta_c(X)\}] > 0$*,* $\beta(\eta_B) > B$ *implies* $E_Q[\delta_c^-(X)1\{\delta_y(X) = \eta_B\delta_c(X)\}] > 0$*, and* $a_1, a_2 \in [0,1]$*. Furthermore, for* f^* *defined as in Theorem 3.1 with* $\eta_B, a_1,$ *and* a_2 *as above, when* $\beta(0) > B$ *it holds that*

$$E_Q[\delta_c(X)f_B^*(x)] = B.$$

Proof of Lemma A.8.

Proof of (i): To show $\beta(b)$ is non-increasing in b, write

$$\beta(b) = E_Q\left[\delta_c^+(X)1\{\delta_y(X) > b\delta_c(X)\}\right] - E_Q\left[\delta_c^-(X)1\{\delta_y(X) > b\delta_c(X)\}\right], \tag{51}$$

By definition of $\delta_c^+(x)$ and $\delta_c^-(x)$, $\delta_c^+(x)1\{\delta_y(x) - b\delta_c(x)\}$ is non-increasing in b and $\delta_c^-(x)1\{\delta_y(x) - b\delta_c(x)\}$ is non-decreasing in b for all $x \in \mathcal{X}$. It follows that $\beta(b)$ is non-increasing in b.

Checking $0 \le \eta_B < \infty$ translates to verifying that our form of policy assignment rule can meet the budget requirement. Let $\{b_n\}$ be any non-negative sequence such that $b_n \to \infty$. Then,

$E_Q|\delta_c(X)| < \infty$ and $E_Q|\delta_y(X)| < \infty$, equation (51), and an application of the dominated convergence theorem yield

$$\begin{aligned}\lim_{n\to\infty}\beta(b_n) &= \lim_{n\to\infty} E_Q\left[\delta_c^+(X)1\{\delta_y(X) > b_n\delta_c(X)\}\right] - \lim_{n\to\infty} E_Q\left[\delta_c^-(X)1\{\delta_y(X) > b_n\delta_c(X)\}\right] \\ &= 0 - E_Q[\delta_c^-(X)] < B.\end{aligned}$$

The inequality follows from the assumption that $B > E_Q[\delta_c(X)1\{\delta_c(X) < 0\}] = -E_Q[\delta_c^{-1}(X)]$. As $\beta(b)$ is non-increasing, we have either $\{b \geq 0 : \beta(b) \leq B\} = [r, \infty)$ or $\{b \geq 0 : \beta(b) \leq B\} = (r, \infty)$ for some $r \in \mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$. It follows that $0 \leq \eta_B < \infty$.

Proof of (ii): Let $b \in \mathbb{R}$. For any sequence $b_n \uparrow b$, by the dominated convergence theorem we have

$$\begin{aligned}\lim_{b_n\uparrow b}\beta(b_n) =& \lim_{b_n\uparrow b} E_Q[\delta_c^+(X)1\{\delta_y(X) > b_n\delta_c(X)\}] - \lim_{b_n\uparrow b} E_Q[\delta_c^-(X)1\{\delta_y(X) > b_n\delta_c(X)\}] \\ =& E_Q[\delta_c^+(X)1\{\delta_y(X) \geq b\delta_c(X)\}] - E_Q[\delta_c^-(X)1\{\delta_y(X) > b\delta_c(X)\}]. \\ =& E_Q[\delta_c^+(X)1\{\delta_y(X) > b\delta_c(X)\}] + E_Q[\delta_c^+(X)1\{\delta_y(X) = b\delta_c(X)\}] \\ & - E_Q[\delta_c^-(X)1\{\delta_y(X) > b\delta_c(X)\}]. \\ =& \beta(b) + E_Q[\delta_c^+(X)1\{\delta_y(X) = b\delta_c(X)\}]\end{aligned}$$

This yields

$$\lim_{x\to b^-}\beta(x) = \beta(b) + E_Q\left[\delta_c^+(X)\,1\left\{\delta_y(X) = b\delta_c(X)\right\}\right]. \tag{52}$$

Similar steps now starting with any sequence $b_n \downarrow b$ produce that

$$\lim_{x\to b^+}\beta(x) = \beta(b) - E_Q\left[\delta_c^-(X)\,1\left\{\delta_y(X) = b\delta_c(X)\right\}\right]. \tag{53}$$

As $\beta(\cdot)$ is non-increasing, it has at most countably many discontinuities, which occur at values b for which either $E_Q[\delta_c^+(X)1\{\delta_y(X) = b\delta_c(X)\}] > 0$ or $E_Q[\delta_c^-(X)1\{\delta_y(X) = b\delta_c(X)\}] > 0$ or both.

Now, if $B > \beta(\eta_B)$ and $\eta_B > 0$, by definition of η_B we have that $\beta(\eta') > B$ for any $\eta' < \eta_B$. Combined with (52), we obtain

$$\beta(\eta_B) < B \leq \beta(\eta_B) + E_Q\left[\delta_c^+1\{\delta_y(X) = \eta_B\delta_c(X)\}\right],$$

which implies that $E_Q[\delta_c^+(X)1\{\delta_y(X) = \eta_B\delta_c(X)\}] > 0$ and $a_1 \in [0, 1]$.

Next, if $B < \beta(\eta_B)$, by defnition of η_B we have $\beta(\eta') \leq B$ for any $\eta' > \eta_B$. Combining this with (53), we obtain

$$\beta(\eta_B) - E_Q\left[\delta_c^-(X)\,1\left\{\delta_y(X) = \eta_B\delta_c(X)\right\}\right] \leq B < \beta(\eta_B).$$

This implies $E_Q[\delta_c^-(X)1\{\delta_y(X) = \eta_B\delta_c(X)\}] > 0$ and $a_2 \in [0, 1]$.

For the last claim of (ii), write

$$\begin{aligned}E_Q\left[\delta_c(X)f_B^*(x)\right] =& E_Q\left[\delta_c(X)1\left\{\delta_y(X) > \eta_B\delta_c(X)\right\}\right] \\ & + a_1E_Q\left[\delta_c(X)1\left\{\delta_y(X) = \eta_B\delta_c(X)\right\}1\left\{\delta_c(X) > 0\right\}\right] \\ & + a_2E_Q\left[\delta_c(X)1\left\{\delta_y(X) = \eta_B\delta_c(X)\right\}1\left\{\delta_c(X) < 0\right\}\right].\end{aligned} \tag{54}$$

When $\beta(0) > B$, there are 3 scenarios for $\beta(\eta_B)$: (i) $\beta(\eta_B) = B$ and $\eta_B > 0$; (ii) $\beta(\eta_B) < B$ and $\eta_B > 0$; or (iii) $\beta(\eta_B) > B$ and $\eta_B \geq 0$. For scenario (i), we have $a_1 = a_2 = 0$ and the result holds as $E_Q[\delta_c(X)1\{\delta_y(X) > \eta_B\delta_c(X)\}] = \beta(\eta_B) = B$. For scenario (ii), $a_2 = 0$ and (54) becomes

$$E_Q\left[\delta_c(X)f_B^*(x)\right] = E_Q\left[\delta_c(X)1\left\{\delta_y(X) > \eta_B\delta_c(X)\right\}\right] \\ + \frac{B - \beta(\eta_B)}{E_Q\left[\delta_c^+(X)1\left\{\delta_y(X) = \eta_B\delta_c(X)\right\}\right]} E_Q\left[\delta_c^+(X)1\left\{\delta_y(X) = \eta_B\delta_c(X)\right\}\right] = B.$$

For scenario (iii), $a_1 = 0$ and (54) becomes

$$E_Q\left[\delta_c(X)f_B^*(x)\right] = E_Q\left[\delta_c(X)1\left\{\delta_y(X) > \eta_B\delta_c(X)\right\}\right] \\ - \frac{\beta(\eta_B) - B}{E_Q\left[\delta_c^-(X)1\left\{\delta_y(X) = \eta_B\delta_c(X)\right\}\right]} E_Q\left[\delta_c^-(X)1\left\{\delta_y(X) = \eta_B\delta_c(X)\right\}\right] = B.$$

This completes the proof of (ii). ■

Proof of Theorem 3.1.

The existence of $\eta_B \geq 0$, $a_1, a_2 \in [0, 1]$ such that either $\eta_B = a_1 = a_2 = 0$ (then f^* simplifies to f^*) when $K(f^*) \leq B$ or else (η_B, a_1, a_2) are such that $K(f^*) = B$ when $K(f^*) > B$ follows from Lemma A.8. To see this note $\beta(0) = K(f^*)$, where $\beta(\cdot)$ is defined in Lemma A.8. Thus, the statement about the budget being used entirely when $K(f^*) > B$ is stated directly in Lemma A.8. When $\beta(0) = K(f^*) \leq B$, η_B as defined in Lemma A.8 is equal to zero and then both $a_1 = a_2 = 0$ also from their definitions there.

Next we need to verify that f^* satisfies (6), i.e. is an optimal budget-constrained treatment policy. Let $r : \mathcal{X} \to [0, 1]$ denote any other stochastic treatment assignment rule that satisfies the budget constraint $K(r) \leq B$. As in Sun et al. (2021), we proceed by verifying that

$$E_Q\left[\delta_y(X)f_B^*(x)\right] \geq E_Q\left[\delta_y(X)r(X)\right].$$

By the definition of f^*, when $\delta_y(x) > \eta_B\delta_c(x)$ we also have $f_B^*(x) - r(x) \geq 0$. Hence $\delta_y(x)(f_B^*(x) - r(x)) \geq \eta_B\delta(x)(f_B^*(x) - r(x))$ in this case. When $\delta_y(x) < \eta_B\delta_c(x)$, we have $f_B^*(x) - r(x) \leq 0$ and hence $\delta_y(x)(f_B^*(x) - r(x)) \geq \eta_B\delta(x)(f_B^*(x) - r(x))$ in this case as well. It follows that

$$E_Q\left[\delta_y(X)\left(f_B^*(x) - r(X)\right)\right] \geq \eta_B E_Q\left[\delta_c(X)\left(f_B^*(x) - r(X)\right)\right]. \tag{55}$$

There are two possible scenarios: $K(f^*) \leq B$ or else $K(f^*) > B$. When $K(f^*) \leq B$, we have $\eta_B = 0$ and hence the right-hand-side of (55) is zero implying f^* is optimal. If, alternatively, $K(f^*) > B$, then we know that $K(f^*) = B$ and $K(r) \leq B$. Thus

$$E_Q[\delta_c(X)(f_B^*(x) - r(X))] = K(f^*) - K(r) \geq 0.$$

Now the right-hand-side of (55) is non-negative (as $\eta_B \geq 0$) and f^* is again optimal.

Lastly we need to show that if

$$E_Q[1\{\delta_y(X) = \eta_B\delta_c(X)\}] = 0, \tag{56}$$

then f^* is deterministic and unique (in an almost sure sense). It is clear from the form of f^* that it is almost surely equivalent to $1\{\delta_y(x) > \eta_B\delta_c(x)\}$ in this setting. Additionally, with the choices of η_B, a_1, a_2 given in Lemma A.8 this will be true for all $x \in \mathcal{X}$. To see that this follows from the proof of Lemma A.8, by (52) and (53) there, $\beta(b)$ is continuous in this scenario so that $\beta(\eta_B) = B$ when $\eta_B > 0$; this implies $a_1 = a_2 = 0$ when $\eta_B > 0$. When $\eta_B = 0$, we must have $\beta(0) \leq B$ and then again $a_1 = a_2 = 0$ as defined in Lemma A.8.

To check uniqueness, let $r(x)$ be any other treatment assignment rule that satisfies the budget constraint $K(r) \leq B$ and is not a.s. equal to $f^*_B(x)$. When $\eta_B = 0$, r must then assign treatment for a subset of $\mathcal{X}$ with positive probability that has negative CATE or else fail to assign treatment to some subset of $\mathcal{X}$ that has positive CATE with positive probability (or both). This results in lower expected welfare than f^*, so r cannot be optimal. When $\eta_B > 0$, the argument is similar to that showing f^* is optimal. When $\eta_B > 0$, its definition in Lemma A.8 indicates that $K(f^*) = \beta(0) > B$ (and from (54) in Lemma A.8, it follows that $\eta_B > 0$ in this case must be the unique choice for which the expected budget of f^* is B). $P(f^*_B(x) \neq r(x)) > 0$ then implies that for some subset of $\mathcal{X}$ with positive probability we must have $f^*_B(x) - r(x) > 0$ when $\delta_y(x) > \eta_B \delta_c(x)$ or else $f^*_B(x) - r(x) < 0$ when $\delta_y(x) < \eta_B \delta_c(x)$ (or both). This implies that the inequality in (55) is strict. As f^* uses up the entire budget (55) now implies the left-hand-side is strictly positive, which concludes the proof. ■

Proofs for Subsection 3.2: PAC-Bayesian Setting

Proof of Lemma 3.1. Part (a). There are two possible scenarios. First, if $\Lambda(0) \leq B$, i.e. the "cost" at $u = 0$ is within budget, then $\tilde{\rho}_{A,H,\lambda,0} \in \mathcal{E}_B$ and $\tilde{\rho}_{A,H,\lambda,0} = \rho_{\lambda A,\pi}$ in the notation of Corollary A.1. Then the result follows from Corollary A.1 (a). Note that this scenario captures the case when $B = \infty$, i.e. when there is no budget constraint.

In the second scenario, $\Lambda(0) > B$ (and $B < \infty$). Assume this is case for the remainder of the proof of part (a). First, we will show that this implies $\Lambda(u)$ is (strictly) decreasing in u and that there exists a unique $\overline{u}_B > 0$ such that $\Lambda(\overline{u}_B) = B$. Note below that at any point $u \geq 0$, because the derivatives of the integrands are dominated by integrable functions on intervals of the form $(u - a, u + b)$, some $a, b > 0$, and as $\Lambda(u)$ is easily extended in definition to negative values of u in neighborhoods of 0, we can exchange differentiation and integration. We have

$$
\begin{aligned}
&\frac{d}{du}\Lambda(u) \\
&= \frac{d}{du}\left[\left(\int_\Theta H(\theta)\exp\left[-\lambda\left(A(\theta) + uH(\theta)\right)\right]d\pi(\theta)\right)\left(\int_\Theta \exp\left[-\lambda\left(A(\theta) + uH(\theta)\right)\right]d\pi(\theta)\right)^{-1}\right] \\
&= -\lambda\int_\Theta H^2(\theta)\, d\tilde{\rho}_{A,H,\lambda,u}(\theta) + \lambda\left(\int_\Theta H(\theta)\, d\tilde{\rho}_{A,H,\lambda,u}(\theta)\right)^2 \\
&= -\lambda \mathbb{V}_{\theta\sim\tilde{\rho}_{A,H,\lambda,u}}\left[H(\theta)\right] \\
&< 0,
\end{aligned}
$$

where $\mathbb{V}_{\theta\sim\tilde{\rho}_{A,H,\lambda,u}}[H(\theta)]$ denotes the variance of $H(\theta)$ when $\theta \sim \tilde{\rho}_{A,H,\lambda,u}$. Note the strict inequality of the last line holds because the distribution of $H(\theta)$ induced by $\tilde{\rho}_{A,H,\lambda,u}$ is degenerate only when the distribution of $H(\theta)$ induced by π is degenerate. If this were the case, (16) would imply that $\Lambda(0) < B$. Hence the strict inequality when $\Lambda(0) \geq B$, which includes our current $\Lambda(0) > B$ scenario.

Now, note that (16) implies there exist $\epsilon_1, \eta > 0$ such that $\pi(\{\theta : H(\theta) \leq B - \epsilon_1\}) = \eta > 0$. Let ϵ_2 be such that $0 < \epsilon_2 < \epsilon_1$. Letting M_h and M_a be such that $|H(\theta)| \leq M_h$ and $|A(\theta)| \leq M_a$

for all θ (as these functions are assumed bounded), we have

$$
\begin{aligned}
&\int_\Theta H(\theta)\, d\tilde{\rho}_{A,H,\lambda,u}(\theta) \\
&\leq (B-\epsilon_2) + M_h \int_\Theta 1\{H(\theta) > B-\epsilon_2\}\, d\tilde{\rho}_{A,H,\lambda,u}(\theta) \\
&= (B-\epsilon_2) + M_h \frac{\int_\Theta 1\{H(\theta) > B-\epsilon_2\} \exp[-\lambda(A(\theta)+uH(\theta))]\, d\pi(\theta)}{\int_\Theta (1\{H(\theta) \leq B-\epsilon_1\} + 1\{H(\theta) > B-\epsilon_1\}) \exp[-\lambda(A(\theta)+uH(\theta))]\, d\pi(\theta)} \\
&\leq (B-\epsilon_2) + M_h \frac{\int_\Theta 1\{H(\theta) > B-\epsilon_2\} \exp[-\lambda(A(\theta)+uH(\theta))]\, d\pi(\theta)}{\int_\Theta 1\{H(\theta) \leq B-\epsilon_1\} \exp[-\lambda(A(\theta)+uH(\theta))]\, d\pi(\theta)} \\
&\leq (B-\epsilon_2) + M_h \left(\frac{\exp[-\lambda u(B-\epsilon_2)]}{\exp[-\lambda u(B-\epsilon_1)]}\right)\left(\frac{\exp[\lambda M_a]}{\eta \exp[-\lambda M_a]}\right) \\
&= (B-\epsilon_2) + \exp[-\lambda u(\epsilon_1-\epsilon_2)]\left(\frac{M_h \exp[2\lambda M_a]}{\eta}\right).
\end{aligned}
$$

As $\epsilon_1 - \epsilon_2 > 0$, for large enough values of u it holds that $\Lambda(u) < B$. Then, as $\Lambda(u)$ is continuous and strictly decreasing in u it follows that there is a unique $\overline{u}_B > 0$ such that $\Lambda(\overline{u}_B) = B$.

To finish the proof of Part (a) we need to show that when $\Lambda(0) > B$, $\tilde{\rho}_{A,H,\lambda,\overline{u}_B}$ is the optimal probability measure on Θ for the minimization problem. Replacing A in Corollary A.1 (a) with the $A + \overline{u}_B H$ as given above and noting that $\tilde{\rho}_{A,H,\lambda,\overline{u}_B} = \rho_{\lambda(A+\overline{u}_B H),\pi}$, we have that for any $\rho \in \mathcal{E}_B$,

$$
\begin{aligned}
&\tilde{\rho}_{A,H,\lambda,\overline{u}_B} \\
&= \underset{\rho\in\mathcal{P}_\pi(\Theta)}{\arg\min} \left[\int_\Theta \{A(\theta) + \overline{u}_B H(\theta)\}\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho,\pi)\right] &(57)\\
&= \underset{\rho\in\mathcal{P}_\pi(\Theta)}{\arg\min} \left[\int_\Theta \{A(\theta) + \overline{u}_B H(\theta)\}\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho,\pi) - \overline{u}_B B\right] \\
&= \underset{\mathcal{E}_B}{\arg\min} \left[\int_\Theta A(\theta)\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho,\pi) + \overline{u}_B\left(\int_\Theta H(\theta)\, d\rho(\theta) - B\right)\right] &(58)\\
&= \underset{\{\rho\in\mathcal{P}_\pi(\Theta):\int_\Theta H(\theta)d\rho(\theta)=B\}}{\arg\min} \left[\int_\Theta A(\theta)\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho,\pi) + \overline{u}_B\left(\int_\Theta H(\theta)\, d\rho(\theta) - B\right)\right], &(59)
\end{aligned}
$$

where the third equality holds in our specific setting because $\tilde{\rho}_{A,H,\lambda,\overline{u}_B} \in \mathcal{E}_B$ and $\mathcal{E}_B \subset \mathcal{P}_\pi(\Theta)$. The fourth equality follows similar reasoning. Next note that for any $\rho \in \mathcal{E}_B$, as $\overline{u}_B > 0$,

$$
\begin{aligned}
\int_\Theta A(\theta)\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho,\pi) &\geq \int_\Theta A(\theta)\, d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho,\pi) + \overline{u}_B\left(\int_\Theta H(\theta)\, d\rho(\theta) - B\right) &(60)\\
&\geq \int_\Theta A(\theta)\, d\tilde{\rho}_{A,H,\lambda,\overline{u}_B}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\tilde{\rho}_{A,H,\lambda,\overline{u}_B},\pi), &(61)
\end{aligned}
$$

where (61) follows from (58) and the fact that $\int_\Theta H(\theta) d\tilde{\rho}_{A,H,\lambda,\overline{u}_B}(\theta) = B$. Because the inequality in (60) is strict whenever $\int_\Theta H(\theta) d\rho < B$ it follows from (59) that $\tilde{\rho}_{A,H,\lambda,\overline{u}_B}$ is the argmin when $\Lambda(0) > B$, completing the proof of Part (a).

Part (b). Note that for $B' \geq B$, (16) will also hold for B' whenever it holds for B. We can then apply part (a) with B replaced by B' to obtain results for this (potentially) expanded budget, definition of $\overline{u}_{B'}$, etc. The first equality then follows from part (a) applied with the budget B' in place of B. For the second, let

$$h(u) = \int_\Theta A(\theta) d\tilde{\rho}_{A,H,\lambda,u}(\theta) + u \left(\int_\Theta H(\theta) d\tilde{\rho}_{A,H,\lambda,u}(\theta) - B' \right) + \frac{1}{\lambda} D_{\mathrm{KL}} \left(\tilde{\rho}_{A,H,\lambda,u}, \pi \right).$$

By the definition of $\overline{u}_{B'}$, we need to show that the supremum of $h(u)$ over $u \geq 0$ is achieved at $\overline{u}_{B'}$. Observe that by Corollary A.1 (a), as $\tilde{\rho}_{A,H,\lambda,u} = \rho_{\lambda(A+uH),\pi}$ in the notation there,

$$\begin{aligned} &\int_\Theta \{A(\theta) + uH(\theta)\} d\tilde{\rho}_{A,H,\lambda,u}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\tilde{\rho}_{A,H,\lambda,u}, \pi) \\ &= -\frac{1}{\lambda} \log \left[\int_\Theta \exp\left[-\lambda \left(A(\theta) + uH(\theta) \right) \right] d\pi(\theta) \right]. \end{aligned} \tag{62}$$

Utilizing this it is straightforward to derive that

$$\begin{aligned} \frac{d}{du} h(u) &= \frac{\int_\Theta H(\theta) \exp\left[-\lambda \left(A(\theta) + uH(\theta) \right) \right] d\pi(\theta)}{\int_\Theta \exp\left[-\lambda \left(A(\theta') + uH(\theta') \right) \right] d\pi(\theta')} - B' \\ &= \Lambda(u) - B', \end{aligned} \tag{63}$$

where we may exchange differentiation and expectation following similar reasoning as in the proof of Lemma 3.1.

In the proof of part (a) it is shown that $\Lambda(u)$ is strictly decreasing on $[0, \infty)$ when $\Lambda(0) \geq B$. When $\Lambda(0) > B' \geq B$ we then have $\Lambda(\overline{u}_{B'}) = B'$ with $\overline{u}_{B'} > 0$. It follows that the supremum of $h(u)$, which is continuous in u, is achieved at $\overline{u}_{B'}$. This is because, from (63), the derivative of $h(u)$ is positive on $[0, \overline{u}_{B'})$, zero at $\overline{u}_{B'}$ and decreasing on $(\overline{u}'_B, \infty)$. If $\Lambda(0) = B'$, we have that the supremum is achieved at 0, which is $\overline{u}_{B'}$ in this case, as the derivative of $h(u)$ is now zero at $u = \overline{u}_{B'} = 0$ and negative for $u \in (0, \infty)$. Conversely, if $\Lambda(0) < B'$, nearly identical steps to those in the proof of Lemma 3.1 with minor adjustments show that $\Lambda(u)$ is non-increasing in u for $u \geq 0$. Hence in this case the derivative of $h(u)$ is negative for $u \in [0, \infty)$ and the supremum is achieved at 0, which by the definition, is the value of $\overline{u}_B$ when $\Lambda(0) \leq B$.

Part (c). As in the proof of part (a), we again have $\frac{d}{du}\Lambda(u) = -\lambda \mathbb{V}_{\theta \sim \hat{\rho}_{A,H,\lambda,u}}[H(\theta)] < 0$ where now the strict inequality follows from the assumption in (18), which guarantees that the distribution of $H(\theta)$ induced by π is not degenerate. This in turn implies that the distribution of $H(\theta)$ induced by $\hat{\rho}_{A,H,\lambda,u}$ is not degenerate. Hence $\mathbb{V}_{\theta \sim \hat{\rho}_{A,H,\lambda,u}}[H(\theta)] > 0$ and the inequality is strict and $\Lambda(u)$ is strictly decreasing.

Next, we now check the first equality in (17). The proof is similar to that of part (a). There are two scenarios we must consider. The first is when $\Lambda(0) \leq B'$. As $\Lambda(u)$ is strictly decreasing, this occurs only if $\tilde{u} = 0$ and $B' = B_{\tilde{u}} = \Lambda(0)$. Also in this case, by definition $\overline{u}_{B'} = 0$ and the result follows from Corollary A.1 (a) as in the proof of part (a). The second scenario to consider is when $\Lambda(0) > B'$. Now as $\Lambda(u)$ is continuous in u, strictly decreasing, takes the value $\lambda(0) > B'$ at 0, and the value $B_{\tilde{u}} \leq B'$ at u, it follows that there is a unique positive real number $\overline{u}_{B'} \in (0, u]$ satisfying $\Lambda(\overline{u}_{B'}) = B'$. To see that the minimum is achieved at $\tilde{\rho}_{A,H,\lambda,\overline{u}_{B'}}$ when $\Lambda(0) > B'$, identical steps to the corresponding part of the proof of part (a) (at the end of that proof) can be followed with B' in place of B.

Lastly, we need to check the second equality in (17). The proof is similar to that of part (b). Let $h(u)$ for $u \geq 0$ be as in the proof of part (b). We need to verify that the supremum of $h(u)$ over $u \geq 0$ is attained at $\overline{u}_{B'}$. As in the proof of part (b), we obtain that $\frac{d}{du}h(u) = \Lambda(u) - B'$. By the assumption in (18), we have that $\Lambda(u)$ is strictly decreasing on $[0, \infty)$. First consider when $\Lambda(0) < B'$. Then $\overline{u}_{B'} = 0$ and

$$\Lambda(u) - B' \leq \Lambda(0) - B' < 0, \ u \geq 0.$$

Hence the derivative is strictly negative and $h(u)$ is decreasing for $u \geq 0$ so the supremum is attained at $0 = \overline{u}_{B'}$. Next consider when $\Lambda(0) = B'$. Then $\Lambda(u) - B'$ is 0 at $u = 0 = \overline{u}_{B'}$ and is negative on $(0, \infty)$, hence again the supremum is attained at $0 = \overline{u}_{B'}$. Lastly consider when $\Lambda(0) > B'$. In this case, we have shown there is a unique $\overline{u}_{B'} > 0$ with $\Lambda(\overline{u}_{B'}) = B'$. Using the continuity of $\Lambda(u)$ and the fact that it is decreasing in u, we obtain that the derivative of $h(u)$ is positive on $[0, \overline{u}_{B'})$, zero at $u = \overline{u}_{B'}$, and negative on $(\overline{u}_{B'}, \infty)$ so that again the supremum is attained at $\overline{u}_{B'}$. ■

A.3 Proofs for Section 4

A.3.1 Proofs for Subsection 4.1

Proof of Theorem 4.1. Part (a). When $V_n(\theta) = R_n(\theta)$, $V(\theta) = R(\theta)$, and $M_\ell = M_y$, we have the setup for Theorem A.1 with

$$\ell_v(Z, \theta) = \left(\frac{YD}{e(X)} - \frac{Y(1-D)}{1-e(X)} \right) (f^*(X) - f_\theta(X)),$$

$L(\theta) = V(\theta)$, and $L_n(\theta) = V_n(\theta)$. Note that, by Assumption 3.1, parts (iii) and (iv), we have that $-M_\ell/2\kappa \leq \ell_v(Z, \theta) \leq M_\ell/2\kappa$ a.s. Similarly, when $V_n(\theta) = K_n(\theta)$, $V(\theta) = K(\theta)$, and $M_\ell = M_c$, we have the setup for Theorem A.1 now with

$$\ell_v(Z, \theta) = \left(\frac{CD}{e(X)} - \frac{C(1-D)}{1-e(X)} \right) f_\theta(X),$$

and again taking $L(\theta) = V(\theta)$ and $L_n(\theta) = V_n(\theta)$. Again Assumption 3.1, parts (iii) and (iv), yields that $-M_\ell/2\kappa \leq \ell_v(Z, \theta) \leq M_\ell/2\kappa$ a.s.

Given this setup, we apply Theorem A.1 in the same way for either of the settings for $L(\theta), L_n(\theta)$ and M_ℓ. We need an appropriate choice for $D(\cdot, \cdot)$ and to then verify the condition in (33). Importantly, in either setting we have that, for any $\theta \in \Theta$, $\ell_v(Z_1, \theta), \ldots, \ell_v(Z_n, \theta)$ is an iid set of random variables taking values in $[-M_\ell/2\kappa, M_\ell/2\kappa]$ almost surely. For either $s \in \{-1, 1\}$, take $D[L_n(\theta), L(\theta)] = s(L_n(\theta) - L(\theta))$. We need to verify the condition in (33) and determine an appropriate $f(\lambda, n)$. Start with $s = 1$. Then, by Hoeffding's lemma (see, for example, Massart (2007), page 21), for any $\theta \in \Theta$,

$$\begin{aligned}
E_{P^n}\left[\exp\left(\lambda\left[L_n(\theta) - L(\theta)\right]\right)\right] &= E_{P^n}\left[\exp\left(\frac{\lambda}{n}\sum_{i=1}^{n}\left(\ell_v(Z_i, \theta) - E_P\left[\ell_v(Z_i, \theta)\right]\right)\right)\right] \\
&= \prod_{i=1}^{n} E_P\left[\exp\left\{\frac{\lambda}{n}\left(\ell_v(Z_i, \theta) - E_P\left[\ell_v(Z_i, \theta)\right]\right)\right\}\right] \\
&\leq \prod_{i=1}^{n} \exp\left(\frac{\lambda^2 M_\ell^2}{8\kappa^2 n^2}\right) = \exp\left(\frac{\lambda^2 M_\ell^2}{8\kappa^2 n}\right) \qquad (64)
\end{aligned}$$

Nearly identical steps in the $s=-1$ case, now applying Hoeffding's lemma to $-\ell_v(Z_i,\theta)$ produce that

$$E_{P^n}\left[\exp\left(\lambda\left[L(\theta)-L_n(\theta)\right]\right)\right]\le\exp\left(\frac{\lambda^2M_\ell^2}{8\kappa^2n}\right).\tag{65}$$

Integrating with respect to π, (64) and (65) yield that

$$\int_\Theta E_{P^n}\left[\lambda s\left(R_n(\theta)-R(\theta)\right)\right]d\pi(\theta)\le\exp\left(\frac{\lambda^2M_\ell^2}{8\kappa^2n}\right),\quad s\in\{-1,1\}.$$

We can reverse the order of integration on the left-hand of the above inequality, as π is independent of the sample by Assumption 3.3. Therefore, condition (33) in Theorem A.1 holds with $f(\lambda,n)=\lambda^2M_\ell^2/(8n\kappa^2)$. Applying Theorem A.1 completes the proof for Part (a).

Part (b). We utilize the same notation in terms of $\ell_v(Z,\theta)$ in the two scenarios for $V_n(\theta)$, $V(\theta)$, and M_ℓ as in part (a). Let E_1 denote the event that the following inequality holds,

$$D_{\mathrm{KL}}\left(\hat\rho_{\lambda,u},\rho^*_{\lambda,u}\right)\le\frac{\lambda\sqrt2\left(M_y+uM_c\right)}{\kappa\sqrt n}\sqrt{\log\left(2\sqrt n\right)+\log\frac2\epsilon}+\frac{\lambda^2\left(M_y+uM_c\right)^2}{2n\kappa^2}.\tag{66}$$

Note that by Lemma A.3, $P^n(E_1)\ge1-\epsilon/2$.

Next, let E_2 denote the event that the following inequality holds,

$$\left(\int_\Theta\left[V_n(\theta)-V(\theta)\right]d\hat\rho_{\lambda,u}(\theta)\right)^2\le\frac{M_\ell^2}{2n\kappa^2}\left[D_{\mathrm{KL}}\left(\hat\rho_{\lambda,u},\rho^*_{\lambda,u}\right)+\log\left(2\sqrt n\right)+\log\frac2\epsilon\right].\tag{67}$$

In the setup of Theorem A.2 (b), take

$$\ell(Z,\theta)=\left(\ell_v(Z,\theta)+\frac{M_\ell}{2\kappa}\right)\left(\frac{\kappa}{M_\ell}\right).$$

Then, for any $\theta\in\Theta$, $\ell(Z,\theta)\in[0,1]$ (P almost surely). Applying Theorem A.2 (b) yields that $P^n(E_2)\ge1-\epsilon/2$.

Then, the following a union bound argument,

$$\begin{aligned}P^n\left(E_1\cap E_2\right)&=1-P^n\left(E_1^c\cup E_2^c\right)\\&\ge1-P^n\left(E_1\right)-P^n\left(E_2\right)\\&\ge1-\frac\epsilon2-\frac\epsilon2=1-\epsilon,\end{aligned}$$

yields that events E_1 and E_2 occur jointly with probability greater than $1-\epsilon$. In the intersection of these events, plugging (66) into (67) produces the result in part (b).

Part (c). The proof follows similar steps to that in part (b). Define E_1 the same way as in part (b). Now, E_2 is defined to be the event that

$$\int_\Theta s\left[V_n(\theta)-V(\theta)\right]d\hat\rho_{\lambda,u}(\theta)\le\frac1\lambda D_{\mathrm{KL}}(\rho,\pi)+\frac1\lambda\left[\frac{\lambda^2M_\ell^2}{8n\kappa^2}+\log\frac2\epsilon\right].$$

By part (a), $P^n(E_1)\ge1-\epsilon/2$. Then, event E_2 is defined the same way is in the proof of part (b), $P(E_1\cap E_2)>1-\epsilon$ by a union bound argument, and combining the inequalities in E_1 and E_2 produces the statement of part (c). ■

Proof of Theorem 4.2.

Part (a). Let E_1 denote the event that, for all $\rho \in \mathcal{P}_\pi(\Theta)$ simultaneously it holds that

$$\int_\Theta R(\theta) d\rho(\theta) \leq \int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right]. \tag{68}$$

Let E_2 denote the event that, for all $\rho \in \mathcal{P}_\pi(\Theta)$ simultaneously it holds that

$$\int_\Theta R_n(\theta) d\rho(\theta) \leq \int_\Theta R(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right]. \tag{69}$$

Lastly, let $u^* = u^*(B, \lambda/2)$ as specified in Definition 3.3 and let E_3 denote the event that

$$\int_\Theta K_n(\theta) d\rho^*_{\lambda/2,u^*}(\theta) - \int_\Theta K(\theta) d\rho^*_{\lambda/2,u^*}(\theta) \leq \sqrt{\frac{M_c^2 \log\frac{3}{\epsilon}}{2n\kappa^2}}, \tag{70}$$

where $\rho^*_{\lambda/2,u}$ is given in Definition 3.2.

By Theorem 4.1 (a), applied to each $s \in \{-1, 1\}$ with $V_n(\theta) = R_n(\theta)$, $V(\theta) = R(\theta)$, and by Lemma A.4, respectively, we have

$$P^n(E_1) \geq 1 - \frac{\epsilon}{3}, \ P^n(E_2) \geq 1 - \frac{\epsilon}{3}, \text{ and } P^n(E_3) \geq 1 - \frac{\epsilon}{3}.$$

Applying a union bound argument as in the proof of Theorem 4.1 (b), it holds that $P^n(E_1 \cap E_2 \cap E_3) \geq 1 - \epsilon$. From the remainder of the proof, we work assuming the intersection of these three events. We show the event in Theorem 4.2 (a) is implied by their intersection, hence the event of interest contains this intersection and has probability greater than or equal $1 - \epsilon$.

We consider two possible scenarios in conjuncture with events E_1, E_2, and E_3. In the first scenario, suppose that

$$\int_\Theta K_n(\theta) d\rho^*_{\lambda/2,u^*}(\theta) \leq B. \tag{71}$$

In this case $\rho^*_{\lambda/2,u^*} \in \widehat{\mathcal{E}}_B$, where $\widehat{\mathcal{E}}_B$ is given by (20). Starting from (68) with $\rho = \hat{\rho}_{\lambda,\hat{u}}$,

$$\begin{aligned}
\int_\Theta R(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) &\leq \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,\hat{u}}, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] \\
&= \min_{\rho \in \widehat{\mathcal{E}}_B}\left\{\int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right]\right\} \\
&\leq \int_\Theta R_n(\theta) d\rho^*_{\lambda/2,u^*}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda/2,u^*}, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right].
\end{aligned}$$

Now, consider (69) with $\rho = \rho^*_{\lambda/2,u^*}$. Plugging this inequality into the right-hand side of the above inequality produces

$$\begin{aligned}\int_\Theta R(\theta)d\hat{\rho}_{\lambda,\hat{u}}(\theta) &\leq \int_\Theta R(\theta)d\rho^*_{\lambda/2,u^*}(\theta) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho^*_{\lambda/2,u^*},\pi) + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] \\ &= \min_{\rho\in\mathcal{E}_B}\left\{\int_\Theta R(\theta)d\rho(\theta) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho,\pi) + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right]\right\} \\ &\leq \min_{\rho\in\mathcal{E}_B}\left\{\int_\Theta R(\theta)d\rho(\theta) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho,\pi) + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right]\right\} + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}},\end{aligned}$$

where the equality in the second row follows from Corollary 3.1 (ii) and the final inequality holds as $\hat{u} \geq 0$. Thus the result of part (a) holds in the first scenario described by (71), noting that for $\rho \in \mathcal{P}(\Theta)$, $R(f_{G,\rho}) = \int_\Theta R(\theta)d\rho(\theta)$.

In the second and only remaining scenario, we consider when

$$\int_\Theta K_n(\theta)d\rho^*_{\lambda/2,u^*}(\theta) > B. \tag{72}$$

If we set

$$B' = \int_\Theta K_n(\theta)d\rho^*_{\lambda/2,u^*}(\theta), \tag{73}$$

then it holds that $\rho^*_{\lambda/2,u^*} \in \widehat{\mathcal{E}}_{B'}$. Again starting from the event in (68) with $\rho = \hat{\rho}_{\lambda,\hat{u}}$, we obtain

$$
\begin{aligned}
&\int_\Theta R(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) \\
&\leq \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,\hat{u}}, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] \\
&= \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,\hat{u}}, \pi) + \hat{u}\left(\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) - B\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] && (74) \\
&= \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) + \hat{u}\left(\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) - B'\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,\hat{u}}, \pi) \\
&\quad + \hat{u}\left(\int_\Theta K_n(\theta) d\rho^*_{\lambda/2,u^*}(\theta) - B\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] && (75) \\
&\leq \sup_{u\geq 0}\left[\int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + u\left(\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) - B'\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \pi)\right] \\
&\quad + \hat{u}\left(\int_\Theta K_n(\theta) d\rho^*_{\lambda/2,u^*}(\theta) - B\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] \\
&= \min_{\rho\in\widehat{\mathcal{E}}_{B'}}\left\{\int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi)\right\} \\
&\quad + \hat{u}\left(\int_\Theta K_n(\theta) d\rho^*_{\lambda/2,u^*}(\theta) - B\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] && (76) \\
&\leq \min_{\rho\in\widehat{\mathcal{E}}_{B'}}\left\{\int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi)\right\} + \hat{u}\sqrt{\frac{M_c^2 \log\frac{3}{\epsilon}}{2n\kappa^2}} + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] && (77) \\
&\leq \int_\Theta R_n(\theta) d\rho^*_{\lambda/2,u^*}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda/2,u^*}, \pi) + \hat{u}\sqrt{\frac{M_c^2 \log\frac{3}{\epsilon}}{2n\kappa^2}} + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] && (78) \\
&\leq \int_\Theta R(\theta) d\rho^*_{\lambda/2,u^*}(\theta) + \frac{2}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda/2,u^*}, \pi) + \hat{u}\sqrt{\frac{M_c^2 \log\frac{3}{\epsilon}}{2n\kappa^2}} + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] && (79) \\
&= \min_{\rho\in\mathcal{E}_B}\left\{\int_\Theta R(\theta) d\rho(\theta) + \frac{2}{\lambda} D_{\mathrm{KL}}(\rho, \pi) + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2 \log\frac{3}{\epsilon}}{2n\kappa^2}}\right\}. && (80)
\end{aligned}
$$

In the above, step (74) follows from the properties of $\hat{u} = \hat{u}(B, \lambda)$ in Corollary 3.1 (i). In step (75) we simply added and subtracted $\hat{u}B'$ with B' given in (73). Step (76) follows from Corollary 3.1 (i). Step (77) follows from (70) and the observation that $\int_\Theta K(\theta) d\rho^*_{\lambda/2,u^*}(\theta)$ is always less than or equal to B by Corollary 3.1 (ii). Step (78) follows from fact that $\rho^*_{\lambda/2,u^*} \in \widehat{\mathcal{E}}_{B'}$ by the construction of B' in (73). (79) follows from (69) with $\rho = \rho^*_{\lambda/2,u^*}$ and lastly (80) follows from Corollary 3.1 (ii).

It follows that the result in part (a) also holds in the second scenario in (72) which completes the proof for this part.

Part (b). Now, let E_1 denote the event that, for all $\rho \in \mathcal{P}_\pi(\Theta)$ simultaneously it holds that

$$\int_\Theta R(\theta) d\rho(\theta) \leq \int_\Theta R_n(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right]. \tag{81}$$

Let E_2 denote the event that

$$\begin{aligned}
&\int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + u \int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\hat{\rho}_{\lambda,u}, \pi\right) \\
&\leq \int_\Theta R(\theta) d\rho^*_{\lambda,u}(\theta) + u \int_\Theta K(\theta) d\rho^*_{\lambda,u}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,u}, \pi) + \sqrt{\frac{(M_y + uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}},
\end{aligned} \tag{82}$$

and let E_3 denote the event that

$$\begin{aligned}
&\int_\Theta K(\theta) d\hat{\rho}_{\lambda,u}(\theta) - \int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) \\
&\leq \frac{\sqrt{2}\,(M_y + uM_c)}{\kappa\sqrt{n}} \sqrt{\log\left(2\sqrt{n}\right) + \log\frac{4}{\epsilon}} + \frac{\lambda\,(M_y + uM_c)^2}{2n\kappa^2} + \frac{1}{\lambda}\left[\frac{\lambda^2 M_c^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right] \\
&= U_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_c^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right]
\end{aligned} \tag{83}$$

By Theorem 4.1 (a), applied with $s = -1$, $V_n(\theta) = R_n(\theta)$, $V(\theta) = R(\theta)$, and $M_\ell = M_y$, we have that $P^n(E_1) = \epsilon/4$. By Lemma A.7, $P^n(E_2) = \epsilon/4$. And lastly, by Theorem 4.1 (c) with $V_n(\theta) = K_n(\theta)$, $V(\theta) = K(\theta)$, and $M_\ell = M_c$, it holds that $P^n(E_3) = \epsilon/2$. Again applying a union bound argument similar to that in the proof of Theorem 4.1 (b), we have

$$P^n\left(E_1 \cap E_2 \cap E_3\right) \geq 1 - \epsilon.$$

As in part (a), we prove the result by showing that the intersection of these events implies the event in the result.

Recall,

$$B\left(\hat{\rho}_{\lambda,u}\right) = \int_\Theta K(\theta) d\hat{\rho}_{\lambda,u}(\theta) \text{ and } \widehat{B}\left(\hat{\rho}_{\lambda,u}\right) = \int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u}(\theta).$$

Then, the event E_3 described in (83) can be stated

$$B\left(\hat{\rho}_{\lambda,u}\right) - \widehat{B}\left(\hat{\rho}_{\lambda,u}\right) \leq U_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_c^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right]. \tag{84}$$

Now, starting from (81) with $\rho = \hat{\rho}_{\lambda,u}$,

$$
\begin{aligned}
&\int_\Theta R(\theta) d\hat{\rho}_{\lambda,u}(\theta) \\
&\leq \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right] \\
&= \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + u\left(\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u} - \widehat{B}\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \pi) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right] \\
&= \int_\Theta R_n(\theta) d\hat{\rho}_{\lambda,u}(\theta) + u\int_\Theta K_n(\theta) d\hat{\rho}_{\lambda,u} + \frac{1}{\lambda} D_{\mathrm{KL}}(\hat{\rho}_{\lambda,u}, \pi) \\
&\qquad - uB\left(\hat{\rho}_{\lambda,u}\right) + u\left(B\left(\hat{\rho}_{\lambda,u}\right) - \widehat{B}\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right] \\
&\leq \int_\Theta R(\theta) d\rho^*_{\lambda,u}(\theta) + u\int_\Theta K(\theta) d\rho^*_{\lambda,u}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,u}, \pi) + \sqrt{\frac{(M_y + uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} \\
&\qquad - uB\left(\hat{\rho}_{\lambda,u}\right) + u\left(B\left(\hat{\rho}_{\lambda,u}\right) - \widehat{B}\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right] \qquad (85) \\
&= \int_\Theta R(\theta) d\rho^*_{\lambda,u}(\theta) + u\left(\int_\Theta K(\theta) d\rho^*_{\lambda,u}(\theta) - B\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,u}, \pi) + \sqrt{\frac{(M_y + uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} \\
&\qquad + u\left(B\left(\hat{\rho}_{\lambda,u}\right) - \widehat{B}\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{4}{\epsilon}\right] \\
&\leq \int_\Theta R(\theta) d\rho^*_{\lambda,u}(\theta) + u\left(\int_\Theta K(\theta) d\rho^*_{\lambda,u}(\theta) - B\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,u}, \pi) + \sqrt{\frac{(M_y + uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} \\
&\qquad + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2 + uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right] \qquad (86) \\
&\leq \sup_{a\geq 0}\left[\int_\Theta R(\theta) d\rho^*_{\lambda,a}(\theta) + a\left(\int_\Theta K(\theta) d\rho^*_{\lambda,a}(\theta) - B\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,a}, \pi)\right] \\
&\qquad + \sqrt{\frac{(M_y + uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2 + uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right] \\
&= \min_{\rho\in\mathcal{E}_{B(\hat{\rho}_{\lambda,u})}}\left\{\int_\Theta R(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho, \pi)\right. \\
&\qquad\qquad \left. + \sqrt{\frac{(M_y + uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2 + uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right]\right\}. \qquad (87)
\end{aligned}
$$

In the above, (85) follows from plugging in (82), (86) follows from plugging in (84), and lastly (87) follows from Corollary 3.2. Switching the notation to $\int_\Theta R(\theta) d\rho(\theta) = R(f_{G,\rho})$ for $\rho \in \mathcal{P}(\Theta)$ and utilizing the definition of $U_2\left(\epsilon; \lambda, u, n\right)$, the above yields the statement in part (b) of the Theorem. ■

A.3.2 Proofs for Subsection 4.2

Proof of Theorem 4.3. We will use the following properties in the proofs of part (a) and (b). For treatment assignment rules of the form in (10), when $\|\theta\| \neq 0$, it holds that $f_\theta(x) = f_{\theta/\|\theta\|}(x)$ for all $x \in \mathcal{X}$. As we are presuming that $\overline{\theta} \neq 0$ and $\overline{\theta}_u \neq 0$ (almost surely), with probability one we can find a values $\overline{\theta}$ and $\overline{\theta}_u$ such that $\|\overline{\theta}\| = 1$ and $\|\overline{\theta}_u\| = 1$. We assume $\overline{\theta}$ and $\overline{\theta}_u$ are selected to have this property for the remainder of the proof. Below, for integration over $\Theta = \mathbb{R}^q$, we write $\int \ldots$ in place of $\int_{\mathbb{R}^q} \ldots$

Observe that for $\theta, \theta_1 \in \mathbb{R}^q$ such that $\|\theta_1\| = 1$ and $\|\theta\| \neq 0$,

$$
\begin{aligned}
R(\theta) - R(\theta_1) &= W(f_\theta) - W(f_{\theta_1}) && (88)\\
&= E_Q\left[(Y_1 - Y_0)\left(f_\theta(X) - f_{\theta_1}(X)\right)\right]\\
&\leq M_y E_P\left[\left|1\left\{\phi(X)^\intercal\theta > 0\right\} - 1\left\{\phi(X)^\intercal\theta_1 > 0\right\}\right|\right] && (89)\\
&= M_y P\left[\left(\phi(X)^\intercal\theta\right)\left(\phi(X)^\intercal\theta_1\right) < 0\right]\\
&= M_y P\left[\left(\phi(X)^\intercal\frac{\theta}{\|\theta\|}\right)\left(\phi(X)^\intercal\theta_1\right) < 0\right]\\
&\leq M_y \nu \left\|\frac{\theta}{\|\theta\|} - \theta_1\right\| && (90)\\
&\leq M_y 2\nu \left\|\theta - \theta_1\right\|, && (91)
\end{aligned}
$$

where (88) follows from the definition of welfare regret, (89) follows from Assumption 3.1 (iii) and the fact that the distribution of X is determined by P as well as Q, (90) follows from Assumption 4.3, and (91) follows from the fact that with θ, θ_1 as above,

$$\left\|\frac{\theta}{\|\theta\|} - \theta_1\right\| \leq \|\theta - \theta_1\|.$$

As a consequence of (91), for any $\sigma > 0$,

$$
\begin{aligned}
\int R(\theta) d\Phi_{\theta_1,\sigma^2}(\theta) &= R(\theta_1) + \int \left[R(\theta) - R(\theta_1)\right] d\Phi_{\theta_1,\sigma^2}(\theta)\\
&\leq R(\theta_1) + 2M_y\nu \int \|\theta - \theta_1\| \, d\Phi_{\theta_1,\sigma^2}(\theta)\\
&\leq R(\theta_1) + 2M_y\nu\sigma\sqrt{q}, && (92)
\end{aligned}
$$

where we have used the fact that for $\theta \sim \Phi_{\theta_1,\sigma^2}$, $||\theta - \theta_1|| \sim \sigma H^{1/2}$ with $H \sim \chi^2(q)$. Then, by Jensen's inequality, $E\sigma H^{1/2} \leq \sigma(EH)^{1/2} = \sigma(q)^{1/2}$.

Following nearly identical steps, now starting with the definition of the expected costs $K(\theta)$ and $K(\overline{\theta})$, it is straightforward to derive that, for $\theta, \theta_1 \in \mathbb{R}^q$ such that $\|\theta_1\| = 1$ and $\|\theta\| \neq 0$,

$$K(\theta) - K(\theta_1) \leq M_c 2\nu \|\theta - \theta_1\|,$$

and for $\sigma > 0$,

$$\int K(\theta) d\Phi_{\theta_1,\sigma^2}(\theta) \leq K(\theta_1) + 2M_c\nu\sigma\sqrt{q}. \tag{93}$$

Lastly, before considering part (a) and (b) separately, note that by Lemma A.5, with $\sigma_\pi = 1/\sqrt{q}$, $\sigma_\rho = 1/(2\sqrt{nq})$, and $\|\theta_1\| = 1$,

$$D_{\mathrm{KL}}\left(\Phi_{\theta_1,\sigma_\rho^2}, \Phi_{0,\sigma_\pi^2}\right) = \frac{q}{2}\left[\frac{1}{4n} + \log(4n)\right]. \tag{94}$$

Part (a). We consider the posterior distribution $\widetilde{\rho} = \Phi_{\overline{\theta},\sigma_\rho^2}$ with $\sigma_\rho = 1/(2\sqrt{nq})$ so that $D_{\mathrm{KL}}(\widetilde{\rho},\pi)$ is given by (94). Next, define

$$B' = B + \frac{\nu M_c}{\sqrt{n}}. \tag{95}$$

Assumptions 3.2 and 3.3 are met and Assumptions 4.3 and 3.4 are assumed to hold so we can apply Theorem 4.2 (a). Starting from there, with probability at least $1-\epsilon$ we have

$$\begin{aligned}
&\int R(\theta)d\hat{\rho}_{\lambda,\hat{u}}(\theta) \\
&\le \min_{\rho\in\mathcal{E}_B}\left\{\int R(\theta)d\rho(\theta) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho,\pi)\right\} + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} \\
&= \int R(\theta)d\rho^*_{\lambda/2,u^*}(\theta) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho^*_{\lambda/2,u^*},\pi) + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} \qquad (96) \\
&= \int R(\theta)d\rho^*_{\lambda/2,u^*}(\theta) + u^*\left(\int K(\theta)d\rho^*_{\lambda/2,u^*}(\theta) - B\right) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho^*_{\lambda/2,u^*},\pi) \\
&\qquad + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} \qquad (97) \\
&= \int R(\theta)d\rho^*_{\lambda/2,u^*}(\theta) + u^*\left(\int K(\theta)d\rho^*_{\lambda/2,u^*}(\theta) - B'\right) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho^*_{\lambda/2,u^*},\pi) \\
&\qquad + u^*\left(B'-B\right) + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} \\
&\le \sup_{u\ge 0}\left[\int R(\theta)d\rho^*_{\lambda/2,u} + u\left(\int K(\theta)d\rho^*_{\lambda,u}(\theta) - B'\right) + \frac{2}{\lambda}D_{\mathrm{KL}}(\rho^*_{\lambda,u},\pi)\right] \\
&\qquad + u^*\left(B'-B\right) + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} \\
&= \min_{\rho\in\mathcal{E}_{B'}}\left\{\int R(\theta)d\rho(\theta) + \frac{2}{\lambda}D_{\mathrm{KL}}\left(\rho,\pi\right)\right\} + u^*\frac{\nu M_c}{\sqrt{n}} + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} \qquad (98)
\end{aligned}$$

In the above, (96) and (97) follow from Corollary 3.1 (ii) while (98) follows from applying Corollary 3.1 (ii) and the definition of B' in (95).

From (93) with $\overline{\theta}$ in the place of θ_1 and with $\sigma_\rho = 1/(2\sqrt{nq})$, we have

$$\int K(\theta)d\widetilde{\rho}(\theta) = \int K(\theta)d\Phi_{\overline{\theta},\sigma_\rho^2}(\theta) \le K\left(\overline{\theta}\right) + \frac{\nu M_c}{\sqrt{n}} \le B' \tag{99}$$

as, by the definition of $\overline{\theta}$, $K(\overline{\theta}) \le B$. Therefore $\widetilde{\rho}\in\mathcal{E}_{B'}$. From (98), we have, with probability at least $1-\epsilon$,

$$\begin{aligned}
&\int R(\theta) d\hat{\rho}_{\lambda,\hat{u}}(\theta) \\
&\leq \inf_{\rho\in\mathcal{E}_{B'}} \left\{\int R(\theta)d\rho(\theta) + \frac{2}{\lambda}D_{\mathrm{KL}}\left(\rho,\pi\right)\right\} + u^*\frac{\nu M_c}{\sqrt{n}} + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} \\
&\leq \int R(\theta)d\widetilde{\rho}(\theta) + \frac{2}{\lambda}D_{\mathrm{KL}}\left(\widetilde{\rho},\pi\right) + u^*\frac{\nu M_c}{\sqrt{n}} + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}} \\
&\leq R\left(\overline{\theta}\right) + \frac{\nu M_y}{\sqrt{n}} + \frac{q}{\lambda}\left[\frac{1}{4n} + \log\left(4n\right)\right] + u^*\frac{\nu M_c}{\sqrt{n}} + \frac{2}{\lambda}\left[\frac{\lambda^2 M_y^2}{8n\kappa^2} + \log\frac{3}{\epsilon}\right] + \hat{u}\sqrt{\frac{M_c^2\log\frac{3}{\epsilon}}{2n\kappa^2}}.
\end{aligned}$$

In the last step, we have applied the properties in (92) and (94) with $\overline{\theta}$ taking the role of θ_1. Plugging in $\lambda = \kappa\sqrt{nq}/M_y$ and rearranging terms then produces the result in (a) with

$$\overline{U}_1(n;q) = \sqrt{\frac{q}{n}}\left[\frac{\nu M_y}{\sqrt{q}} + \frac{M_y}{\kappa}\left(\frac{1}{4} + \frac{1}{4n}\right)\right].$$

Part (b). As a starting point, we utilize the setup and initial steps of the proof of Theorem 4.2 (b). Assume the same the definitions of events E_1, E_2 and E_3 as in (81), (82), (83), respectively. Following that proof up to (86), we have that with probability at least $1-\epsilon$,

$$\begin{aligned}
&\int R(\theta)d\hat{\rho}_{\lambda,u}(\theta) \\
&\leq \int R(\theta)d\rho^*_{\lambda,u}(\theta) + u\left(\int K(\theta)d\rho^*_{\lambda,u}(\theta) - B\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda}D_{\mathrm{KL}}(\rho^*_{\lambda,u},\pi) + \sqrt{\frac{(M_y+uM_c)^2\log(4/\epsilon)}{2n\kappa^2}} \\
&\quad + uU_1\left(\epsilon;\lambda,u,n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2+uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right], \qquad (100)
\end{aligned}$$

where $B(\hat{\rho}_{\lambda,u}) = \int K(\theta)d\hat{\rho}_{\lambda,u} = K(f_{G,\hat{\rho}_{\lambda,u}})$ and $U_1\left(\epsilon;\lambda,u,n\right)$ is defined in Theorem 4.2 (b).

Now we will consider the posterior $\widetilde{\rho} = \Phi_{\overline{\theta}_u,\sigma_\rho^2}$ with $\sigma_\rho = 1/(2\sqrt{nq})$. Utilizing (94) with π as described in the theorem, now with $\overline{\theta}_u$ in place of θ_1, with probability one we have

$$D_{\mathrm{KL}}(\widetilde{\rho},\pi) = \frac{q}{2}\left[\frac{1}{4n} + \log\left(4n\right)\right]. \qquad (101)$$

Additionally, we now define

$$B' = B\left(\hat{\rho}_{\lambda,u}\right) + \frac{\nu M_c}{\sqrt{n}}. \qquad (102)$$

From (93) with $\overline{\theta}_u$ in the place of θ_1 and with $\sigma_\rho = 1/(2\sqrt{nq})$, with probability one we have

$$\int K(\theta)d\widetilde{\rho}(\theta) = \int K(\theta)d\Phi_{\overline{\theta}_u,\sigma_\rho^2}(\theta) \leq K\left(\overline{\theta}_u\right) + \frac{\nu M_c}{\sqrt{n}} \leq B', \qquad (103)$$

because, by the definition of $\overline{\theta}_u$ we have $K(\overline{\theta}_u) \leq B\left(\hat{\rho}_{\lambda,u}\right)$ (a.s.). It follows that with probability one, $\widetilde{\rho} \in \mathcal{E}_{B'}$.

Returning to (100), we have, with probability at least $1-\epsilon$,

$$
\begin{aligned}
&\int R(\theta) d\hat{\rho}_{\lambda,u}(\theta) \\
&\leq \int R(\theta) d\rho^*_{\lambda,u}(\theta) + u\left(\int K(\theta) d\rho^*_{\lambda,u}(\theta) - B\left(\hat{\rho}_{\lambda,u}\right)\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,u}, \pi) + \sqrt{\frac{(M_y+uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} \\
&\quad + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2+uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right] \\
&\leq \int R(\theta) d\rho^*_{\lambda,u}(\theta) + u\left(\int K(\theta) d\rho^*_{\lambda,u}(\theta) - B'\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,u}, \pi) + \sqrt{\frac{(M_y+uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} \\
&\quad + u\left(B' - B\left(\hat{\rho}_{\lambda,u}\right)\right) + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2+uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right] \\
&\leq \sup_{a\geq 0}\left[\int R(\theta) d\rho^*_{\lambda,a}(\theta) + u\left(\int K(\theta) d\rho^*_{\lambda,a}(\theta) - B'\right) + \frac{1}{\lambda} D_{\mathrm{KL}}(\rho^*_{\lambda,a}, \pi)\right] \\
&\quad + \sqrt{\frac{(M_y+uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} + u\left(\frac{\nu M_c}{\sqrt{n}}\right) + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2+uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right] \qquad (104) \\
&= \inf_{\rho\in\mathcal{E}_{B'}}\left[\int R(\theta) d\rho(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\rho, \pi\right)\right] + \sqrt{\frac{(M_y+uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} \\
&\quad + u\left(\frac{\nu M_c}{\sqrt{n}}\right) + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2+uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right] \qquad (105) \\
&\leq \int R(\theta) d\widetilde{\rho}(\theta) + \frac{1}{\lambda} D_{\mathrm{KL}}\left(\widetilde{\rho}, \pi\right) + \sqrt{\frac{(M_y+uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} \\
&\quad + u\left(\frac{\nu M_c}{\sqrt{n}}\right) + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2+uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right] \qquad (106) \\
&\leq R\left(\overline{\theta}_u\right) + \frac{\nu M_y}{\sqrt{n}} + \frac{q}{2\lambda}\left[\frac{1}{4n} + \log\left(4n\right)\right] + \sqrt{\frac{(M_y+uM_c)^2 \log(4/\epsilon)}{2n\kappa^2}} \\
&\quad + u\left(\frac{\nu M_c}{\sqrt{n}}\right) + uU_1\left(\epsilon; \lambda, u, n\right) + \frac{1}{\lambda}\left[\frac{\lambda^2\left(M_y^2+uM_c^2\right)}{8n\kappa^2} + (1+u)\log\frac{4}{\epsilon}\right] \qquad (107)
\end{aligned}
$$

Above, (104) follows from (102) and the fact that the supremum there is greater than or equal to the object it replaces, (105) follows from Corollary 3.2 (ii), (106) follows from having $\widetilde{\rho} \in \mathcal{E}_{B'}$ with probability one, and lastly (107) follows from (92), with $\overline{\theta}_u$ in place of θ_1 and $\sigma_\rho = 1/(2\sqrt{nq})$ in place of σ, and utilizing (101). Plugging in λ as given in part (b), straightforward manipulations of the expression in (107) show that the inequality can be written

$$
R\left(f_{G,\hat{\rho}_{\lambda,u}}\right) \leq R\left(\overline{\theta}_u\right) + \frac{M_y+uM_c}{\kappa}\left[\overline{U}_2(n;q,u,\epsilon) + \overline{U}_3(n;q,u,\epsilon) + \overline{U}_4(n;q,u)\right],
$$

where

$$
\overline{U}_2(n;q,u,\epsilon) = \frac{\sqrt{q}\log\left(2\sqrt{n}\right) + \sqrt{2}u\sqrt{\log\left(2\sqrt{n}\right) + \log\frac{4}{\epsilon}}}{\sqrt{n}} = \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right),
$$

$$\overline{U}_3(n;q,u,\epsilon) = \frac{\sqrt{\frac{\log(4/\epsilon)}{2}} + \frac{1}{\sqrt{q}}(1+u)\log\frac{4}{\epsilon}}{\sqrt{n}} = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right),$$

and

$$\overline{U}_4(n;q,u) = \frac{\kappa\nu + \sqrt{q}\left(\frac{1}{8n} + \frac{u}{2}\right)}{\sqrt{n}} + \sqrt{\frac{q}{n}}\left(\frac{M_y^2 + uM_c^2}{8\left(M_y + uM_c\right)^2}\right) = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

■

A.3.3 Proofs for Subsection 4.3

Proof of Theorem 4.4. First, note that

$$f_{\mathrm{mv},\rho}(x) = 1\left\{\int_\Theta f_\theta(x)d\rho(\theta) > \frac{1}{2}\right\} \leq 2\int_\Theta f_\theta(x)d\rho(\theta) = 2f_{G,\rho}(x). \tag{108}$$

To see this, note that when $\int_\Theta f(x)d\rho(\theta) \leq 1/2$, $f_{\mathrm{mv},\rho} = 0$ hence the left-hand size of the above inequality is zero while the right-hand size is non-negative and the inequality holds. When $\int_\Theta f_\theta(x)d\rho(\theta) > 1/2$, the left hand side is 1 while the right hand side must be greater than 1, so the inequality holds in all cases.

Next we will show that for any $x \in \mathcal{X}$,

$$\left(\delta_y(x) - \eta_{B(\rho)}\delta_c(x)\right)\left(f^*_{B(\rho)}(x) - f_{\mathrm{mv},\rho}(x)\right) \leq 2\left(\delta_y(x) - \eta_{B(\rho)}\delta_c(x)\right)\left(f^*_{B(\rho)}(x) - f_{G,\rho}(x)\right). \tag{109}$$

To see this, first consider $x \in \mathcal{X}$ such that $f^*_{B(\rho)}(x) = 1\{\delta_y(x) - \eta_{B(\rho)}\delta_c(x) > 0\} = 0$. In this case, $\delta_y(x) - \eta_{B(\rho)}\delta_c(x) \leq 0$ and we have

$$\begin{aligned}
\left(\delta_y(x) - \eta_{B(\rho)}\delta_c(x)\right)\left(f^*_{B(\rho)}(x) - f_{\mathrm{mv},\rho}(x)\right) &= \left|\delta_y(x) - \eta_{B(\rho)}\delta_c(x)\right| f_{\mathrm{mv},\rho}(x) \\
&\leq 2\left|\delta_y(x) - \eta_{B(\rho)}\delta_c(x)\right| f_{G,\rho}(x) \\
&= 2\left(\delta_y(x) - \eta_{B(\rho)}\delta_c(x)\right)\left(f^*_{B(\rho)}(x) - f_{G,\rho}(x)\right),
\end{aligned}$$

where the inequality follows from (108). To verify (109), we now need to check that it holds for $x \in \mathcal{X}$ such that $f^*_{B(\rho)}(x) = 1\{\delta_y(x) - \eta_{B(\rho)}\delta_c(x) > 0\} = 1$. In this case, $\delta_y(x) - \eta_{B(\rho)}\delta_c(x) > 0$, so (109) reduces to

$$\left(f^*_{B(\rho)}(x) - f_{\mathrm{mv},\rho}(x)\right) \leq 2\left(f^*_{B(\rho)}(x) - f_{G,\rho}(x)\right). \tag{110}$$

First consider $x \in \mathcal{X}$ such that $f_{\mathrm{mv},\rho}(x) = 1$. Then the left-hand size is zero while the right hand side is non-negative as $f_{G,\rho} \in [0,1]$ and $f^*_{B(\rho)(x)} = 1$ in the current assumed setting, so the condition holds. Lastly, if $f_{\mathrm{mv},\rho}(x) = 0$, so that $\int_\Theta f_\theta(x)d\rho(\theta) \leq 1/2$, in the current setting with $f^*_{B(\rho)}(x) = 1$ we then have

$$\begin{aligned}
2\left(f^*_{B(\rho)}(x) - f_{G,\rho}(x)\right) &= 2\left(1 - \int_\Theta f_\theta(x)d\rho(\theta)\right) \\
&\geq 2\left(1 - \frac{1}{2}\right) \\
&= 1 \\
&= f^*_{B(\rho)}(x) - f_{\mathrm{mv},\rho}(x).
\end{aligned}$$

Hence (109) holds for all $x \in \mathcal{X}$. Taking the expectation of both sides of that inequality with respect to a draw of X from Q then yields that

$$L_{B(\rho)}\left(f_{\mathrm{mv},\rho}\right) \leq 2L_{B(\rho)}\left(f_{G,\rho}\right). \tag{111}$$

To complete the proof, we need to verify that

$$L_{B(\rho)}\left(f_{G,\rho}\right) = R_{B(\rho)}\left(f_{G,\rho}\right). \tag{112}$$

Now there are two possibilities to consider. The first is when $\eta_{B(\rho)} = 0$. In this case, we have

$$\begin{aligned} L_{B(\rho)}\left(f_{G,\rho}\right) &= E_Q\left[\delta_y(X)\left(f^*_{B(\rho)} - f_{G,\rho}\right)\right] \\ &= W\left(f^*_{B(\rho)}\right) - W\left(f_{G,\rho}\right) = R_{B(\rho)}\left(f_{G,\rho}\right). \end{aligned}$$

And in the only remaining case, when $\eta_{B(\rho)} > 0$, we also have $K(f^*_{B(\rho)}) = B(\rho)$ by Theorem 3.1. As, by the definition of $B(\rho)$, it also holds that $K(f_{G,\rho}) = B(\rho)$, we have

$$\begin{aligned} L_{B(\rho)}\left(f_{G,\rho}\right) &= E_Q\left[\delta_y(X)\left(f^*_{B(\rho)} - f_{G,\rho}\right)\right] - \eta_{B(\rho)} E_Q\left[\delta_c(X)\left(f^*_{B(\rho)} - f_{G,\rho}\right)\right] \\ &= E_Q\left[\delta_y(X)\left(f^*_{B(\rho)} - f_{G,\rho}\right)\right] - \eta_{B(\rho)}\left[K\left(f^*_{B(\rho)}\right) - K\left(f_{G,\rho}\right)\right] \\ &= E_Q\left[\delta_y(X)\left(f^*_{B(\rho)} - f_{G,\rho}\right)\right] \\ &= R_{B(\rho)}\left(f_{G,\rho}\right). \end{aligned}$$

Hence (112) holds and combined with (111) this completes the proof. ∎

References

Alquier, P., Ridgway, J., and Chopin, N. (2016). On the properties of variational approximations of gibbs posteriors. *The Journal of Machine Learning Research*, 17(1):8374–8414.

Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

Athey, S. and Wager, S. (2021). Policy learning with observational data. *Econometrica*, 89(1):133–161.

Beygelzimer, A. and Langford, J. (2009). The offset tree for learning with partial labels. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 129–138.

Bhattacharya, D. and Dupas, P. (2012). Inferring welfare maximizing treatment assignment under budget constraints. *Journal of Econometrics*, 167(1):168–196.

Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration inequalities: A nonasymptotic theory of independence*. Oxford University Press.

Catoni, O. (2004). *Statistical Learning Theory and Stochastic Optimization: Ecole d'Eté de Probabilités de Saint-Flour, XXXI-2001*, volume 1851. Springer Science & Business Media.

Catoni, O. (2007). Pac-bayesian supervised classification: The thermodynamics of statistical learning. *arXiv preprint arXiv:0712.0248*.

Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience.

Csiszár, I. (1975). I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, pages 146–158.

Dehejia, R. H. (2005). Program evaluation as a decision problem. *Journal of Econometrics*, 125(1-2):141–173.

Del Moral, P., Doucet, A., and Jasra, A. (2006). Sequential monte carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):411–436.

Donsker, M. D. and Varadhan, S. R. S. (1975). Asymptotic evaluation of certain markov process expectations for large time. *Communications on Pure and Applied Mathematics*, 28.

Elliott, G. and Lieli, R. P. (2013). Predicting binary outcomes. *Journal of Econometrics*, 174(1):15–26.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Group, O. H. S. (2012). The oregon health insurance experiment: Evidence from the first year. *The Quarterly Journal of Economics*, 127(3):1057–1106.

Freund, Y., Mansour, Y., and Schapire, R. E. (2004). Generalization Bounds for Averaged Classifiers. *The Annals of Statistics*, 32(4):1698 – 1722.

Germain, P., Lacasse, A., Laviolette, F., March, M., and Roy, J.-F. (2015). Risk bounds for the majority vote: From a pac-bayesian analysis to a learning algorithm. *Journal of Machine Learning Research*, 16(26):787–860.

Guedj, B. and Robbiano, S. (2018). Pac-bayesian high dimensional bipartite ranking. *Journal of Statistical Planning and Inference*, 196:70–86.

Hendren, N. and Sprung-Keyser, B. (2020). A unified welfare analysis of government policies. *The Quarterly Journal of Economics*, 135(3):1209–1318.

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

Huang, X. and Xu, J. (2020). Estimating individualized treatment rules with risk constraint. *Biometrics*, 76(4):1310–1318.

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

Kuecken, M., Thuilliez, J., and Valfort, M.-A. (2014). Does malaria control impact education? a study of the global fund in africa.

Langford, J. and Shawe-Taylor, J. (2003). Pac-bayes & margins. In *Advances in Neural Information Processing Systems 15: Proceedings of the 2002 Conference*, volume 15, page 439. MIT Press.

Lengeler, C. (1998). Insecticide treated bednets and curtains for malaria control. *Cochrane database of systematic reviews*, (2).

Lever, G., Laviolette, F., and Shawe-Taylor, J. (2010). Distribution-dependent pac-bayes priors. In *International Conference on Algorithmic Learning Theory*, pages 119–133. Springer.

London, B. and Sandler, T. (2019). Bayesian counterfactual risk minimization. In *International Conference on Machine Learning*, pages 4125–4133. PMLR.

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

Massart, P. (2007). *Concentration Inequalities and Model Selection: Ecole d'Eté de Probabilités de Saint-Flour XXXIII-2003*. Springer.

Maurer, A. (2004). A note on the pac bayesian theorem. *arXiv preprint cs/0411099*.

Mbakop, E. and Tabord-Meehan, M. (2021). Model selection for treatment choice: Penalized welfare maximization. *Econometrica*, 89(2):825–848.

McAllester, D. (2003a). Simplified pac-bayesian margin bounds. In *Learning theory and Kernel machines*, pages 203–215. Springer.

McAllester, D. A. (1999a). Pac-bayesian model averaging. In *Proceedings of the twelfth annual conference on Computational learning theory*, pages 164–170.

McAllester, D. A. (1999b). Some pac-bayesian theorems. *Machine Learning*, 37(3):355–363.

McAllester, D. A. (2003b). Pac-bayesian stochastic model selection. *Machine Learning*, 51(1):5–21.

McDiarmid, C. (1989). On the method of bounded differences. *Surveys in Combinatorics*, 141(1):148–188.

Qian, M. and Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *Annals of statistics*, 39(2):1180.

Ridgway, J., Alquier, P., Chopin, N., and Liang, F. (2014). Pac-bayesian auc classification and scoring. *Advances in Neural Information Processing Systems*, 27.

Seeger, M. (2002). Pac-bayesian generalisation error bounds for gaussian process classification. *Journal of Machine Learning Research*, 3(Oct):233–269.

Shawe-Taylor, J. and Williamson, R. C. (1997). A pac analysis of a bayesian estimator. In *Proceedings of the tenth annual conference on Computational learning theory*, pages 2–9.

Sun, H., Du, S., and Wager, S. (2021). Treatment allocation under uncertain costs. *arXiv preprint arXiv:2103.11066*.

Sun, L. (2021). Empirical welfare maximization with constraints. *arXiv preprint arXiv:2103.15298*.

Teklehaimanot, A., McCord, G. C., and Sachs, J. D. (2007). Scaling up malaria control in africa: An economic and epidemiological assessment. *The American Journal of Tropical Medicine and Hygiene*, 77(6 Suppl):138 – 144.

Thiemann, N., Igel, C., Wintenberger, O., and Seldin, Y. (2017). A strongly quasiconvex pac-bayesian bound. In Hanneke, S. and Reyzin, L., editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 466–492. PMLR.

Viviano, D. (2019). Policy targeting under network interference. *arXiv preprint arXiv:1906.10258*.

Wang, Y., Fu, H., and Zeng, D. (2018). Learning optimal personalized treatment rules in consideration of benefit and risk: with an application to treating type 2 diabetes patients with insulin therapies. *Journal of the American Statistical Association*, 113(521):1–13.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499):1106–1118.