Functional Sequential Treatment Allocation

Anders Bredahl Kock
University of Oxford
CREATES, Aarhus University
anders.kock@economics.ox.ac.uk

David Preinerstorfer
ECARES, SBS-EM
Université libre de Bruxelles
david.preinerstorfer@ulb.ac.be

Bezirgen Veliyev
CREATES, Aarhus University
bveliyev@econ.au.dk

December 27, 2018

Abstract

In this paper we study a treatment allocation problem with multiple treatments, in which the individuals to be treated arrive sequentially. The goal of the policy maker is to treat every individual as well as possible. Which treatment is "best" is allowed to depend on various characteristics (functionals) of the individual-specific outcome distribution of each treatment. For example measures of welfare, inequality, or poverty. We propose the Functional Sequential Allocation policy, and provide upper bounds on the regret it incurs compared to the oracle policy that *knows* the best treatment for each individual. These upper bounds increase sublinearly in the number of treatment assignments and we show that this regret guarantee is minimax optimal. In addition, the assumptions under which this regret guarantee is established are as weak as possible — even a minimal weakening of them will imply non-existence of policies with regret increasing sub-linearly in the number of assignments. Furthermore, we provide an upper bound on the number of suboptimal assignments made by the FSA policy and show that every policy must make a number of suboptimal assignments at least of this order.

JEL Classification: C18, C22, J68.

Keywords: Sequential Treatment Allocation, Distributional Policy Effects, Statistical Decision Theory, Minimax Optimal Regret, Multiple Treatments.

1 Introduction

In many treatment programs the individuals to be treated arrive sequentially. For example, workers become unemployed throughout the year or patients to be treated in a hospital fall at different points in time. Therefore, one must *sequentially* treat the individuals and thus sequentially learn about the treatment-specific outcome distribution as treatment outcomes are observed. In this paper, we study a setting in which a policy maker's objective is to treat as many individuals as well as possible in the course of the treatment period. We measure the attractiveness of a treatment by a (combination of) functionals of the distribution of treatment outcomes. Thus, the goal of the policy maker can be, *for example*, to assign as often as possible the treatment with the highest welfare (according to some welfare measure) or quantile, lowest uncertainty of outcome (according to a measure of dispersion), lowest inequality or poverty, or best tradeoff between several distributional characteristics of the treatments. In order to treat as many individuals as well as possible the policy maker must sequentially learn the distribution of treatment outcomes for the K available treatments, yet assign as few as possible individuals to suboptimal treatments.

We propose the Functional Sequential Allocation (FSA) policy for the above problems and establish its properties. We first analyze a homogeneous treatment distribution setting, i.e.,the outcome distribution for each of the treatments is assumed to be the same for all individuals. To begin, we provide an upper bound on the maximal expected *regret* of the FSA policy compared to what could have been obtained had the population outcome distributions of the K treatments been known from the outset and one had always assigned the treatment maximizing the (combination of) functional characteristics of interest, cf. Theorem 3.1. Next, we show that FSA policy is near minimax optimal. More precisely, we show in Theorem 3.6 that every policy must incur of at least almost the same order as FSA policy.

In the setting of heterogeneous treatment effects we assume that the policy maker observes a vector of characteristics (covariates) of the individual to be treated prior to treatment allocation. For each individual the best is now the one that maximizes the (combination of) functionals of the conditional distribution given the vector of characteristics. In Theorem 4.4 (cf. also Corollary 4.5) we show how the FSA policy can be adjusted in such a way that its maximal regret vis-à-vis the infeasible oracle policy (the policy that *knows* the population conditional distributions of treatment outcomes given characteristics and always assigns the best one) can be bounded from above and shown to increase sublinearly in the number of assignments. Theorem 5.2 shows that the assumptions under which the upper bounds on regret are established are essentially minimal as every policy must incur a regret which is linear in the number of assignments without such assumptions. Furthermore, Theorem 5.3 establishes the minimax optimality of the FSA policy up to logarithmic factors.

Next, if for a large proportion of vectors of characteristics the best and second best treatment are not too similar, we show that the upper bound on the maximal regret compared to the infeasible oracle can be further sharpened, cf. Theorem 4.7. Furthermore, we show in Theorem 4.8 that the expected number of suboptimal assignments made increases sublinearly in the number of assignments. The latter can be interpreted as an ethical guarantee on the FSA policy; only few persons will receive a treatment which is not optimal for them. Furthermore, Theorem 5.1 (and its proof) shows that i) no policy can achieve much lower regret than the FSA policy, ii) every policy must make at least as many suboptimal assignments as the FSA policy. In this sense, the FSA policy is (minimax) optimal.

Finally, in Theorem 5.4, we show that even though our sublinear regret bounds are increasing in the dimension d of the vector of characteristics, one should never ignore these, as maximal regret of any policy must increase even linearly in the number of assignments made if one ignores the characteristics.

We note that the sequential setting considered in this paper differs from the classic treatment setting in which one often presupposes the existence of a data set that has already been sampled. Based on this data set, one is then interested in estimating the *effect* of a treatment. This effect is most often the difference in a suitable (conditional) expectation, i.e.,the focus is on the first moment of the distribution of treatment outcomes — a special case of the functionals studied in this paper. We also stress that the goal of the policy maker in this paper is to treat as many individuals as well as possible. This is not equivalent to assigning treatments in such a way that "information" about the treatments is maximized after the last treatment. Assigning treatments sequentially in a way that maximizes "information" at the end of the treatment period is also an interesting goal, which warrants further study. The latter objective, however, can sometimes be problematic as an algorithm designed for this purpose *knowingly* treats individuals in a suboptimal way in order to obtain information.

1.1 Related literature

Our paper is related to several strands of literature. First, our work relates to the literature on statistical treatment allocation rules. Here Manski (2004) did seminal work in proposing conditional empirical success rules which take a finite partition of the covariate space and on each set of this partition dictate to assign the treatment with the highest sample average. In choosing this partition, one faces a tradeoff between individualization and having enough observations for each group to estimate treatment effects precisely. In particular, Manski (2004) studies when full individualization is optimal. Stoye (2009) shed further light on this by showing that if one does not restrict how outcomes vary with covariates then full individualization is alway minimax optimal. This result relies on the fact that without any restrictions on how the outcome distribution varies with covariates, this relationship could be arbitrarily wiggly such that even seemingly similar individuals may carry no information about how treatments affect the other person.

Furthermore, our work is related to the recent paper by Kitagawa and Tetenov (2018) who consider treatment allocation through an *empirical welfare maximization* lens. The authors take the view that realistic policies are often constrained to be simple due to ethical, legislative, or political reasons. Using techniques from empirical risk minimization they show how their procedure is minimax optimal within the considered class of realistic policies. Furthermore, Athey and Wager (2017) have used concepts from semiparametric efficiency theory to establish regret bounds that scale with the semiparametrically efficient variance. Finally, Kitagawa and Tetenov (2017) have considered treatment allocation in a setting in which one targets "equality-minded" social welfare functions. Other papers on statistical treatment rules in econometrics focusing on the case where the sample is given include Chamberlain (2000); Dehejia (2005); Hirano and Porter (2009); Bhattacharya and Dupas (2012); Stoye (2012); Tetenov (2012). Furthermore, Rothe (2010) has done work on inference on policy effects that need not be restricted to the mean of the distribution, cf. also Rothe (2012).

The most important distinguishing features of our work compared to the classic literature on statistical treatment rules above is that we are working in a sequential setting where the individuals to be treated arrive gradually. Thus, we do not have a data set of size N at our disposal from the outset based on which the best treatment must be found. The sequential setting poses new challenges such as not maltreating too many individuals in the search for the best treatment. Furthermore, in contrast to most work, we focus on the problem of a policy maker who targets a general (combination of) functionals of the outcome distribution of the treatments.

Second, the sequential setting adopted in this paper is related to the literature on multi-armed bandit problem which study sequential decision making under uncertainty. While this literature generally focuses on problem without covariates in which one is interested in the mean of the distribution only, we give an overview of the most related papers here. Robbins (1952) was the first to introduce an algorithm with provable performance guarantees. In particular, he showed that the average reward will converge to the mean of the best arm.

The algorithm most relevant for our work is the Upper Confidence Bound (UCB) strategy of Lai and Robbins (1985), for targeting unconditional means, which was further refined in Auer et al. (2002). The underlying idea of our FSA policy is similar to the one of UCB. However, as will be seen, the analysis of the FSA policy is very different from the one of UCB since the FSA policy is designed to target (combinations of) general functionals of the distribution of treatment outcomes instead of the mean only. Furthermore, we allow for the presence of covariates on which the distribution of treatment outcomes can depend. In this sense, the works of Rigollet and Zeevi (2010) and Perchet and Rigollet (2013) are related to our paper — both consider a setting targeting the distribution with the highest conditional mean in the presence of covariates. The former paper studies a UCB type policy while the latter studies the successive elimination algorithm and verifies its minimax optimality. Kock and Thyrsgaard (2017) considered a setting in which the policy maker is also interested in how risky a treatment is and takes this into account by targeting a tradeoff between expected outcome and variance of the treatments. Finally, Cassel et al. (2018) consider bandit problems where the target can be a general (risk) criterion defined on the empirical distribution functions of the path of assignments. For a good general overview of multi-armed bandit problems for targeting the unconditional mean of the distributions to be sampled from we refer to Bubeck and Cesa-Bianchi (2012).

2 Sequential treatment allocation without covariates

To set the stage for the general setting, we begin by considering a treatment allocation problem where no covariates are observed prior to assigning the treatment. This can also be interpreted as (correctly) assuming that the outcome of the treatment is independent of the (observed) covariates. In most settings this is clearly not a realistic assumption. However, this stripped case clearly illustrates the main ideas of our policy .

2.1 Notation

We here present some notation used throughout the paper. For any $x \in \mathbb{R}^d$, $||x||$ denotes the ℓ_2-norm. Furthermore, for any $a < b$ and $d \in \mathbb{N}$, $\mathcal{B}([a,b]^d)$ denotes the Borel σ-field on $[a,b]^d$ equipped with the usual topology. Subsequently, let $D_{cdf}(\mathbb{R})$ denote the set of cumulative distribution functions (cdfs) on $\mathbb{R}$, and for real numbers $a < b$ let $D_{cdf}([a,b])$ denote the set of elements $F \in D_{cdf}(\mathbb{R})$ such that $F(a-) = 0$ and $F(b) = 1$. Furthermore, we shall denote the supremum metric on $D_{cdf}(\mathbb{R})$ by $\|.\|_\infty$.

2.2 Setup

Consider a setting where at each point in time $t = 1, ..., N$ a policy maker must assign one of K treatments to an individual. Thus, t can also be thought of as indexing individuals instead of time. We shall allow the total number of treatments to be made, N, to be a random variable. This reflects that in many treatment problems the policy maker does not know a priori how many treatments will be made. For example, one does not know at the beginning of the year how many individuals will become unemployed. The aim of the policy maker is to assign as many of the N individuals as well as possible. However, the policy maker does not know which of the K treatments is the best. The observational structure is as follows: after assigning a treatment, its outcome is observed but the policy maker does not observe the counterfactuals, i.e., what would have happened if another treatment had been assigned. Upon observing the outcome of treatment $t \in 1, ..., N-1$, individual t arrives and must be assigned to a treatment based on the information gathered from all *previous* assignments. Thus, the data set is gradually constructed in the course of the treatment program, and the policy maker seeks to sample in such a way as to maximize cumulative welfare by assigning the *unknown* best treatment as often as possible.

To be precise, let $a < b$ and let $Y_{i,t} \in [a, b]$ be the outcome of treatment $i \in \mathcal{I} := \{1, ..., K\}$ at time/individual $t \in \{1, ..., N\}$. A policy is a (recursively) defined sequence of (Borel measurable) functions $\pi = \{\pi_t\}_{t=1}^{\infty}$ which can depend on observed variables only: the first treatment is some element in $\mathcal{I}$ not depending on any non-existing previous treatment outcomes. The second treatment can depend on $Z_1 := Y_{\pi_1,1}$ such that $\pi_2 : [a, b] \to \mathcal{I}$. In general,

$$\pi_t : [a, b]^{(t-1)} \to \mathcal{I}$$

and we write $\pi_t(Z_{t-1})$ where $Z_{t-1} := (Y_{\pi_{t-1}(Z_{t-1}),t-1}, ..., Y_{\pi_1,1})$. Thus, Z_{t-1} is the information available after the $(t-1)$-th treatment outcome was observed. For convenience, the dependence of π_t on Z_{t-1} is often suppressed. We stress that the policy maker can assign only one treatment $\pi_t \in \mathcal{I}$ to each individual and observes the outcome of that treatment, i.e., $Y_{\pi_t,t}$, only. This is in accordance with most real life situations: one does generally not observe the counterfactuals. Note also that restricting attention to problems where only one of the K treatments can be assigned does not exclude that a treatment consists of a combination of several treatments (for example a combination of several drugs) — one simply defines this combined treatment as a separate treatment at the expense of increasing the set of potential treatments.

While we assume that $Y_t = (Y_{1,t}, ..., Y_{K,t})$ are distributed identically and independently across t, the joint distribution of the treatment outcomes is left unspecified. In particular, given t, the dependence structure of $Y_{i,t}$ and $Y_{j,t}$ is not restricted. Let $\mathbb{P}^i$ denote the outcome distribution on $\mathcal{B}([a, b])$ of treatment i with corresponding cdf F^i. Ideally, the policy maker would like to assign every individual to the "best" treatment, in the sense that the outcome distribution for this treatment maximizes a functional $\mathsf{T} : D_{cdf}([a, b]) \to \mathbb{R}$, where we recall that $D_{cdf}([a, b])$ denotes the set of cdfs F such that $F(a-) = 0$ and $F(b) = 1$, i.e., the set of cdfs with support $[a, b]$. The specific functional used depends on the application, and encodes the particular characteristic of the distribution the policy maker is interested in. To give specific examples, the functional could be a (combination of) welfare-, inequality-, or poverty-measures, see Appendix A. It could also be a quantile, a (trimmed) moment, a U-functional, or an

L-functional see Appendix C. Given T, the goal is to find a policy π that minimizes the "regret"

$$R_N(\pi) = \sum_{t=1}^{N} \left(\max_{i \in \mathcal{I}} \mathsf{T}(F^i) - \mathsf{T}(F^{\pi_t}) \right). \tag{1}$$

The loss for assigning treatment π_t instead of a best treatment is $\max_{i \in \mathcal{I}} \mathsf{T}(F^i) - \mathsf{T}(F^{\pi_t})$. Thus, we consider a policy maker whose goal is to incur as little loss as possible for as many individuals as possible. Of course there are situations in which this is not the goal of the policy maker. For example, one may only be interested in using the treatments to gather as much information as possible about a characteristic of the treatment outcome distributions by the end of N treatments without regard to the loss each individual incurs. While this exception is definitely also an interesting problem, we shall focus on finding policies that minimize (1) since treating as many individuals as possible as well as possible is a common objective for policy makers.

For every treatment i define $\Delta_i := \max_{1 \leq k \leq K} \mathsf{T}(F^k) - \mathsf{T}(F^i)$ as the loss due to assigning treatment i instead of an optimal one. Then, the regret can also be written as

$$R_N(\pi) = \sum_{i:\Delta_i > 0} \Delta_i \sum_{t=1}^{N} 1_{\{\pi_t = i\}} = \sum_{i:\Delta_i > 0} \Delta_i S_i(N) \tag{2}$$

where $S_i(N) = \sum_{t=1}^{N} 1_{\{\pi_t = i\}}$ is the number of times treatment i is assigned in the course of N treatments.

Throughout the paper, we shall assume that the functional T of interest satisfies the following assumption:

Assumption 2.1. *The functional T is well defined on $D_{cdf}([a,b])$ for $a < b$, and for $\mathscr{D} \subseteq D_{cdf}([a,b])$ there exists a real number C such that:*

$$|\mathsf{T}(F) - \mathsf{T}(G)| \leq C \|F - G\|_\infty \quad \textit{for every} \quad F \in \mathscr{D} \textit{ and every } G \in D_{cdf}([a,b]). \tag{3}$$

Remark 2.2. The set $\mathscr{D}$ appearing in Assumption 2.1 encodes the assumptions imposed on the cdfs of each treatment outcome, i.e., on $F^1, \ldots, F^K$. In particular, the larger $\mathscr{D}$, the less restrictive are the assumptions imposed on $F^1, \ldots, F^K$. Ideally, one would thus like $\mathscr{D} = D_{cdf}([a,b])$, which, however, is too much to ask for many functionals. Furthermore, there is a trade-off between C and $\mathscr{D}$, in the sense that a "larger" class $\mathscr{D}$ leads to a larger constant C. Note also that Assumption 2.1 implies that the *restriction* of T to $\mathscr{D}$ is Lipschitz continuous w.r.t. $\|.\|_\infty$. But, Assumption 2.1 does *not* require T to be Lipschitz continuous on all of $D_{cdf}([a,b])$.

Remark 2.3. Assumption 2.1 is satisfied for many popular functionals arising in applied economics. We provide a detailed discussion together with formal results in Appendix A, where we consider many important inequality-, welfare-, and poverty measures. The inequality measures we discuss in Appendix A.1 include the Schutz-coefficient, the Gini-index, the class of linear inequality measures of Mehran (1976), the generalized entropy family (which includes Theil's index), the Atkinson family of inequality indices (Atkinson (1970)), and the family of Kolm-indices (Kolm (1976a)). In most cases, we discuss both relative and absolute versions of these measures. In Appendix A.2 we provide results for welfare measures based on inequality measures discussed in Appendix A.1. The poverty measures we discuss in

Appendix A.3 are the headcount ratio, the family of poverty measures of Sen (1976) (in the generalized form of Kakwani (1980)), and the family of inequality measures suggested by Foster et al. (1984). We emphasize that the results in Appendices A.1, A.2 and A.3 are obtained from a set of more abstract results of independent interest that we establish in Appendix C. These results establish Assumption 2.1 for U-functionals (i.e., population versions of U-statistics), quantiles, L-functionals (population versions of L-statistics), and for truncated versions of U-functionals.

We also need an assumption that guarantees that the functional T evaluated at empirical cdfs is measurable.

Assumption 2.4. *For every $m \in \mathbb{N}$ the function that maps $x \in [a, b]^m$ to T evaluated at $m^{-1} \sum_{j=1}^{m} \mathbb{1}_{x_j \leq \cdot}$, the empirical cdf corresponding to $x = (x_1, \ldots, x_m)$, is Borel measurable.*

Assumption 2.4 is typically satisfied and poses no practical restrictions.

3 Functional sequential allocation policy and regret bounds

We now turn to describing our treatment policy, the *Functional Sequential Allocation* (FSA) policy, and its properties. The policy is inspired by the UCB strategy of Lai and Robbins (1985) for multi-armed bandit problems. While the UCB policy was designed for targeting the mean of a distribution, the FSA policy can target any functional. Furthermore, Section 4 allows covariates to influence treatment outcomes.

We need some more notation to introduce the FSA policy. Given a policy π, a natural number t and treatment i, we shall define the random set

$$\mathbb{S}_{i,t}(\pi) = \{s \in \{1, \ldots, t\} : \pi_s = i\} \subseteq \{1, \ldots, t\}.$$

That is, $\mathbb{S}_{i,t}(\pi)$ contains all those individuals $s \in \{1, \ldots, t\}$ the policy π has assigned to treatment i. We shall often just write $\mathbb{S}_{i,t}$ instead of $\mathbb{S}_{i,t}(\pi)$. Furthermore, we shall denote the cardinality of $\mathbb{S}_{i,t}(\pi)$ by $S_i(t)$. Note that $S_i(t) = \sum_{s=1}^{t} \mathbb{1}_{\{\pi_s = i\}}$ holds, and that $S_i(t)$ is the number of times treatment i has been assigned in the first t treatments. To formulate our algorithm, we also define the empirical cdf

$$\hat{F}_{i,t}(.) := S_i(t)^{-1} \sum_{s \in \mathbb{S}_{i,t}} \mathbb{1}_{\{Y_{i,s} \leq .\}}, \tag{4}$$

Observe that this just denotes the empirical cdf based on the individuals assigned to treatment i up to time t. The policy we analyze is defined as follows:

Functional Sequential Allocation: Let C be the Lipschitz coefficient in Assumption 2.1. Then, the FSA policy $\hat{\pi}$ with parameter $\beta > 0$ proceeds as:

1. If $t \in \{1, \ldots, K\}$, assign treatment t, i.e. $\hat{\pi}_t = t$.
2. If $t \geq K + 1$, assign

$$\hat{\pi}_t \in \arg\max_{i \in \mathcal{I}} \left\{ \mathsf{T}(\hat{F}_{i,t-1}) + C\sqrt{\beta \log(t)/2S_i(t-1)} \right\}$$

After the K initialization rounds, the FSA policy assigns a treatment that i) is promising in the sense that $\mathsf{T}(\hat{F}_{i,t-1})$ is large or ii) has not been well explored in the sense that $S_i(t-1)$ is small. The parameter β is chosen by the researcher and indicates the weight put on assigning scarcely explored treatments, i.e treatments with low $S_i(t-1)$. A regret minimizing choice of β is given after Theorem 3.1 below. Note also that the FSA policy does *not* require knowledge of the, often unknown, number of treatments to be made (N). In this sense, the policy falls in the class of "*anytime* strategies", in the sense that it does not use the number of assignments to be made when making assignments and its regret guarantees hold for any termination point.

Below, we use the notation $\overline{\log}(x) = \max(\log(x), 1)$. Note that $\overline{\log}(x) = \log(x)$ if $x \geq e$.

Theorem 3.1. *Suppose that the number of treatments, N, has expectation n and is independent of treatment outcomes. Under Assumptions 2.1 and 2.4, the cumulative regret of the FSA policy $\hat{\pi}$ with parameter $\beta > 2$ satisfies*

$$\sup \mathbb{E}[R_N(\hat{\pi})] \leq c\sqrt{Kn\overline{\log}(n)} \tag{5}$$

where the supremum is over all K-tuples of $\mathscr{D}$ and $c = c(\beta, C)$ is a constant, defined in the proof of the theorem, that depends on β and C.

Theorem 3.1 provides an upper bound on the maximal regret incurred by the FSA policy in the absence of covariates. As seen in Theorem 3.6 below, this bound is minimax optimal in n up to the factor $\sqrt{\log(n)}$. Note that the "expected per person regret" $\mathbb{E}(R_N(\hat{\pi}))/n$ tend to zero as n tends to infinity. The choice parameter β can be chosen optimally as a function of C to minimize c. In particular, inspection of the proof shows that $\beta = 2 + \sqrt{2}$ minimizes $c(\beta, C)$ and implies $c \leq \sqrt{11}C$. Finally, we remark that it is sensible that the upper bound on regret is increasing in the number of available treatments K as it becomes harder to find the best treatment as the number of available treatments increases.

We now turn to showing the near-minimax optimality of the upper bound on maximal regret in Theorem 3.1. It suffices to show that for *any* policy the maximal regret must be large against a certain family of K-tuples of distributions of treatment outcomes (we shall consider $K = 2$). To this end, consider the following family of distributions on $\mathcal{B}([0,1])$, but cf. also Remark 3.5 below.

Definition 3.2. *Let $\mathcal{H} = \{\mathbb{P}_{h_a} : a \in (-1, \infty)\}$ where $\mathbb{P}_{h_a}$ is the distribution on $\mathcal{B}([0,1])$ with density $h_a(y) = (1+a)y^a 1_{\{y>0\}}$, and corresponding cdf H_a with $H_a(y) = y^{a+1}$ for $y \in [0,1]$, $H_a(y) = 0$ for $y \in (-\infty, 0)$ and $H_a(y) = 1$ else.*

Assumption 3.3. *Assume that there exists $\bar{a} > -1, c > 0$ and $\delta > 0$ such that $H_a \in \mathscr{D}$ for all $a \in [\bar{a} - \delta, \bar{a} + \delta] \subseteq (-1, \infty)$. Furthermore, either*

$$\mathsf{T}(H_{a_2}) - \mathsf{T}(H_{a_1}) \geq c(a_2 - a_1) \tag{6}$$

holds for all $a_1, a_2 \in [\bar{a} - \delta, \bar{a} + \delta] \subseteq (-1, \infty)$ such that $a_1 \leq a_2$, or

$$\mathsf{T}(H_{a_2}) - \mathsf{T}(H_{a_1}) \leq -c(a_2 - a_1)$$

holds for all $a_1, a_2 \in [\bar{a} - \delta, \bar{a} + \delta] \subseteq (-1, \infty)$ such that $a_1 \leq a_2$.

Remark 3.4. The requirement on T in Assumption 3.3 is a local uniform monotonicity condition on $a \mapsto \mathsf{T}(H_a)$. It is rather mild and satisfied, for example, if $a \mapsto \mathsf{T}(H_a)$ is continuously differentiable on $(\bar{a}-\delta, \bar{a}+\delta)$ with a derivative bounded away from zero (this can be seen by the mean value theorem). This requirement is, in turn, easily seen to be satisfied when T is any moment or quantile. Intuitively, Assumption 3.3 rules out that T is too flat. For example, all policies would incur zero regret if T is constant and it is thus sensible that some strict monotonicity is needed in order to prove non-trivial lower bounds on regret. We stress that the local uniform monotonicity is needed only at *one* point $\bar{a} \in (-1, \infty)$.

Remark 3.5. The only property of $\mathcal{H}$ that is used in the proof of Theorem 3.6 below is that the Kullback-Leibler divergence between any two members of $\mathcal{H}$ is *sub-quadratic* in a sense made precise in Lemma D.3 and surrounding discussion. Thus, the family $\mathcal{H}$ can be replaced by any other family of distributions with this sub-quadratic property as well as satisfying Assumption 3.3.

Out next result shows that the maximal/uniform regret in Theorem 3.1 is optimal as a function of n up to a multiplicative factor $\sqrt{\log(n)}$. It suffices to consider N non-random, $[a, b] = [0, 1]$ and $K = 2$.

Theorem 3.6. *Let Assumptions 2.1 and 3.3 be satisfied and consider a treatment problem with $N = n$ non-random and $K = 2$ treatments. Then, for any policy π, there exists a $c_l > 0$ such that*

$$\sup \mathbb{E} R_n(\pi) \geq c_l \sqrt{n},$$

where the supremum is over all two-tuples of distributions on $\mathcal{B}([0, 1])$.

Remark 3.7. While the FSA policy does not need to know N, inspection of the proof of Theorem 3.6 shows that even when N is non-random (such that $N = \mathbb{E}[N] = n$) and *known*, even a policy that requires knowledge of n must incur a maximal regret of order $\sqrt{n}$. The FSA policy is guaranteed to incur a maximal regret not much more than this even without knowing n, cf. Theorem 3.1 above.

4 Treatment allocation with covariates

The results up to this point have been for treatment allocation problems without any covariates being observed on an individual prior to treatment allocation. While the results for this problem will be useful in the present section, it is often too restrictive to assume that the treatment outcomes do not depend on the characteristics of the person to be treated. For example, one medicine may work very well (in terms of the functional of interest) for one person while it may be outright dangerous to another person if he is allergic to some of the substances.

We now suppose that prior to assigning each treatment the policy maker observes a vector of covariates. More precisely, let $X_t \in [0, 1]^d$, $d \in \mathbb{N}$ be the vector of covariates observed on individual t prior to the treatment assignment and assume that the random vector $(Y_{1,t}, \ldots, Y_{K,t}, X_t)$ is iid across t. For each $x \in [0, 1]^d$, let $F^i(\cdot, x)$ be the distribution function of $Y_{i,t}$ conditional on $X_t = x$. The corresponding probability measure on $\mathcal{B}([a, b])$ is denoted $\mathbb{P}^i(\cdot, x)$.

A policy is now a (recursively) defined sequence of Borel measurable functions $\pi = \{\pi_t\}_{t=1}^{\infty}$ which can depend on observed variables only: the first treatment can depend on X_1 only, so $\pi_1 : [0, 1]^d \to \mathcal{I}$. The second treatment can depend on X_2 and $Z_1 := (Y_{\pi_1(X_1)}, X_1)$ such that $\pi_2 : [0,1]^d \times [a,b] \times [0,1]^d \to \mathcal{I}$. In general

$$\pi_t : [0,1]^d \times [[0,1]^d \times [a,b]]^{t-1} \to \mathcal{I}$$

and we write $\pi_t(X_t, Z_{t-1})$ where $Z_{t-1} := (Y_{\pi_{t-1}(X_t, Z_{t-1})}, X_{t-1}, ..., Y_{\pi_1(X_1)}, X_1)$. We shall often use that it is convenient to suppress the dependence of π_t on Z_{t-1} and write $\pi_t(X_t)$. In this case, π_t is, of course, a random function as it implicitly depends on Z_{t-1}.

The benchmark for our policy will be the infeasible (oracle) policy which *knows* the true conditional treatment outcome distributions $F^i(\cdot, x)$ and for an individual with characteristics x assigns the treatment with the optimal conditional distribution, i.e assigns [1]

$$\pi^\star(x) = \arg\max_{i \in \mathcal{I}} \mathsf{T}(F^i(\cdot, x))$$

where ties are broken arbitrarily such that $\mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) = \max_{i=1,\dots,K} \mathsf{T}(F^i(\cdot, x))$. This way of defining the oracle reflects the general difference to the setting without covariates: we now attempt to get as close as possible to the welfare we could have obtained had we known for *each* individual which treatment is best. The best treatment now depends on the characteristics of the individual and thus conditional distributions replace the unconditional ones in Section 2. One can still target all functionals T of the conditional distributions $F^i(\cdot, x)$ as long as Assumption 2.1 is satisfied, cf. the examples given in Appendix A.

In the presence of covariates, the regret of a policy π is defined as

$$R_N(\pi) = \sum_{t=1}^{N} \mathsf{T}\big(F^{\pi^\star(X_t)}(\cdot, X_t)\big) - \sum_{t=1}^{N} \mathsf{T}\big(F^{\pi_t(X_t)}(\cdot, X_t)\big) \tag{7}$$

Our goal is to provide sharp upper bounds on the expected value of the regret. Thus, we strive to get as close as possible to the welfare we could have attained had we assigned the optimal treatment $\pi^\star(x)$ for each individual. We stress again that the optimal treatment now depends on the individual's characteristics through x.

Remark 4.1. Without any assumptions on the map $x \mapsto F^i(y, x)$, the problem does not have any interesting solution in the sense that *any* policy has maximal regret that increases linearly in n (since Assumption 2.1 implies that T is bounded, such that no policy has regret of larger order than n, this also implies that any policy is minimax optimal). In fact we show in Theorem 5.2 below that even if $x \mapsto F^i(y, x)$ is continuous on (and thus also uniformly continuous on $[0,1]^d$), the maximal regret of *any* policy must increase linearly in n. Hence, we impose the following, minimally stronger, Hölder continuity condition on $F^i(y, x)$ which will be just enough to ensure existence of (near) minimax optimal policies with sub-linear regret.

[1] Note that we allow π_t, the assignment of a policy π for individual t, to be a function of X_t *and* Z_{t-1}. Thus, one may argue that the oracle should assign the treatment $i \in \mathcal{I}$ for which the conditional distribution of $Y_{i,t}$ given X_t *and* Z_{t-1} maximizes T. However, by independence of $(Y_{i,t}, X_t)$ and Z_{t-1}, this distribution *is* $\mathbb{P}^i(\cdot, x)$, i.e.,the conditional distribution of $Y_{i,t}$ given X_t.

Assumption 4.2. *There exist $\gamma \in (0,1]$ and $L > 0$ such that for all $i \in \mathcal{I}$*

$$|F^i(y, x_1) - F^i(y, x_2)| \leq L||x_1 - x_2||^\gamma \text{ for all } y \in \mathbb{R} \text{ and all } x_1, x_2 \in [0,1]^d.$$

In addition, for each $i \in \mathcal{I}$ and $x \in [0,1]^d$, $F^i(\cdot, x)$ belongs to $\mathscr{D}$ of Assumption 2.1.

Assumption 4.2 requires that the distribution functions $F^i(\cdot, x_1)$ and $F^i(\cdot, x_2)$ are close to each other whenever x_1 and x_2 are close. The assumption essentially requires that individuals with similar characteristics have similar outcome distributions for each treatment. Since any policy must generically incur linear in n maximal regret without Assumption 4.2, Hölder continuity is in this sense the weakest possible form of continuity ensuring existence of policies with non-trivial upper bounds on regret.

Assumption 4.3. *The distribution $\mathbb{P}_X$ of covariate X_t has a density with respect to the Lebesgue measure on $\mathcal{B}([0,1]^d)$ that is bounded above and below by $\bar{c}$ and $\underline{c} > 0$, respectively.*

Assumption 4.3 restricts the covariates to be continuous but finitely discrete covariates can also be allowed for by simply running a separate policy for each of the values of the discrete covariates.

4.1 The functional sequential allocation policy in the presence of covariates

Heuristically, the idea of the FSA policy in thepresence of covariates is to group together individuals with similar values of the covariates, and then implement the FSA policy without covariates from Section 3 on each group separately. This amounts to targeting the treatment that is best on average in each group instead of fully individualizing the treatments. It thus strikes a middle ground between individualization and and having enough observations to estimate the treatment outcome distributions in each group. In other words, the grouping amounts to choosing a rectangular kernel in a *sequential* nonparametric estimation problem of $x \mapsto F^i(y, x)$.

More precisely, let $\{B_1, ..., B_F\} \subseteq \mathcal{B}([0,1]^d)$ be a partition of the space of covariates $[0,1]^d$, i.e. $\cup_{i=1}^F B_i = [0,1]^d$ and $B_i \cap B_j = \emptyset$ for $i \neq j$. In addition, let $V_j = \sup_{x,y \in B_j} \|x - y\|$ be the maximal distance between any two points in B_j and $\bar{B}_j = \lambda_d(B_j) > 0$ be the Lebesgue measure of B_j.

In order to define the FSA policy $\bar{\pi}$ in the presence of covariates precisely, let $N_j(t) = \sum_{s=1}^t 1_{\{X_s \in B_j\}}$ be the number of individuals with covariates in B_j in the course of t treatment assignments and let $\hat{\pi}_{B_j,r}$ be the assignment made by the FSA policy without covariates to r-th individual in group B_j with the parameter $\beta > 2$.[2] We then define $\bar{\pi}_t : [0,1]^d \to \mathcal{I}$ as

$$\bar{\pi}_t(x) = \hat{\pi}_{B_j, N_j(t)} \qquad \text{if } x \in B_j \tag{8}$$

where $\hat{\pi}_{B_j,N_j(t)}$ of course depends on the $N_j(t)$ previous treatment outcomes observed for individuals with covariates in B_j. Note also that the covariates are only used to assign group membership to an individual. Since the FSA policy without covariates is used separately for each group, we are effectively targeting a treatment that is best for the average individual in each group. For group j, this means targeting a treatment that attains $\max_{i \in \mathcal{I}} \mathsf{T}(F_j^i)$ where F_j^i is the cumulative distribution function defined as

$$F_j^i(y) := \frac{1}{\mathbb{P}_X(B_j)} \int_{B_j} F^i(y, x) \mathbb{P}_X(dx). \tag{9}$$

[2]We recommend running the FSA policy for each group with $\beta = 2 + \sqrt{2}$ as this minimizes c in Theorem 3.1.

4.2 An upper bound on the maximal regret of the FSA policy with covariates

Let $\mathcal{S}_0 = \mathcal{S}_0(\gamma, L, \underline{c}, \bar{c})$ be the set of distributions of $(Y_{1,t}, ..., Y_{K,t}, X_t)$ that satisfy Assumptions 4.2 and 4.3.

Theorem 4.4. *Suppose that the number of treatments, N, has expectation n and is independent of treatment outcomes and covariates. Consider a partition characterized by $\{V_1, ..., V_F\}$ and $\{\bar{B}_1, ..., \bar{B}_F\}$. Under Assumptions 2.1, 2.4, 4.2 and 4.3, there exists a $c > 0$ such that*

$$\sup_{\mathcal{S}_0} \mathbb{E}[R_N(\bar{\pi})] \leq \sum_{j=1}^{F} \left(c\sqrt{K\bar{c}\bar{B}_j n \overline{\log}(\bar{c}\bar{B}_j n)} + 2CLV_j^{\gamma} n\bar{c}\bar{B}_j \right). \tag{10}$$

Theorem 4.4 gives an upper bound on the regret of the FSA policy in the presence of covariates for *any* choice of grouping individuals. This flexibility may be useful since a policy maker is sometimes constrained by ethical or legislative reasons in the way he groups individuals such that he can not choose the partition $\{B_1, ..., B_F\}$ that minimizes the upper bound on regret in (10).

The upper bound on regret consists of two parts. The first part is very similar to the uniform part of Theorem 3.1; the difference being that the total number of individuals expected to be treated, n, has now been replaced by an upper bound on the number of individuals expected to fall in group B_j, $\bar{c}n\bar{B}_j$. Thus, the first part of the upper bound on regret is the regret we expect to accumulate on each group compared to always assigning the treatment that is best for the average individual in that group. The second part of the upper bound on expected regret is the approximation error incurred due to targeting the treatment that is best for the average individual in group B_j, i.e.,targeting $\max_{i \in \mathcal{I}} \mathsf{T}(F_j^i)$, instead of targeting $\mathsf{T}\big(F^{\pi^\star(x)}(\cdot, x)\big)$. It is sensible that this approximation error is increasing in the size of the groups as measured by V_j and $\bar{B}_j$.

A specific type of partition that achieves near-minimax optimal regret over the class of distributions $\mathcal{S}_0$ is "quadratic groups". It uses hard thresholds for each entry of X_t to create hypercubes that partition $[0,1]^d$. The groups thus created do not only attain low regret but are also relevant in practice due to their simplicity and resemblance to real ways of grouping people. More precisely, fix $P \in \mathbb{N}$ and define

$$B_k = \left\{ x \in [0,1]^d : \frac{k_l - 1}{P} \leq x_l \leq^{(*)} \frac{k_l}{P},\ l = 1, ..., d \right\} \tag{11}$$

for $k = (k_1, ..., k_d) \in \{1, ..., P\}^d$ where $\leq^{(*)}$ is weak for $k_l = P$ and strict otherwise. Thus, P is the number of splits along each dimension of X_t creating a partition of $[0,1]^d$ consisting of P^d smaller hypercubes $B_1, ..., B_{P^d}$ with side lengths $1/P$.

Corollary 4.5. *Suppose that the horizon N has expectation n and is independent of treatment outcomes and covariates. Use the partition in (11) with $P = \lceil n^{1/(2\gamma+d)} \rceil$. Under Assumptions 2.1, 2.4, 4.2 and 4.3, we obtain that*

$$\sup_{\mathcal{S}_0} \mathbb{E}\left[R_N(\bar{\pi})\right] \leq c\sqrt{K\overline{\log}(n)}\, n^{1-\frac{\gamma}{2\gamma+d}}. \tag{12}$$

for some $c > 0$.

Corollary 4.5 reveals that it is possible to achieve sublinear (in n) regret under the smoothness on the conditional distributions guaranteed by Assumption 4.2. Note that a curse of dimensionality is present in the sense that the upper bound on regret gets close to linear in n as d, the number of covariates, increases. The presence of this effect is due to the fact that as part of the regret minimization, we *sequentially* estimate the conditional distributions of the treatment outcomes and each $F^i(y, \cdot)$ is a function of a d-dimensional variable. Finally, it is also to be expected that the regret is increasing in the number of available treatments K as more observations must be used for experimentation when more treatments are available.

The upper bound on the maximal regret in Corollary 4.5 is near-minimax optimal as we shall make precise in Theorem 5.3 below. Thus, if there is nothing prohibiting the choice of groups in (11), not much can be gained from a maximal regret point-of-view in searching for "better" partitions under the given set of assumptions.

Finally, we remark that the grouping in (11) with $P = \lceil n^{1/(2\gamma+d)} \rceil$ requires knowledge of the expected number of treatments n. If the exact number of treatments is known a priori (as it is in many classical treatment problems) then n is trivially known. If, however, n is unknown one can instead use the doubling trick to attain upper bound on the maximal regret that are of the same order as in Corollary 4.5, but with slightly higher multiplicative constants. In essence, the doubling trick works by resetting the policy at time 2^m, $m \in \mathbb{N}$. The name "doubling trick" comes from the fact that the length between subsequent resets of the policy doubles between subsequent resets. Importantly, the length of each treatment period is known. The doubling trick is a general tool in games of unknown horizon and we refer to Shalev-Shwartz (2012) for more details.

4.3 Stronger regret guarantees and number of suboptimal assignments

So far, our results in the case where covariates are present have only assumed that the conditional distribution of the treatment outcomes is Hölder continuous, cf. Assumption 4.2. If, furthermore, it is also the case that the best and second best treatment are "well-separated", the upper bound on maximal regret in Section 4.2 can be lowered slightly. Formally, introduce the second best treatment $\pi^\sharp(x)$ that is, for any $x \in [0,1]^d$, if $\min_{i \in \mathcal{I}} \mathsf{T}(F^i(\cdot, x)) < \mathsf{T}\big(F^{\pi^\star(x)}(\cdot, x)\big)$, then $\pi^\sharp(x)$ satisfies

$$\mathsf{T}\big(F^{\pi^\sharp(x)}(\cdot, x)\big) = \max_{i \in \mathcal{I}} \{\mathsf{T}(F^i(\cdot, x)) : \mathsf{T}(F^i(\cdot, x)) < \mathsf{T}\big(F^{\pi^\star(x)}(\cdot, x)\big)\},$$

and $\pi^\sharp(x) = 1$ if $\min_{i \in \mathcal{I}} \mathsf{T}(F^i(\cdot, x)) = \mathsf{T}\big(F^{\pi^\star(x)}(\cdot, x)\big)$. We can now introduce the *margin condition.*

Assumption 4.6. *There exists $\alpha \in (0,1)$ and $C_0 > 0$ such that*

$$\mathbb{P}_X\Big(0 < \mathsf{T}\big(F^{\pi^\star(X)}(\cdot, X)\big) - \mathsf{T}\big(F^{\pi^\sharp(X)}(\cdot, X)\big) \leq \delta\Big) \leq C_0 \delta^\alpha \text{ for all } \delta \in [0,1].$$

The margin condition restricts how likely it is that the best and second best treatment are close to each other. In particular, it limits the probability of these treatments being almost equally good, i.e., it limits how likely it is that the best and second-best treatment are within a δ-margin. Assumptions of the margin condition type have previously been used in the works of Mammen and Tsybakov (1999), Tsybakov (2004), Audibert and Tsybakov (2007) in the statistics literature. In the context of statistical treatment rules, the margin condition has recently been used in the work of Kitagawa and Tetenov (2018) who considered a static (non-sequential arrival of individuals to be treated/information) treatment allocation problem. Finally, the margin condition was also used in the work of Perchet and Rigollet (2013) in the context of a multi-armed bandit problem.

The margin condition does not only allow us to prove sharper upper bounds on maximal regret than in Section 4.2, it also allows us to provide an upper bound on the number of suboptimal assignments made by the FSA policy. In particular, we shall define the total number of suboptimal assignments for a policy π over the course of a total of N assignments as

$$S_N(\pi) = \sum_{t=1}^{N} 1_{\left\{\pi_t(X_t, Z_{t-1}) \notin \arg\max\{\mathsf{T}(F^i(\cdot, X_t)), i=1,\dots,K\}\right\}}.$$

In Theorem 4.8 we establish an upper bound on $\mathbb{E}(S_N(\bar{\pi}))$ which is near minimax optimal.

Let $\mathcal{S} = \mathcal{S}(\gamma, L, \underline{c}, \bar{c}, \alpha, C_0)$ be the set of K-tuples of conditional distributions of $(Y_{1,t}, \dots, Y_{K,t})$ given X_t that satisfy Assumptions 4.2, 4.3 and 4.6. The maximal regret over $\mathcal{S}$ of the FSA policy over can be bounded as follows.

Theorem 4.7. *Suppose that the horizon N has expectation n and is independent of treatment outcomes and covariates. Consider the partition in* (11) *and set $P = \lceil n^{1/(2\gamma+d)} \rceil$. Under Assumptions 2.1, 2.4, 4.2, 4.3 and 4.6, we obtain that*

$$\sup_{\mathcal{S}} \mathbb{E}[R_N(\bar{\pi})] \leq cK\overline{\log}(n) n^{1-\frac{\gamma(1+\alpha)}{2\gamma+d}} \tag{13}$$

for some constant $c > 0$.

Compared to Corollary 4.5 the exponent on n in the upper bound on regret is now smaller. Thus, in the presence of the margin condition (Assumption 4.6), the regret guarantee on the FSA policy is stronger. Of course, since $\mathcal{S} \subset \mathcal{S}_0$, this is not altogether surprising. We shall see in Theorem 5.1 below that the upper bound on maximal regret is minimax optimal in n up to logarithmic factors.

Our next result shows that the upper bound on maximal regret does not come at the price of excessive experimentation leading to many suboptimal assignments. In fact, Lemma D.5 in the appendix shows that under the margin condition, an upper bound on $\mathbb{E}[S_N(\pi)]$ for any policy π follows as a *consequence* of an upper bound on regret.

Theorem 4.8. *Suppose that the horizon N has expectation n and is independent of treatment outcomes and covariates. Consider the partition in* (11) *and set $P = \lceil n^{1/(2\gamma+d)} \rceil$. Under Assumptions 2.1, 2.4, 4.2, 4.3 and 4.6, we obtain that*

$$\sup_{\mathcal{S}} \mathbb{E}[S_N(\bar{\pi})] \leq c[K\overline{\log}(n)]^{\frac{\alpha}{1+\alpha}} n^{1-\frac{\alpha\gamma}{2\gamma+d}}$$

for some constant $c > 0$.

The upper bound in Theorem 4.8 on the maximal number of suboptimal assignments made is a useful theoretical guarantee since it limits the number of individuals who receive suboptimal treatments. A policy which only ensures high total welfare (low regret), may not be ethically viable if too many individuals are maltreated. Finally we note that *no* policy can be expected to make substantially fewer suboptimal assignments than the FSA policy since Step 4 of the proof of Theorem 5.1 below actually shows that for *any* policy there exist distributions in $\mathcal{S}$ for which the number of suboptimal assignments must be at least of order $n^{1-\frac{\alpha\gamma}{2\gamma+d}}$.

5 Lower bounds on maximal regret and minimax optimality of the FSA policy

In this section we prove formally the impossibility results mentioned in Remark 4.1 and establish (near) minimax optimality of the FSA policy in several settings. As was the case in the setting without covariates in Section 3, only a local uniform monotonicity of T over $\mathcal{H}$ suffices to establish tight lower bounds on maximal regret. We stress, however, that while the assumptions imposed in the present section are the same as in the case without covariates, the proofs are more involved as one must carefully construct conditional distributions of treatment outcomes satisfying Assumptions 4.2 and 4.6 against which large regret must be incurred by *any* policy.

For all lower bounds, we consider the case of $K = 2$ available treatments. Fix a functional T and let Π denote the set of all policies π. For any 2-tuple of conditional distributions (F^1, F^2) of $Y_{1,t}$ and $Y_{2,t}$ given X_t in $\mathcal{S}$ and policy π, we make the dependence of regret on (F^1, F^2) explicit by

$$R_n(\pi) = R_n(\pi, F^1, F^2) = \sum_{t=1}^{n} \left|\mathsf{T}\big(F^1(\cdot, X_t)\big) - \mathsf{T}\big(F^2(\cdot, X_t)\big)\right| 1_{\{\pi^\star(X_t) \neq \pi_t(X_t, Z_{t-1})\}}. \tag{14}$$

Theorem 5.1. *Suppose that X_t is uniformly distributed on $[0,1]^d$ (thus $\underline{c} = \bar{c} = 1$ in Assumption 4.3). Then, under Assumptions 2.1 and 3.3, there exists a $c_l > 0$ such that*

$$\inf_{\pi \in \Pi} \sup_{(F^1, F^2) \in \mathcal{S}} \mathbb{E}[R_n(\pi, F^1, F^2)] \geq c_l n^{1 - \frac{\gamma(1+\alpha)}{2\gamma + d}}. \tag{15}$$

Theorem 5.1 shows that the upper bound on maximal regret of the FSA policy obtained in Theorem 4.7 is only improvable by logarithmic factors. Put differently, the FSA policy is minimax optimal up to logarithmic factors.

5.1 Maximal regret is linear without Assumption 4.2

Let $C[0,1]^d$ denote the set of (uniformly) continuous functions on $[0,1]^d$ and let $\mathcal{S}_C$ denote the 2-tuples of distributions (F^1, F^2) of $Y_{1,t}$ and $Y_{2,t}$ given X_t such that $F^1, F^2 \in C[0,1]^d$. The following theorem, which is a consequence of Theorem 5.1, shows that without the Hölder continuity imposed in Assumption 4.2, no policy exists that has sub-linear maximal regret in n. Furthermore, since every policy is guaranteed to incur no more than linear (in n) regret, this shows that the problem is not well posed without Assumption 4.2. More precisely, even when restricting attention to (uniformly) continuous F^1 and F^2, maximal regret of *any* policy is linear in n. Thus, the following Theorem makes precise Remark 4.1 prior to Assumption 4.2.

Theorem 5.2. *Suppose that X_t is uniformly distributed on $[0,1]^d$ (thus $\underline{c} = \bar{c} = 1$ in Assumption 4.3). Then, under Assumptions 2.1 and 3.3, there exists a $c_l > 0$ such that*

$$\inf_{\pi \in \Pi} \sup_{(F^1, F^2) \in \mathcal{S}_C} \mathbb{E}[R_n(\pi, F^1, F^2)] \geq c_l n.$$

5.2 Lower bound on maximal regret without margin condition

The next result, which gives a lower bound on maximal regret in the absence of the margin condition in Assumption 4.6, is again a consequence of Theorem 5.1.

Theorem 5.3. *Suppose that X_t is uniformly distributed on $[0,1]^d$ (thus, $\underline{c} = \bar{c} = 1$ in Assumption 4.3). Then, under Assumptions 2.1 and 3.3, for all $\varepsilon > 0$ there exists a $c_l(\varepsilon) > 0$ such that*

$$\inf_{\pi \in \Pi} \sup_{(F^1,F^2) \in \mathcal{S}_0} \mathbb{E}[R_n(\pi, F^1, F^2)] \geq c_l(\varepsilon) n^{1-\frac{\gamma}{2\gamma+d}} n^{-\varepsilon}. \tag{16}$$

Comparing the lower bound on maximal regret in Theorem 5.3 to the upper bound on maximal regret of the FSA policy established in Corollary 4.5 reveals that the FSA policy is near-optimal also in this setting. If a policy with strictly smaller maximal regret exists, the order of this improvement must be $o(n^{\varepsilon})$ for all $\varepsilon > 0$, e.g., logarithmic.

5.3 Ignoring covariates

Theorem 3.1 shows that the maximal regret for the FSA policy is guaranteed to increase not much faster than rate $\sqrt{n}$ in the absence of covariates. On the other hand, if $\frac{\gamma(1+\alpha)}{2\gamma+d} < \frac{1}{2}$ (which occurs in case $2\gamma\alpha < d$), the maximal regret must increase faster than $\sqrt{n}$ for any policy in the presence of covariates, cf. Theorem 5.1. At first sight, one could be led to believe that it may sometimes be advantageous to ignore the covariates in order to achieve a lower maximal regret. Note, however, that the oracle targets are defined differently in the context of Theorems 3.1 and 5.1. More precisely, if covariates are available one targets the "best" conditional distribution. Our next result shows, that it is in fact a very bad idea to ignore the covariates unless one *knows* that these are irrelevant (which one rarely does). To be precise, consider polices that are a (recursively) defined sequence of Borel measurable functions $\pi = \{\pi_t\}_{t=1}^{\infty}$ which can depend on observed treatment outcomes (but not covariates): the first treatment is some element in $\mathcal{I}$ not depending on any non-existing previous treatment outcomes. The second treatment can depend on $Z_1 := Y_{\pi_1,1}$ such that $\pi_2 : [0,1] \to \mathcal{I}$. In general, $\pi_t : [0,1]^{(t-1)} \to \mathcal{I}$ and we write $\pi_t(Z_{t-1})$ where $Z_{t-1} := (Y_{\pi_{t-1}(Z_{t-1})}, ..., Y_{\pi_1})$. Thus, Z_{t-1} is the allowed to be used after the previous treatment outcome is observed. Let $\tilde{\Pi}$ denote the collection of such policies. Note the similarity to the definition of a policy in the setting without covariates, cf. Section 2.2.

Theorem 5.4. *Suppose that X_t is uniformly distributed on $[0,1]^d$ (thus $\underline{c} = \bar{c} = 1$ in Assumption 4.3). Then, under Assumptions 2.1 and 3.3, there exists a $c_l > 0$ such that*

$$\inf_{\pi \in \tilde{\Pi}} \sup_{(F^1,F^2) \in \mathcal{S}} \mathbb{E}[R_n(\pi, F^1, F^2)] \geq c_l n. \tag{17}$$

Thus, the maximal regret of any policy ignoring covariates must increase at the worst-case linear rate in n. To illustrate the connections to the FSA policy note that for this policy ignoring the covariates amounts to assigning all individuals to the same group, i.e. $F = 1$ and thus $V_1 = \sqrt{d}, \bar{B}_1 = 1$. Using these quantities in Theorem 4.4 results in a upper bound on regret which is linear in n.

A Verification of Assumption 2.1 for some inequality-, welfare- and poverty-measures

To keep the statements in the subsequent examples simple, and to give some concrete meaning to the discussion in Remark 2.2, we shall now define some sets of cdfs $\mathscr{D}$ that will show up frequently in the following discussion: Given real numbers $a < b$, we shall denote (i) the subset of all cdfs F in $D_{cdf}([a, b])$ that are continuous on $[a, b]$ and right-differentiable on (a, b) with right-sided derivative F^+, say, satisfying $F^+(x) \leq s$ for all $x \in (a, b)$ by $\mathscr{D}^s([a, b])$; (ii) we shall denote the subset of all cdfs F in $D_{cdf}([a, b])$ that are continuous on $[a, b]$ and right-differentiable on (a, b) with right-sided derivative F^+, say, satisfying $F^+(x) \geq r$ for all $x \in (a, b)$ by $\mathscr{D}_r([a, b])$; and (iii) we shall denote the subset of all cdfs F in $D_{cdf}([a, b])$ that are continuous on $[a, b]$ and right-differentiable on (a, b) with right-sided derivative F^+, say, satisfying $r \leq F^+(x) \leq s$ for all $x \in (a, b)$ by $\mathscr{D}_r^s([a, b])$. Furthermore, we shall denote the subset of $\mathscr{D}^s([a, b])$ consisting of all cdfs $F \in \mathscr{D}^s([a, b])$ that are continuous on *all of* $\mathbb{R}$ (and not only on $[a, b]$) by $\mathscr{C}^s([a, b])$, and correspondingly define $\mathscr{C}_r([a, b])$ and $\mathscr{C}_r^s([a, b])$.

To illustrate the scope of our results, and to facilitate their implementation in practice, we shall now discuss several functionals of interest in applied economics that satisfy Assumption 2.1. We emphasize that Appendix C contains a catalog of *general* methods for verifying Assumption 2.1, and that the results in the following three sections are established using these techniques. Therefore, in addition to their intrinsic importance, the following results, and in particular their proofs, also provide a pattern as to how Assumption 2.1 can be verified for functionals that are not explicitly discussed.

A.1 Inequality measures

In this section we verify Assumption 2.1 for functionals that measure the inequality inherent to an, e.g., "income", "wealth" or "productivity", distribution F. Such "inequality measures" are obviously relevant in situations where one intends to select, among various candidates, that treatment (e.g., the introduction of a certain tax) which leads to the most "equal" outcome distribution. To avoid possible misunderstanding, we emphasize that it is neither our goal to discuss theoretical foundations of inequality measures nor to point out their relative advantages and disadvantages. The functional must be chosen by the applied economist, who can—in making such a choice—rely on excellent book-length treatments, e.g., Lambert (2001), Chakravarty (2009) or Cowell (2011), as well as the original sources some of which we shall point out further below.

We first discuss inequality measures that are derived from the Lorenz curve. The first such inequality measure we consider is the *Schutz-coefficient* $\mathsf{S}_{rel}(F)$, say, which is also known as the *Hoover-index* or *Robin Hood-index*. In plain words, this coefficient measures the maximal vertical distance between the 45° line and the Lorenz curve corresponding to F (cf. Gastwirth (1971) or Equation (22) below for a formal definition of the Lorenz curve). It can be shown (e.g., Lambert (2001)) that $\mathsf{S}_{rel}(F)$ coincides with half the *relative mean deviation* index, i.e.,

$$\mathsf{S}_{rel}(F) = \frac{1}{2\mu(F)} \int |x - \mu(F)| dF(x), \tag{18}$$

provided the expression is well defined. Here, the index "rel" signifies that this index is defined "relative" to the mean (as a consequence, multiplying each income by the same amount does not change the inequality index). A corresponding "absolute" variant (i.e., a measure which remains unchanged if one adds to every income the same amount) is obtained by multiplying the relative measure by the mean functional, and is denoted by $\mathsf{S}_{abs}(F) = \frac{1}{2}\int |x - \mu(F)| dF(x)$. We refer to Kolm (1976a,b) for a discussion concerning relative and absolute inequality measures. The following lemma provides conditions under which the relative and absolute Schutz-coefficient satisfies Assumption 2.1.

Lemma A.1. *Let $a < b$ be real numbers. Then, the absolute Schutz-coefficient $\mathsf{T} = \mathsf{S}_{abs}$ satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a,b])$ and $C = b - a$. Next, assume that $a \geq 0$, and define for every $\delta \in (a,b)$ and every $s > 0$ the set*

$$\mathscr{D}(s,\delta) := \{F \in \mathscr{C}^s([a,b]) : \mu(F) \geq \delta\}. \tag{19}$$

*Then, for every $\delta \in (a,b)$ and every $s > 0$ the relative Schutz-coefficient $\mathsf{T} = \mathsf{S}_{rel}$ (*defined *as 0 for the cdf corresponding to point mass 1 at 0) satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}(s,\delta)$ and constant $C = (b-a)(2s + \delta^{-1}) + 5$.*

The next index we consider is the *Gini-index*, which for a cdf F is defined as the area between the Lorenz curve and the 45° line, and which can be written as (cf. again Lambert (2001))

$$\mathsf{G}_{rel}(F) = \frac{1}{2\mu(F)} \int\int |x_1 - x_2| dF(x_1) dF(x_2), \tag{20}$$

provided that the expression is well defined. A corresponding absolute inequality measure is $.5\int\int |x_1 - x_2| dF(x_1)dF(x_2)$, which we denote by $\mathsf{G}_{abs}(F)$. The following lemma provides conditions under which Assumption 2.1 is satisfied for these two Gini-indices:

Lemma A.2. *Let $a < b$ be real numbers, and let $\mathscr{D} = D_{cdf}([a,b])$. For this choice of a, b and $\mathscr{D}$, the functional $\mathsf{T} = \mathsf{G}_{abs}$ satisfies Assumption 2.1 with constant $C = b - a$. Next, assume that $a \geq 0$, and define for every $\delta \in (a,b)$ the set*

$$\mathscr{D}(\delta) := \{F \in \mathscr{D}_{cdf}([a,b]) : \mu(F) \geq \delta\}. \tag{21}$$

*Then, for every $\delta \in (a,b)$ the relative Gini-index $\mathsf{T} = \mathsf{G}_{rel}$ (*defined *as 0 for the cdf corresponding to point mass 1 at 0) satisfies Assumption 2.1 with constant $C = 4\delta^{-1}(b-a)$.*

It can be verified that the Gini-index belongs to the class of *linear inequality measures* (cf. Mehran (1976)). Linear inequality measures are functionals of the form

$$F \mapsto \int_{[0,1]} (u - L(F,u)) dW(u), \quad \text{where} \quad L(F,u) := \mu(F)^{-1} \int_{[0,u]} q_\alpha(F) d\alpha, \tag{22}$$

where W denotes a function on $[0,1]$ (independent of F) with finite total variation. Note that $L(F;u)$ is the Lorenz curve corresponding to F evaluated at u (cf. Gastwirth (1971) and our Equation (94)). The following lemma provides conditions under which a linear inequality measure as defined in Equation (22) satisfies Assumption 2.1. The result relies on Lipschitz-type properties of the Lorenz curve established in Lemma C.14 in the appendix. The generality is bought at the price of comparably strong regularity conditions. This becomes apparent by comparing the regularity conditions to the ones in Lemma A.2.

Lemma A.3. *Let $a < b$ be positive real numbers and let $r > 0$. Assume that $W : [0,1] \to \mathbb{R}$ has finite total variation κ, say. Then, the functional defined in Equation* (22) *satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{C}_r([a,b])$, and constant $C = \kappa a^{-1}(r^{-1} + (b-a)a^{-1})$.*

An absolute version of the linear inequality measure in Equation (22) is obtained after multiplication with $\mu(F)$, and is given by

$$F \mapsto \int_{[0,1]} (\mu(F)u - U(F,u))dW(u), \quad \text{where} \quad U(F,u) := \int_{[0,u]} q_\alpha(F)d\alpha. \tag{23}$$

The following result provides conditions under which such absolute linear inequality measures satisfy Assumption 2.1. As usual, the regularity conditions required for absolute versions of an inequality measure are weaker than the ones needed for the relative version:

Lemma A.4. *Let $a < b$ be real numbers and let $r > 0$. Assume that $W : [0,1] \to \mathbb{R}$ has finite total variation κ, say. Furthermore, denote $|\int_{[0,1]} udW(u)| =: c$. Then, the functional defined in Equation* (23) *satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{C}_r([a,b])$, and constant $C = c(b-a) + r^{-1}\kappa$.*

Another important family of inequality indices is the so-called *generalized entropy family*, cf. Cowell (1980): Given a parameter $\alpha \in \mathbb{R}$, an inequality measure is obtained via (if the involved expressions are well defined)

$$\mathsf{E}_c(F) = \begin{cases} \frac{1}{c(c-1)} \int \left[\left(x/\mu(F) \right)^c - 1 \right] dF(x) & \text{if } c \notin \{0,1\} \\ \int \left(x/\mu(F) \right) \log \left(x/\mu(F) \right) dF(x) & \text{if } c = 1 \\ \int \log \left(x/\mu(F) \right) dF(x) & \text{if } c = 0. \end{cases} \tag{24}$$

The inequality measures corresponding to $c = 1$ is known as Theil's entropy index (cf. also Theil (1967)), and the measure corresponding to $c = 0$ is known as the mean logarithmic deviation (cf. Lambert (2001), p.112). A formal result providing conditions under which generalized entropy measures in the previous display satisfy Assumption 2.1 is presented next. The regularity conditions we need to impose depend on c. In particular, support assumptions inherent in the definition of $\mathscr{D}$ are somewhat weaker in case $c \in (0,1)$.

Lemma A.5. *Let $0 \leq a < b$ be real numbers, and let $c \in \mathbb{R}$.*

1. *If $c \in (0,1)$, then, for every $\delta \in (a,b)$ the functional $\mathsf{T} = \mathsf{E}_c$ (*defined *as 0 for the cdf corresponding to point mass 1 at 0) satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}(\delta)$ (cf. Equation 21) and constant $C = |c(c-1)|^{-1} \left[\delta^{-c}(b^c - a^c) + \delta^{-1}(b-a) \right]$.*

2. *If $c \notin [0,1]$ and $a > 0$, then the functional $\mathsf{T} = \mathsf{E}_c$ satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a,b])$ and $C = |c(c-1)|^{-1}[a^{-c}|b^c - a^c| + |c| \max\left((a/b)^{2c-1}, (b/a)^{2c-1} \right) a^{-1}(b-a)]$.*

3. *If $c \in \{0,1\}$ and $a > 0$, then the functional $\mathsf{T} = \mathsf{E}_c$ satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a,b])$ and with constant $C = a^{-1} + \log(b/a)$ if $c = 0$, and with constant $C = a^{-1} \int_{[a,b]} |1 + \log(x)|dx + ba^{-1} \left\{ |\log(b/a)|(b-a) + a^{-1} \right\}$ if $c = 0$.*

We continue with a family of (relative) inequality indices introduced in Atkinson (1970). This family depends on an "inequality aversion" parameter $\varepsilon \in (0,1) \cup (1,\infty)$. For a fixed ε in that range, the index obtained equals (if the involved quantities are well defined)

$$\mathsf{A}_\varepsilon(F) = 1 - \frac{1}{\mu(F)} \left[\int x^{1-\varepsilon} dF(x) \right]^{1/(1-\varepsilon)}. \tag{25}$$

It is well known (cf., e.g., Lambert (2001) p.112) that A_ε can be written as

$$\mathsf{A}_\varepsilon(F) = 1 - [\varepsilon(\varepsilon - 1)\mathsf{E}_{1-\varepsilon}(F) + 1]^{1/(1-\varepsilon)}. \tag{26}$$

Together with Lemma A.5 this relation can be used to obtain the following result:

Lemma A.6. *Let $0 \leq a < b$ be real numbers, let $\varepsilon \in (0,1) \cup (1,\infty)$ and set $c(\varepsilon) = 1 - \varepsilon$.*

1. *If $\varepsilon \in (0,1)$, then, for every $\delta \in (a,b)$ the functional $\mathsf{T} = \mathsf{A}_\varepsilon$ (*defined *as* 0 *for the cdf corresponding to point mass* 1 *at* 0*) satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}(\delta)$ (cf. Equation 21) and constant $C = c(\varepsilon)^{-1}\left[\delta^{-c(\varepsilon)}(b^{c(\varepsilon)} - a^{c(\varepsilon)}) + \delta^{-1}(b-a)\right]$.*

2. *If $\varepsilon \in (1,\infty)$ and $a > 0$, then the functional $\mathsf{T} = \mathsf{A}_\varepsilon$ satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a,b])$ and constant*

$$C = c(\varepsilon)^{-1}(b/a)^\varepsilon \left\{ [a^{-c(\varepsilon)}|b^{c(\varepsilon)} - a^{c(\varepsilon)}| + |\varepsilon||c(\varepsilon)|^2 (a/b)^{2c(\varepsilon)-1} a^{-1}(b-a)] \right\}.$$

As the last example in this section, we proceed to an important family of (absolute) inequality indices, the *Kolm-indices* (Kolm (1976a), cf. also the discussion in Section 1.8.1 of Chakravarty (2009)). Given a parameter $\kappa > 0$ the corresponding index is defined as

$$\mathsf{K}_\kappa(F) = \kappa^{-1} \log\left(\int e^{\kappa[\mu(F) - x]} dF(x) \right). \tag{27}$$

The following lemma verifies Assumption 2.1 for this class of inequality indices:

Lemma A.7. *Let $a < b$ and let $\kappa > 0$. Then, the functional $\mathsf{T} = \mathsf{K}_\kappa$ satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a,b])$ and constant $C = \kappa e^{\kappa b}(b-a) + [e^{-\kappa a} - e^{-\kappa b}]$.*

A.2 Welfare measures

The most elementary class of social welfare functions are of the form (cf. Atkinson (1970))

$$F \mapsto \int u(x) dF(x), \tag{28}$$

for a utility function u. Such functionals can directly be dealt with the theory developed in Kock and Thyrsgaard (2017). We therefore refer to this article for corresponding results.

There are many important social welfare functions, however, that are not of the simple form (28), and can thus not be treated with the results in Kock and Thyrsgaard (2017). Many such exceptional measures are related to a relative inequality measure $F \mapsto \mathsf{I}_{rel}(F)$, say, via the transformation

$$\mathsf{W}(F) = \mu(F)(1 - \mathsf{I}_{rel}(F)); \tag{29}$$

or (correspondingly) to an absolute inequality measure $F \mapsto \mathsf{I}_{abs}(F)$, say, via the transformation

$$\mathsf{W}(F) = \mu(F) - \mathsf{I}_{abs}(F). \tag{30}$$

The Gini social welfare function (obtained via the previous display upon choosing $\mathsf{I}_{abs} = \mathsf{G}_{abs}$, see the discussion after Equation (20) for a definition of G_{abs}) is one of the most important examples. The following result allows one to use the results from the preceding section in establishing Assumption 2.1 for the two types of social welfare functions in (29) and (30).

Lemma A.8. *Let $a < b$ be real numbers. Then, the following holds:*

1. *Let the relative inequality measure I_{rel} satisfy Assumption 2.1 with $\mathscr{D}_{rel}$ and C. Suppose further that $|1 - \mathsf{I}_{rel}| \leq \gamma < \infty$ holds.*[3] *Then, the social welfare function W derived via Equation (29) satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}_{rel}$ and constant $\gamma(b - a) + \max(|a|, |b|)C$.*

2. *Let the absolute inequality measure I_{abs} satisfy Assumption 2.1 with $\mathscr{D}_{abs}$ and C. Then, the social welfare function W derived via Equation (30) satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}_{abs}$ and with constant $(b - a) + C$.*

The preceding lemma together with Lemma A.2 implies that the Gini social welfare function (defined directly after Equation (30) above) satisfies Assumption 2.1 with $a < b$ real numbers, $\mathscr{D} = D_{cdf}([a, b])$, and constant $C = 2(b - a)$. Similar statements can easily be obtained for social welfare functions corresponding to the class of linear inequality measures via Lemma A.4 (which then covers the class of social welfare functions considered recently in Kitagawa and Tetenov (2017)), or via the other inequality measures discussed in the preceding section.

A.3 Poverty measures

Poverty indices are typically based on a *poverty line*, i.e., a threshold z below which an, e.g., income is classified as "poor". There are two basic approaches to defining z: the absolute approach considers z as fixed (i.e., independent of the underlying income distribution F), whereas the relative approach views $\mathsf{z} = \mathsf{z}(F)$ as a function of the "income distribution" F. In particular, in the relative approach the poverty line adapts to growth/decline of the economy. To make this formal and to give an example, the following poverty line functional combines both approaches (cf. Kakwani (1986) and Lambert (2001), p.139) in taking a convex combination of a fixed amount z_0 and a centrality measure of the underlying income distribution:

$$\mathsf{z}_{\mathsf{m}, z_0, \delta}(F) = z_0 + \delta(\mathsf{m}(F) - z_0) \tag{31}$$

[3]Note that for relative inequality measures I_{rel} defined on $D_{cdf}([a, b])$ for $0 \leq a < b$ it is typically the case that $\mathsf{I}_{rel}(F) \in [0, 1]$, and hence $1 - \mathsf{I}_{rel}(F) \in [0, 1]$ as well.

where $z_0 > 0$, $0 \leq \delta \leq 1$, and m is a location functional that either coincides with the mean functional μ, or the median functional $q_{1/2}$. Note in particular that $\mathsf{z}_{m,z_0,0} = z_0$, i.e., this definition nests both an absolute and a relative approach. For concreteness, Lemma B.1 in Appendix B summarizes conditions under which the poverty line functionals in the family (31) satisfy Assumption 2.1.

The first poverty measure we shall consider is the so-called *headcount ratio*, which is the proportion in a population F that qualifies as poor according to a given poverty line z:

$$\mathsf{H}_{\mathsf{z}}(F) = F(\mathsf{z}(F)). \tag{32}$$

For the sake of generality, the following lemma establishes conditions under which the headcount ratio satisfies Assumption 2.1 under high-level conditions concerning the poverty line functional z. Specific constants and domains for the concrete family of poverty lines defined in Equation (31) can immediately be obtained with Lemma B.1 in Appendix B. An analogous statement applies to the poverty measures introduced further below, and will not be restated.

Lemma A.9. *Let $a = 0 < b$, and let $\mathsf{z} : D_{cdf}([a, b]) \to \mathbb{R}$ denote a poverty line functional that satisfies Assumption 2.1 with $\mathscr{D}_{\mathsf{z}}$ and constant C_{z}, say. Let $s > 0$. Then, H_{z} satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}_{\mathsf{z}} \cap \mathscr{D}^s([a, b])$ and $C = C_{\mathsf{z}}s + 1$.*

Certain disadvantages of the headcount ratio motivated Sen (1976) to introduce a different family of poverty measures, using an axiomatic approach. We shall now discuss this family in the generalized form of Kakwani (1980). Given a poverty line z and a "sensitivity parameter" $\kappa \geq 1$, say, each element of this family of poverty indices is written as

$$\mathsf{P}_{SK}(F; \mathsf{z}, \kappa) = (\kappa + 1) \int_{[0,\mathsf{z}(F)]} \left[1 - \frac{x}{\mathsf{z}(F)}\right] \left[1 - \frac{F(x)}{F(\mathsf{z}(F))}\right]^{\kappa} dF(x), \tag{33}$$

with the convention that $0/0 := 0$. A result discussing conditions under which $\mathsf{P}_{SK}(F; \mathsf{z}, \kappa)$ satisfies Assumption 2.1, and which is again established under high-level assumptions on the poverty line z, is provided next.

Lemma A.10. *Let $a = 0 < b$, $\kappa \geq 1$, and let $\mathsf{z} : D_{cdf}([a, b]) \to \mathbb{R}$ denote a poverty line functional that satisfies Assumption 2.1 with $\mathscr{D}_{\mathsf{z}}$ and constant C_{z}, say. Suppose further that $\mathsf{z} \geq z_* > 0$ holds for some real number z_*. Let $s > 0$. Then $T = \mathsf{P}_{SK}(.; \mathsf{z}, \kappa)$ satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}_{\mathsf{z}} \cap \mathscr{D}^s([a, b])$ and $C = (\kappa + 1)[1 + bz_*^{-2}C_{\mathsf{z}} + \kappa[(2 + s)C_{\mathsf{z}} + 4]]$.*

Second, we consider a family, each element of which can be written as

$$\mathsf{P}_{FGT}(F; \mathsf{z}, \Lambda) = \int_{[0,\mathsf{z}(F)]} \Lambda\left(1 - [x/\mathsf{z}(F)]\right) dF(x), \tag{34}$$

where $\Lambda : [0, 1] \to [0, 1]$ is non-decreasing, surjective and convex. This class contains (at least after monotonic transformations), e.g., the measures of Foster et al. (1984) or Chakravarty (1983) as special cases (cf. Lambert (2001) Chapter 6.3, and also the more recent review in Foster et al. (2010)). The following result provides conditions under which P_{FGT} satisfies Assumption 2.1. Again the result is established under high-level assumptions on the poverty line z (note in particular that the poverty line in Equation (31) is greater or equal to $(1 - \delta)z_0 > 0$ in case F is supported on $[0, \infty)$).

Lemma A.11. *Let $a = 0 < b$ and let $\mathsf{z} : D_{cdf}([a,b]) \to \mathbb{R}$ denote a poverty line functional that satisfies Assumption 2.1 with $\mathscr{D}_{\mathsf{z}}$ and constant C_{z}, say. Suppose further that $\mathsf{z} \geq z_* > 0$ holds for some real number z_*, that $\Lambda : [0,1] \to [0,1]$ is non-decreasing and surjective, and that Λ is the restriction of a convex real-valued function Λ^* defined on an open interval in $\mathbb{R}$ containing $[0,1]$. Denote the Lipschitz constant of Λ by C_Λ. Then, $T = \mathsf{P}_{FGT}(.;\mathsf{z},\Lambda)$ satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}_{\mathsf{z}}$ and $C = bz_*^{-2}C_\Lambda C_{\mathsf{z}} + 1$.*

As a direct application of Lemma A.11, we note that given a poverty line z the poverty measure of Foster et al. (1984) is obtained upon setting $\Lambda(x) = x^\alpha$ in Equation (34). The conditions in the preceding lemma are satisfied for $\alpha \geq 1$ (in which case $C_\Lambda = \alpha$). The preceding lemma does not cover the range $\alpha \in [0,1)$. Note that the functional corresponding to $\Lambda(x) = x^\alpha$ with $\alpha = 0$ coincides with the headcount ratio, which is already covered via Lemma A.9. Furthermore, the range $\alpha > 1$ might be considered most important, as only such values of α guarantee that in addition to the "Monotonicity Axiom" ("Given other things, a reduction in the income of a poor household must increase the poverty measure"), which would be satisfied for all $\alpha \geq 0$, the inequality measure obtained also satisfies the "Transfer Axiom" ("Given other things, a pure transfer of income from a poor household to any other household that is richer must increase the poverty measure") of Sen (1976), cf. Proposition 1 in Foster et al. (2010). Both axioms are plausible requirements (albeit not undisputed, an early criticism being Kundu and Smith (1983)), but are not satisfied by the headcount ratio, cf. Sen (1976).

B Proofs for Section A

Proof of Lemma A.1: Given $F, G \in D_{cdf}([a,b])$ it holds that $|\mathsf{S}_{abs}(F) - \mathsf{S}_{abs}(G)|$ is not greater than 1/2-times

$$\int_{[a,b]} \big||x-\mu(F)| - |x-\mu(G)|\big|\, dF(x) + \left|\int_{[a,b]} |x-\mu(G)| dF(x) - \int_{[a,b]} |x-\mu(G)| dG(x)\right|.$$

Using the reverse triangle inequality, the first integral in the previous display can be bounded from above by $|\mu(F)-\mu(G)| \leq (b-a)\|F-G\|_\infty$ (cf. Example C.3 for the inequality). Using Lemma C.2, the remaining expression to the right in the previous display is seen not to be greater than $(b-a)\|F-G\|_\infty$. Hence, the first statement follows (noting that S_{abs} is obviously well defined on all of $D_{cdf}([a,b])$).

Concerning the second claim, we first observe that for every $F \in D_{cdf}([a,b])$ it holds that

$$\frac{1}{2}\int_{[a,b]} |x-\mu(F)| dF(x) = \int_{[a,\mu(F)]} (\mu(F)-x) dF(x). \tag{35}$$

Next, let $s > 0$, $\delta \in (a,b)$, $F \in \mathscr{D}(s,\delta)$ and $G \in D_{cdf}([a,b])$. We consider two cases, and start with the case where $\mu(G) = 0$ (implying that $a = 0$ and that G is the cdf corresponding to point mass at 0). Then, by convention, $\mathsf{S}_{rel}(G) = 0$, and it follows from Equation (35) (recalling that $\mu(F) \geq \delta > 0$) that

$$|\mathsf{S}_{rel}(F) - \mathsf{S}_{rel}(G)| \leq \int_{[a,\mu(F)]} |1 - x/\mu(F)| dF(x) \leq F(\mu(F)). \tag{36}$$

Since F is continuous $0 = F(0) = F(\mu(G))$ holds. It follows that $F(\mu(F)) = |F(\mu(F)) - F(\mu(G))|$. Using the mean value theorem of Minassian (2007) and Example C.3 we conclude that $|F(\mu(F)) - F(\mu(G))| \leq s(b-a)\|F - G\|_\infty$.

Next, we turn to the case where $\mu(G) > 0$. First, we note that

$$|\mathsf{S}_{rel}(F) - \mathsf{S}_{rel}(G)| \leq |F(\mu(F)) - G(\mu(G))| + \left| \int_{[a,\mu(F)]} \frac{x}{\mu(F)} dF(x) - \int_{[a,\mu(G)]} \frac{x}{\mu(G)} dG(x) \right|.$$

Consider the first term in absolute values in the previous display: by the triangle inequality:

$$|F(\mu(F)) - G(\mu(G))| \leq |F(\mu(F)) - F(\mu(G))| + \|F - G\|_\infty.$$

From the mean value theorem for right-differentiable functions as in Minassian (2007), and the definition of $\mathscr{D}^s([a,b])$, we obtain $|F(\mu(F)) - F(\mu(G))| \leq s|\mu(F) - \mu(G)| \leq s(b-a)\|F - G\|_\infty$, the second inequality following from Example C.3. Now, note that (incorporating the considerations in case $\mu(G) = 0$) it remains to show that

$$\left| \int_{[a,\mu(F)]} \frac{x}{\mu(F)} dF(x) - \int_{[a,\mu(G)]} \frac{x}{\mu(G)} dG(x) \right| \leq ((s + \delta^{-1})(b-a) + 4)\|F - G\|_\infty. \tag{37}$$

To this end, denote $m := \min(\mu(F), \mu(G))$, $M := \max(\mu(F), \mu(G))$, let $\tilde{F}$ denote a cdf in $\{F, G\}$ which realizes the latter maximum, and rewrite the difference of integrals inside the absolute value to the left in the preceding display as

$$\int_{[a,m]} \frac{x}{\mu(F)} dF(x) - \int_{[a,m]} \frac{x}{\mu(F)} dG(x) \pm \int_{(m,M]} \frac{x}{\mu(\tilde{F})} d\tilde{F}(x) + \int_{[a,\mu(G)]} \left[\frac{x}{\mu(F)} - \frac{x}{\mu(G)} \right] dG(x),$$

where the $\pm$ is "+" in case $\tilde{F} = F$ and "−" in case $\tilde{F} = G$. Next, denote the difference of the first two integrals in the previous display by A, the third integral by B and the fourth by C, respectively. First, Lemma C.2 (applied with $k = 1$, $c = a$, $d = m$ and $\varphi(x) = x/\mu(F)$) implies (working with the upper bounds $|M^*| \leq 1$ and $C \leq 1$ in Lemma C.2 for the special case under consideration) that $|A| \leq 2\|F - G\|_\infty$. Second, note that the integrand in B is smaller than 1, hence

$$|B| \leq \tilde{F}(M) - \tilde{F}(m) \leq F(M) - F(m) + 2\|F - G\|_\infty \leq s|\mu(F) - \mu(G)| + 2\|F - G\|_\infty \tag{38}$$

where we used $\|\tilde{F} - F\| \leq \|F - G\|_\infty$ for the first inequality, and the mean value theorem of Minassian (2007) for the second. To obtain an upper bound for $|B|$ we now use Example C.3 to see that the right hand side in the previous display is not greater than $[s(b-a) + 2]\|F - G\|_\infty$. Concerning $|C|$ note that

$$|C| \leq \int_{[a,\mu(G)]} \left| \frac{\mu(G)}{\mu(F)} - 1 \right| dG(x) \leq \left| \frac{\mu(G)}{\mu(F)} - 1 \right|.$$

If $\mu(G)/\mu(F) \geq 1$, then the upper bound in the previous display is not greater (cf. Example C.3) than

$$\left[\mu(F) + (b-a)\|F - G\|_\infty\right] / \mu(F) - 1 \leq \delta^{-1}(b-a)\|F - G\|_\infty,$$

and the same bound holds if $\mu(G)/\mu(F) < 1$. Hence, $|C| \leq \delta^{-1}(b-a)\|F - G\|_\infty$. Summarizing,

$$|A| + |B| + |C| \leq ((s + \delta^{-1})(b-a) + 4)\|F - G\|_\infty,$$

which proves the statement in Equation (37). □

Proof of Lemma A.2: The first statement follows from Example C.6. To prove the statement concerning G_{rel}, we first note that G_{rel} is well defined on $D_{cdf}([a,b])$ (note that $\mu(F) \leq 0$ implies that $a = 0$ and that μ_F is point mass at 0, implying that $\mathsf{G}_{rel}(F) = 0$). Let $F, G \in D_{cdf}([a,b])$. We now consider two cases, and start with the case where $\mu(G) = 0$. Then, $\mathsf{G}_{rel}(G) = 0$ and

$$|\mathsf{G}_{rel}(F) - \mathsf{G}_{rel}(G)| = \mathsf{G}_{rel}(F) \leq \delta^{-1}[\mu(F) - \mu(G)] \leq \delta^{-1}(b-a)\|F - G\|_{\infty}, \tag{39}$$

where we used Example C.3 in the last inequality.

Next, in case $\mu(G) \neq 0$, we have $\mu(G) > 0$ (recall that $a \geq 0$), and we set $\varphi(x_1, x_2) = |x_1 - x_2|$. Write

$$|\mathsf{G}_{rel}(F) - \mathsf{G}_{rel}(G)| \leq A + B, \tag{40}$$

where

$$A := \delta^{-1}\left[\int_{[a,b]}\int_{[a,b]} \varphi(x_1,x_2)dF(x_1)dF(x_2) - \int_{[a,b]}\int_{[a,b]} \varphi(x_1,x_2)dG(x_1)dG(x_2)\right], \tag{41}$$

which, by Example C.6 is not greater than $\delta^{-1}2(b-a)\|F - G\|_{\infty}$, and

$$B := \int_{[a,b]}\int_{[a,b]} |(\mu(F)^{-1} - \mu(G)^{-1})\varphi(x_1,x_2)|dG(x_1)dG(x_2). \tag{42}$$

By the reverse triangle inequality (using that $\mu(F) > 0$ and $\mu(G) > 0$), we see that

$$B \leq |\mu(F)^{-1} - \mu(G)^{-1}| \int_{[a,b]}\int_{[a,b]} |x_1 - x_2| dG(x_1)dG(x_2) \leq 2\left|[\mu(G)/\mu(F)] - 1\right|. \tag{43}$$

Using Example C.3, we see that $\mu(F) - (b-a)\|F-G\|_{\infty} \leq \mu(G) \leq \mu(F) + (b-a)\|F-G\|_{\infty}$, from which it is easy to conclude that

$$-(b-a)\|F-G\|_{\infty}/\mu(F) \leq [\mu(G)/\mu(F)] - 1 \leq (b-a)\|F-G\|_{\infty}/\mu(F), \tag{44}$$

from which it follows that $\left|[\mu(G)/\mu(F)] - 1\right| \leq \delta^{-1}(b-a)\|F-G\|_{\infty}$. Hence, in case $\mu(G) \neq 0$, we obtain that

$$|\mathsf{G}_{rel}(F) - \mathsf{G}_{rel}(G)| \leq 4\delta^{-1}(b-a)\|F-G\|_{\infty}. \tag{45}$$

Together with the first case, this proves the result.

□

Proof of Lemma A.3: First of all, the functional under consideration is trivially well defined on $D_{cdf}([a,b])$ (because $a > 0$ holds). Next, we apply Lemma C.12 together with Lemma C.14 to obtain that for every $u \in [0,1]$ the functional $F \mapsto L(F,u)$ satisfies Assumption 2.1 with a, b and $\mathscr{D}$ (as in the statement of the present lemma) as in the statement of the theorem and with constant $a^{-1}(r^{-1} + (b-a)a^{-1})$. The statement immediately follows. □

Proof of Lemma A.4: The triangle inequality, together with Example C.3 and Lemma C.13 (which is applicable due to Lemma C.12) immediately yield the claimed result. □

Proof of Lemma A.5: We start with the first statement. The functional T is obviously everywhere defined on $D_{cdf}([a,b])$ (in case $\mu(F)=0$ it follows that $a=0$ and that F corresponds to point mass 1 at 0 in which case $T(F)=0$, by definition). Next, let $\delta \in (a,b)$, let $F \in \mathscr{D}(\delta)$ and let $G \in D_{cdf}([a,b])$. We consider first the case where $\mu(G)=0$ (implying that $\mathsf{T}(G)=0$). Then, noting that $\int_{[a,b]} x^c dG(x)=0$, we conclude that $|\mathsf{T}(F)-\mathsf{T}(G)|=\mathsf{T}(F)$, the latter being not greater than

$$\frac{1}{c|c-1|\delta^c}\left|\int_{[a,b]} x^c dF(x) - \int_{[a,b]} x^c dG(x)\right| \leq \frac{b^c-a^c}{c|c-1|\delta^c}\|F-G\|_\infty, \tag{46}$$

where we used Lemma C.2 for the last inequality. Next, consider the case where $\mu(G)>0$. We note that

$$\left|\int_{[a,b]} (x/\mu(F))^c dF(x) - \int_{[a,b]} (x/\mu(F))^c dG(x)\right| \tag{47}$$

can be upper bounded by $A+B$ with

$$A := \left|\int_{[a,b]} (x/\mu(F))^c dF(x) - \int_{[a,b]} (x/\mu(F))^c dG(x)\right| \leq \frac{b^c-a^c}{\delta^c}\|F-G\|_\infty \tag{48}$$

the inequality following from Lemma C.2, and

$$B := |(1/\mu(F))^c - (1/\mu(G))^c| \int_{[a,b]} x^c dG(x) \leq |(\mu(G)/\mu(F))^c - 1|, \tag{49}$$

the inequality following from Jensen's inequality (recalling that $c \in (0,1)$). It remains to observe that the simple inequality $|z^c-1| \leq |x-1|$ for $z>0$ it follows that

$$|(\mu(G)/\mu(F))^c - 1| \leq |\mu(G)/\mu(F) - 1| \leq \delta^{-1}(b-a)\|F-G\|_\infty, \tag{50}$$

where the second inequality follows from Example C.3 together with $\mu(F) \geq \delta$. Hence, in case $\mu(G)>0$ we see that

$$|\mathsf{T}(F)-\mathsf{T}(G)| \leq (c|c-1|)^{-1}\left[\frac{b^c-a^c}{\delta^c} + \delta^{-1}(b-a)\right]\|F-G\|_\infty,$$

which proves the first claim. We now prove the second claim. Since $a>0$ holds in this case, $\mu(G)$ and $\mu(F)$ can not be smaller than a. Hence the functional is well defined on all of $D_{cdf}([a,b])$. Furthermore, the expression in Equation (47) is greater than $A+B$, where A and B have been defined above. By Lemma C.2 it holds that A is not greater than $a^{-c}|b^c-a^c|\|F-G\|_\infty$. Furthermore, B is not greater than

$$\begin{aligned} \max\left((a/b)^c,(b/a)^c\right)|(\mu(F)/\mu(G))^c-1| &\leq |c|\max\left((a/b)^{2c-1},(b/a)^{2c-1}\right)|\mu(F)/\mu(G)-1| \\ &\leq |c|\max\left((a/b)^{2c-1},(b/a)^{2c-1}\right)a^{-1}(b-a)\|F-G\|_\infty, \end{aligned} \tag{51}$$

the first inequality following from $|z^c-1| \leq |c|\max((a/b)^{c-1},(b/a)^{c-1})|z-1|$ for $z \in [a/b, b/a]$ (noting that this interval contains 1 and recalling that $c \notin [0,1]$), and the second inequality following from Exercise C.3. We now turn to the last case where $c \in \{0,1\}$ (and $a>0$ guaranteeing that the functional is then well defined on all of $D_{cdf}([a,b])$). We consider first the case where $c = 0$. The statement follows after noting that $|\mathsf{T}(F) - \mathsf{T}(G)|$ is not greater than $C + D$ with

$$C := \left| \int_{[a,b]} \log(x) dF(x) - \log(x) dG(x) \right| \leq \log(b/a) \|F - G\|_\infty \tag{52}$$

the inequality following from Lemma C.2, and

$$\begin{aligned} D := \int_{[a,b]} \left|\log(x/\mu(F)) - \log(x/\mu(G))\right| dG(x) = |\log(\mu(G)/\mu(F))| &\leq \log(1 + a^{-1}\|F - G\|_\infty) \\ &\leq a^{-1}\|F - G\|_\infty. \end{aligned} \tag{53}$$

In case $c = 1$, let $f(x) = (x/\mu(F)) \log(x/\mu(F))$ and $g(x) = (x/\mu(G)) \log(x/\mu(G))$. Write

$$|\mathsf{T}(F) - \mathsf{T}(G)| \leq \left| \int_{[a,b]} f(x) dF(x) - \int_{[a,b]} f(x) dG(x) \right| + \int_{[a,b]} |f(x) - g(x)| dG(x). \tag{54}$$

From Lemma C.2 it follows that the first absolute value in the upper bound is not greater than $a^{-1} \int_{[a,b]} |1 + \log(x)| dx \|F - G\|_\infty$. Finally, noting that for every $x \in [a,b]$ we have

$$\begin{aligned} |f(x) - g(x)| &= |x| \left\{ |\mu^{-1}(F) - \mu^{-1}(G)| |\log(x/\mu(F))| + \mu^{-1}(G) \left|\log(\mu(F)/\mu(G))\right| \right\} \\ &\leq ba^{-1} \left\{ |\log(b/a)|(b-a) + a^{-1} \right\} \|F - G\|_\infty, \end{aligned} \tag{55}$$

where (in addition to $a > 0$) we used Exercise C.3 and the inequality for $\log(\mu(G)/\mu(F))$ established above. The final claim follows. □

Proof of Lemma A.6: We start with Part 1: Let $\delta \in (a,b)$. From the first part of Lemma A.5 we obtain that in case $\varepsilon \in (0,1)$ the functional $\mathsf{E}_{c(\varepsilon)}$ satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{D}(\delta)$ (cf. Equation 21) and constant $|\varepsilon c(\varepsilon)|^{-1} \left[\delta^{-c(\varepsilon)} (b^{c(\varepsilon)} - a^{c(\varepsilon)}) + \delta^{-1}(b - a) \right]$. It remains to observe that the function

$$z \mapsto 1 - z^{1/c(\varepsilon)} \tag{56}$$

is Lipschitz continuous on $[0,1]$ with constant $c(\varepsilon)^{-1}$. The claim then follows from Lemma C.1 (with $m = 1$), the representation in Equation (26) together with the observation that $0 \leq \varepsilon(\varepsilon - 1)\mathsf{E}_{c(\varepsilon)}(F) + 1 \leq 1$ holds for every $F \in D_{cdf}([a,b])$ as a consequence of Jensen's inequality.

Concerning Part 2: From the second part of Lemma A.5 we obtain that in case $\varepsilon \in (1, \infty)$ the functional $\mathsf{E}_{c(\varepsilon)}$ satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a,b])$ and constant (note that $2c(\varepsilon) - 1 < 0$) equal to

$$|c(\varepsilon)\varepsilon|^{-1} [a^{-c(\varepsilon)} |b^{c(\varepsilon)} - a^{c(\varepsilon)}| + |c(\varepsilon)| (a/b)^{2c(\varepsilon)-1} a^{-1} (b-a)].$$

We shall now argue similarly as in Part 1. The function in Equation (56) is Lipschitz continuous on $[(b/a)^{c(\varepsilon)}, (a/b)^{c(\varepsilon)}]$ with constant $c(\varepsilon)^{-1}(b/a)^{\varepsilon}$. From Equation (26) and because $(b/a)^{c(\varepsilon)} \leq \varepsilon(\varepsilon - 1)\mathsf{E}_{1-\varepsilon}(F) + 1 \leq (a/b)^{c(\varepsilon)}$ holds for every $F \in D_{cdf}([a,b])$ ($a > 0$ and $c(\varepsilon) < 0$) the claim follows from Lemma C.1 (with $m = 1$). □

Proof of Lemma A.7: Clearly T is well defined on all of $D_{cdf}([a,b])$. Let $F \in D_{cdf}([a,b])$. Then, by Jensen's inequality:

$$\int_{[a,b]} e^{\kappa[\mu(F)-x]} dF(x) \geq 1. \tag{57}$$

Since log restricted to $[1,\infty)$ is Lipschitz continuous with constant 1, we obtain for any $G \in D_{cdf}([a,b])$ that

$$|\mathsf{T}(F) - \mathsf{T}(G)| \leq \kappa^{-1} \left| \int_{[a,b]} e^{\kappa[\mu(F)-x]} dF(x) - \int_{[a,b]} e^{\kappa[\mu(G)-x]} dG(x) \right|. \tag{58}$$

Since the term in absolute values in the previous display is not greater than

$$\left| e^{\kappa\mu(F)} - e^{\kappa\mu(G)} \right| + \left| \int_{[a,b]} e^{-\kappa x} dF(x) - \int_{[a,b]} e^{-\kappa x} dG(x) \right| \tag{59}$$

which can be bounded from above by

$$\kappa e^{\kappa b}(b-a)\|F-G\|_\infty + [e^{-\kappa a} - e^{-\kappa b}]\|F-G\|_\infty, \tag{60}$$

where we used Example C.3 and Lemma C.2 $\square$

Proof of Lemma A.8: Obviously, the welfare function W is well defined on $D_{cdf}([a,b])$ in both parts of the lemma. The first statement in the present lemma follows from the assumptions and Example C.3, noting that $x_1x_2 - y_1y_2 = (x_1-y_1)x_2 - y_1(y_2-x_2)$ holds for real numbers x_i, y_i, $i=1,2$. The second statement follows directly from the assumptions and Example C.3. $\square$

Lemma B.1. *Let $a < b$ be real numbers, $z_0 > 0$ and $0 \leq \delta \leq 1$. Then, the following holds:*

1. *If $\delta = 0$, then $\mathsf{z}_{\mathsf{m},z_0,\delta}$ satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a,b])$, and $C = 0$.*
2. *If $\delta > 0$ and $\mathsf{m} = \mu(.)$, then $\mathsf{z}_{\mathsf{m},z_0,\delta}$ satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a,b])$ and $C = \delta(b-a)$.*
3. *If $\delta > 0$ and $\mathsf{m} = q_{1/2}(.)$, then, for every $r > 0$ the poverty line $\mathsf{z}_{\mathsf{m},z_0,\delta}$ satisfies Assumption 2.1 with $\mathscr{D} = \mathscr{C}_r([a,b])$, and $C = r^{-1}\delta$.*

Proof of Lemma B.1: Recall from Equation 31 that by definition $\mathsf{z}_{\mathsf{m},z_0,\delta}(F) = z_0 + \delta(\mathsf{m}(F) - z_0)$. The first statement is trivial; the second follows directly from Example C.3; and the third follows from Lemma C.12 and Example C.10. $\square$

Proof of Lemma A.9: Since z satisfies Assumption 2.1 the functional z is well defined on $D_{cdf}([a,b])$. Thus H_z is well defined on $D_{cdf}([a,b])$ as well. Finally, given $F \in \mathscr{D}$ and $G \in D_{cdf}([a,b])$, note that by definition and the triangle inequality:

$$|\mathsf{H}_\mathsf{z}(F) - \mathsf{H}_\mathsf{z}(G)| \leq |F(\mathsf{z}(F)) - F(\mathsf{z}(G))| + \|F-G\|_\infty \leq (C_\mathsf{z}s + 1)\|F-G\|_\infty, \tag{61}$$

where we used that z satisfies Assumption 2.1 together with a mean-value theorem as in Minassian (2007) for the last inequality. $\square$

Proof of Lemma A.10: Obviously, $\mathsf{P}_{SK}(.;\mathsf{z},\kappa)$ is well defined on $D_{cdf}([a,b])$ because $\mathsf{z} \geq z_* > 0$ holds by assumption, and due to our convention that $0/0 := 0$ (noting also that $F(x) = 0$ for every $x \in [0, \mathsf{z}(F)]$ in case $F(\mathsf{z}(F)) = 0$). Next, fix $F \in \mathscr{D}$ and $G \in D_{cdf}([a,b])$. Define for all $x \in \mathbb{R}$

$$f(x) := \max(1 - [x/\mathsf{z}(F)], 0)[1 - [F(x)/F(\mathsf{z}(F))]]^{\kappa},$$

and analogously

$$g(x) := \max(1 - [x/\mathsf{z}(G)], 0)[1 - [G(x)/G(\mathsf{z}(G))]]^{\kappa}.$$

Define $m := \min(\mathsf{z}(F), \mathsf{z}(G))$ and $M := \max(\mathsf{z}(F), \mathsf{z}(G))$, and the following partition of $[a, M]$ (using our convention $0/0 := 0$):

$$A := \left\{ x \in [a,m] : \frac{F(x)}{F(\mathsf{z}(F))} > \frac{G(x)}{G(\mathsf{z}(G))} \right\}, \quad B := [a,m] \backslash A, \quad \text{and } C := (m, M],$$

where $C = \emptyset$ in case $m = M$. Next, write

$$\frac{\mathsf{P}_{SK}(F;\mathsf{z},\kappa) - \mathsf{P}_{SK}(G;\mathsf{z},\kappa)}{\kappa + 1} = \int_{[a,M]} [f(x) - g(x)] dF(x) + \left[\int_{[a,b]} g(x) F(x) - \int_{[a,b]} g(x) G(x) \right],$$

noting that f and g vanish for $x > M$; and denote the right-hand side by $S_1 + S_2$, S_2 denoting the term in brackets to the far right. Since $g([a,b]) \subseteq [0,1]$ and because g is right-continuous (G is a cdf) and non-increasing, it hence follows from Lemma C.7 that $|S_2| \leq \|F - G\|_\infty$. Concerning S_1, note that for every $x \in [a, M]$ it holds that $|f(x) - g(x)|$ is not greater than the sum of

$$\begin{aligned} |\max([1 - \frac{x}{\mathsf{z}(F)}], 0) - \max([1 - \frac{x}{\mathsf{z}(G)}], 0)| &\leq x|\mathsf{z}(F)^{-1} + \mathsf{z}(G)^{-1}| \\ &\leq x z_*^{-2} |\mathsf{z}(F) + \mathsf{z}(G)| \leq b z_*^{-2} C_{\mathsf{z}} \|F - G\|_\infty, \end{aligned} \tag{62}$$

(where we used that $\mathsf{z} \geq z_0$ to obtain the second inequality, and that z satisfies Assumption 2.1 to obtain the third) and

$$\big| |1 - [F(x)/F(\mathsf{z}(F))]|^{\kappa} - |1 - [G(x)/G(\mathsf{z}(G))]|^{\kappa} \big| \leq \kappa |[F(x)/F(\mathsf{z}(F))] - [G(x)/G(\mathsf{z}(G))]| \tag{63}$$

(where we used $\kappa \geq 1$, the reverse triangle inequality, and Lemma A.9 to obtain the upper bound). It hence follows that $|S_1|$ is bounded from above by the sum of $b z_*^{-2} C_{\mathsf{z}} \|F - G\|_\infty$ and κ times

$$\begin{aligned} \int_A [F(x)/F(\mathsf{z}(F))] - [G(x)/G(\mathsf{z}(G))] dF(x) + \int_B [G(x)/G(\mathsf{z}(G))] - [F(x)/F(\mathsf{z}(F))] dF(x) \\ + \int_C h(x) dF(x), \end{aligned} \tag{64}$$

where for every $x \in C$ we define

$$h(x) := \begin{cases} |F(x)/F(\mathsf{z}(F))| & \text{if } m < M = \mathsf{z}(F) \\ |G(x)/G(\mathsf{z}(G))| & \text{if } m < M = \mathsf{z}(G) \\ 0 & \text{if } m = M. \end{cases} \tag{65}$$

Using $0 \leq h \leq 1$, the mean value theorem in Minassian (2007), and that z satisfies Assumption 2.1, we see that the third integral in Equation (64) satisfies

$$\int_C h(x)dF(x) \leq F(M) - F(m) \leq sC_{\mathsf{z}}\|F - G\|_\infty. \tag{66}$$

To bound the expression in Equation (64), we also recall that Lemma A.9 shows that

$$G(\mathsf{z}(G)) - (C_{\mathsf{z}} + 1)\|F - G\|_\infty \leq F(\mathsf{z}(F)) \leq G(\mathsf{z}(G)) + (C_{\mathsf{z}} + 1)\|F - G\|_\infty. \tag{67}$$

We consider different cases:

Consider first the case where $F(\mathsf{z}(F)) = 0$: Then, the convention $0/0 := 0$ implies $A = \emptyset$ and $B = [a, m]$. Furthermore, the integral over B in Equation (64) vanishes in this case, because $m \leq \mathsf{z}(F)$ implies $F(m) = 0$. By Equation (66), in this case it holds that the expression in Equation (64) does not exceed $sC_{\mathsf{z}}\|F - G\|_\infty$.

Next, consider the case where $G(\mathsf{z}(G)) = 0$ and $F(\mathsf{z}(F)) > 0$. It follows from our convention that then $A = \{x \in [a, m] : F(x)/F(\mathsf{z}(F)) > 0\}$, and that the integral over B in (64) vanishes. The integral over A is not greater than

$$F(m) = F(m) - G(\mathsf{z}(G)) \leq F(\mathsf{z}(F)) - G(\mathsf{z}(G)) \leq (C_{\mathsf{z}} + 1)\|F - G\|_\infty,$$

where we used Equation (67) to obtain the last inequality. Together with Equation (66) we thus see that in this case the expression in Equation (64) does not exceed $(C_{\mathsf{z}}(1 + s) + 1)\|F - G\|_\infty$.

Finally, consider the case where $G(\mathsf{z}(G))$ and $F(\mathsf{z}(F))$ are both positive. Then, we can write the integral over A in Equation (64) as

$$\begin{aligned}
& F(\mathsf{z}(F))^{-1} \int_A F(x) - F(\mathsf{z}(F)) \frac{G(x)}{G(\mathsf{z}(G))} dF(x) \\
\leq \quad & F(\mathsf{z}(F))^{-1} \int_A [F(x) - G(x)] + (C_{\mathsf{z}} + 1)\|F - G\|_\infty \frac{G(x)}{G(\mathsf{z}(G))} dF(x) \\
\leq \quad & \|F - G\|_\infty (C_{\mathsf{z}} + 2),
\end{aligned}$$

where we used Equation (67) to obtain the first inequality. Similarly, the integral over B in Equation (64) can be shown not to be greater than $\|F - G\|_\infty (C_{\mathsf{z}} + 2)$. Summarizing, in this last case the expression in Equation (64) does not exceed $[(2 + s)C_{\mathsf{z}} + 4]\|F - G\|_\infty$. In particular, this bound is bigger than the two bounds in the other two cases. Hence, we conclude that the expression in Equation (64) is not greater than $[(2 + s)C_{\mathsf{z}} + 4]\|F - G\|_\infty$ in general.

It follows that $|S_1|$ is bounded from above by

$$[bz_*^{-2}C_{\mathsf{z}} + \kappa[(2 + s)C_{\mathsf{z}} + 4]]\|F - G\|_\infty. \tag{68}$$

Since $|S_2| \leq \|F - G\|_\infty$, it follows that

$$\frac{\mathsf{P}_{SK}(F; \mathsf{z}, \kappa) - \mathsf{P}_{SK}(G; \mathsf{z}, \kappa)}{\kappa + 1} \leq |S_1| + |S_2| \leq [1 + bz_*^{-2}C_{\mathsf{z}} + \kappa[(2 + s)C_{\mathsf{z}} + 4]]\|F - G\|_\infty.$$

□

Proof of Lemma A.11: Obviously, $\mathsf{P}_{FGT}(.;\mathsf{z},\Lambda)$ is well defined on $D_{cdf}([a,b])$ because $\mathsf{z} \geq z_* > 0$ holds by assumption. Next, fix $F \in \mathscr{D}$ and $G \in D_{cdf}([a,b])$. Note that we can write

$$\mathsf{P}_{FGT}(F;\mathsf{z},\Lambda) = \int_{[a,b]} \Lambda(\max(1-[x/\mathsf{z}(F)],0))dF(x). \tag{69}$$

Abbreviating $f(x) = \Lambda(\max(1-[x/\mathsf{z}(F)],0))$ and $g(x) = \Lambda(\max(1-[x/\mathsf{z}(G)],0))$ we obtain

$$\mathsf{P}_{FGT}(F;\mathsf{z},\Lambda) - \mathsf{P}_{FGT}(G;\mathsf{z},\Lambda) = \int_{[a,b]} [f(x)-g(x)]dF(x) + \left[\int_{[a,b]} g(x)dF(x) - \int_{[a,b]} g(x)dG(x)\right].$$

Denote the first integral on the right by A, and the term in brackets to the far right by B. Noting that $g : [a,b] \to [0,1]$ is continuous (Λ is the restriction of a convex function) and non-increasing, we obtain from Lemma C.7 that $|B| \leq \|F-G\|_\infty$. Concerning A, we note that since Λ is the restriction of a convex function defined on an open interval containing $[0,1]$, Λ is Lipschitz continuous. Hence the constant C_Λ as claimed in the statement of the lemma indeed exists. In particular, using this property and the inequality $|\max(1-z_1,0) - \max(1-z_2,0)| \leq |z_1 - z_2|$ for nonnegative real numbers z_1, z_2, we can bound

$$|A| \leq C_\Lambda b |[1/\mathsf{z}(F)] - [1/\mathsf{z}(G)]| \leq C_\Lambda b z_*^{-2} |\mathsf{z}(F) - \mathsf{z}(G)| \leq C_\Lambda b z_*^{-2} C_{\mathsf{z}} \|F-G\|_\infty, \tag{70}$$

where we used Lipschitz continuity of the map $x \mapsto x^{-1}$ on $[z_*, \infty)$ (with constant z_*^{-2}), and the assumption that z satisfies Assumption 2.1 for obtaining the second inequality. Together with the upper bound on $|B|$ we obtain the claimed statement. □

C General results for establishing Assumption 2.1

We here *summarize* in a self-contained way some results that turn out to be useful for establishing Assumption 2.1 for empirically relevant functionals T. Specific examples were discussed in Appendix A and include inequality measures (cf. Appendix A.1), welfare measures (cf. Appendix A.2), and poverty measures (cf. Appendix A.3). The techniques we describe are based on decomposability-properties of the functional, its specific structural (e.g., linearity) properties, and on properties of quantiles and quantile functions, or related quantities such as Lorenz curves. We emphasize that *all of the results in the present section are fairly elementary, but are difficult to pinpoint in the literature in the form needed.* We first start with a short section concerning notation.

C.1 Notation

We denote by $D(\mathbb{R})$ the Banach space of real-valued bounded càdlàg functions equipped with the supremum norm $\|G\|_\infty = \sup\{|G(x)| : x \in \mathbb{R}\}$. The closed convex subset of $D(\mathbb{R})$ consisting of all cumulative distribution functions (cdfs) shall be denoted by $D_{cdf}(\mathbb{R})$. Furthermore, given two real numbers $a < b$, we define the subset $D_{cdf}((a,b])$ of $D_{cdf}(\mathbb{R})$ as follows: $F \in D_{cdf}((a,b])$ if and only if $F \in D_{cdf}(\mathbb{R})$, $F(a) = 0$ and $F(b) = 1$. Likewise, we define the subset $D_{cdf}([a,b])$ of $D_{cdf}(\mathbb{R})$ as follows: $F \in D_{cdf}([a,b])$ if and only if $F \in D_{cdf}(\mathbb{R})$, $F(a-) = 0$ and $F(b) = 1$. Here $F(a-)$ denotes the left-sided limit of F at a. Given a cdf F we denote by μ_F the (uniquely defined) probability measure on the Borel sets of $\mathbb{R}$ that satisfies

$$\mu_F((-\infty, x]) = F(x) \quad \text{for every } x \in \mathbb{R},$$

and, as usual, we denote the Lebesgue-Stieltjes integral of a μ_F-integrable function $f : \mathbb{R} \to \mathbb{R}$ by $\int_{\mathbb{R}} f(x)dF(x) := \int_{\mathbb{R}} f(x)d\mu_F(x)$.

In the following subsections we shall repeatedly encounter *functionals* T with domain $\mathscr{T} \subseteq D_{cdf}(\mathbb{R})$ and co-domain $\mathbb{R}$, which are Lipschitz continuous ($\mathscr{T}$ being equipped with the metric induced by the supremum norm on $D(\mathbb{R})$): recall that a functional $\mathsf{T} : \mathscr{T} \to \mathbb{R}$ is called *Lipschitz continuous* if there exists a nonnegative real number C so that for every F and $G \in \mathscr{T}$ it holds that

$$|\mathsf{T}(F) - \mathsf{T}(G)| \leq C\|F - G\|_\infty. \tag{71}$$

We then call C *a* Lipschitz constant of T. When we say that a functional T is Lipschitz continuous with constant C, we do not imply that this is the smallest such constant. Recall from Remark 2.2 that if a functional T is Lipschitz continuous on $\mathsf{T} = D_{cdf}([a, b])$ for real numbers $a < b$, then T satisfies Assumption 2.1 with $\mathscr{D} = D_{cdf}([a, b])$.

C.2 Decomposability

Oftentimes functionals can be decomposed into a function of several "simpler" functionals. It is a basic, but useful, fact that if a functional can be written as a composition of functionals that satisfy Assumption 2.1 with a Lipschitz continuous function on the intermediating Euclidean space, this composition satisfies Assumption 2.1 as well. The result is as follows, its proof is trivial and omitted.

Lemma C.1. *Let $a < b$ be real numbers, and let $\mathscr{D} \subseteq D_{cdf}([a, b])$. Suppose that T_i for $i = 1, \ldots, m$ satisfies Assumption 2.1 with a, b and $\mathscr{D}$ and with constant C_i, respectively. Let $\mathcal{I}_i := \mathsf{T}_i(D_{cdf}([a, b]))$ and set $\mathcal{I} = \bigtimes_{i=1}^m \mathcal{I}_i$. Assume $G : \mathcal{I} \to \mathbb{R}$ is Lipschitz continuous with constant C, when $\mathcal{I} \subseteq \mathbb{R}^m$ is equipped with the metric induced by the norm $\|.\|_1$ on $\mathbb{R}^m$. Then, $\mathsf{T} = G \circ (\mathsf{T}_1, \ldots, \mathsf{T}_m)$ satisfies Assumption 2.1 with a, b and $\mathscr{D}$ and with constant $C \sum_{i=1}^m C_i$.*

C.3 U-functionals

We here consider U-functionals (the corresponding sample plug-in variants being traditionally referred to as U-statistics, hence the name). The following result covers examples such as moments, certain concentration measures or dependence measures, cf. Chapter 5 in Serfling (2009), and see also the subsequent discussion for examples.

Lemma C.2. *Let $a < b$ be real numbers and let $\varphi : [a, b]^k \to \mathbb{R}$ for some $k \in \mathbb{N}$. Suppose that φ is bounded, and is symmetric in the sense that for all $x_i \in [a, b]$ for $1, \ldots, k$ it holds that $\varphi(x_1, \ldots, x_k) = \varphi(x_{\pi_1}, \ldots, x_{\pi_k})$ for every permutation $x_{\pi_1}, \ldots, x_{\pi_k}$ of $x_1, \ldots, x_k$. Let $a \leq c < d \leq b$. Suppose that for every $x_2^*, \ldots, x_k^* \in [c, d]^{k-1}$ the function $x \mapsto \varphi(x, x_2^*, \ldots, x_k^*)$ defined on $[c, d]$ is continuous and has* *total variation not greater than $C \in \mathbb{R}$. For $F \in D_{cdf}([a,b])$ define the functional $\mathsf{m}_{\varphi;c,d}$ as the iterated Lebesgue-Stieltjes integral*

$$\mathsf{m}_{\varphi;c,d}(F) := \int_{[c,d]} \cdots \int_{[c,d]} \varphi(x_1, \ldots, x_k) dF(x_1) \ldots dF(x_k), \tag{72}$$

where we write m_φ in case $c = a$ and $d = b$. Then, $\mathsf{m}_{\varphi;c,d}$ is Lipschitz continuous on $D_{cdf}([a,b])$ with constant kC^*, where

$$C^* = \begin{cases} C & \text{if } a = c, b = d \\ C + m^* & \text{if } b = d \\ C + M^* & \text{if } a = c \\ C + m^* + M^* & \text{else,} \end{cases} \tag{73}$$

and where

$$m^* := \sup\{|\varphi(c, x_2^*, \ldots, x_k^*)| : x_2^*, \ldots, x_k^* \in [c,d]^{k-1}\} \tag{74}$$

$$M^* := \sup\{|\varphi(d, x_2^*, \ldots, x_k^*)| : x_2^*, \ldots, x_k^* \in [c,d]^{k-1}\}. \tag{75}$$

Proof. Note first that $\mathsf{m}_{\varphi;c,d}(F)$ is well defined (i.e., φ is integrable w.r.t. the k-fold product measure $\bigotimes_{i=1}^k \mu_F$) on $D_{cdf}([a,b])$ because φ is bounded. Next, we reduce the statement to the case $k = 1$: let F and G be elements of $D_{cdf}([a,b])$, let μ be a measure that dominates both μ_F and μ_G, let f and g denote μ-densities of μ_F and μ_G, respectively, w.r.t. μ. Then, by Fubini's theorem,

$$\mathsf{m}_{\varphi;c,d}(F) = \int_{[c,d]} \cdots \int_{[c,d]} \varphi(x_1, \ldots, x_k) \prod_{j=1}^k f(x_j) d\mu(x_1) \ldots d\mu(x_k), \tag{76}$$

and an analogous expression (replacing the density f by the density g) corresponds to $\mathsf{m}_{\varphi;c,d}(G)$. Recall also that for arbitrary real numbers a_j, b_j for $j = 1, \ldots, k$ we may write

$$\prod_{j=1}^k a_j - \prod_{j=1}^k b_j = \sum_{j=1}^k \left[\left(\prod_{i=1}^{j-1} a_i \right) (a_j - b_j) \prod_{i=j+1}^k b_i \right], \tag{77}$$

where empty products are to be interpreted as 1. Equipped with (77), using Equation (76), and Fubini's theorem, we write $\mathsf{m}_{\varphi;c,d}(F) - \mathsf{m}_{\varphi;c,d}(G)$ as

$$\sum_{j=1}^k \int_{[c,d]} \cdots \int_{[c,d]} \varphi(x_1, \ldots, x_k)[f(x_j) - g(x_j)] d\mu(x_j) dF(x_1) \ldots dF(x_{j-1}) d\mu(x_1) \ldots dG(x_{j-1}) \ldots dG(x_k).$$

Using the triangle inequality to upper bound $|\mathsf{m}_{\varphi;c,d}(F) - \mathsf{m}_{\varphi;c,d}(G)|$, an application of the symmetry condition shows that it suffices to verify that for $x_2^*, \ldots, x_k^*$ in $[c,d]^{k-1}$ arbitrary

$$\left| \int_{[c,d]} \varphi(x, x_2^* \ldots, x_k^*) dF(x) - \int_{[c,d]} \varphi(x, x_2^* \ldots, x_k^*) dG(x) \right| \le C^* \|F - G\|_\infty. \tag{78}$$

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function (possibly depending on $x_2^*, \ldots, x_k^*$) such that $f(x) = \varphi(x, x_2^* \ldots, x_k^*)$ holds for every $x \in [c, d]$, and such that $f(x) \to 0$ as $x \to -\infty$. An application of the integration-by-parts formula (as in, e.g., Exercise 34.b on p.108 in Folland (2013)) gives

$$\int_{[c,d]} \varphi(x, x_2^* \ldots, x_k^*) dF(x) = \int_{[c,d]} f(x) dF(x) = f(d)F(d) - f(c-)F(c-) - \int_{[c,d]} F(x) d\mu_f(x), \tag{79}$$

an analogous statement holding for F replaced by G. Hence, the quantity to the left in the inequality in (78) is seen to be not greater than

$$|f(d)||F(d) - G(d)| + |f(c)||F(c-) - G(c-)| + \left| \int_{[c,d]} F(x) - G(x) d\mu_f(x) \right|. \tag{80}$$

Noting that $|f(d)| \leq M^*$, that $|f(c)| \leq m^*$, that $|F(d) - G(d)| = 0$ if $d = b$, that $|F(c-) - G(c-)| = 0$ if $a = c$, and furthermore noting that $|F(d) - G(d)| \leq \|F - G\|_\infty$ and $|F(c-) - G(c-)| \leq \|F - G\|_\infty$ always hold, the statement in (78) follows from the total variation of μ_f on $[c, d]$ being not greater than C. □

Example C.3 (Mean)**.** Let $a < b$ be real numbers. Let $k = 1$ and set $\varphi(x) = x$, i.e., we consider the mean functional $F \mapsto \mu(F)$, say, defined via

$$F \mapsto \int_{[a,b]} x dF(x). \tag{81}$$

Note that φ is bounded on $[a, b]$, is trivially symmetric, and φ satisfies the continuity condition in Lemma C.2. Furthermore, the total variation of φ is $\int_{[a,b]} |\varphi'(x)| dx = (b - a)$. As a consequence of Lemma C.2 the functional m_φ is thus Lipschitz continuous on $D_{cdf}([a, b])$ with constant $(b - a)$.

Example C.4 (Moments)**.** For simplicity, let $a = 0$ and $b > 0$. Let $k = 1$ and set $\varphi(x) = x^p$ for some $p > 0$, i.e., we consider the p-mean functional

$$F \mapsto \int_{[0,b]} x^p dF(x). \tag{82}$$

Note that φ is bounded on $[a, b]$, is trivially symmetric, and φ satisfies the continuity condition in Lemma C.2. Furthermore, the total variation of φ is $\int_{[0,b]} |\varphi'(x)| dx = \varphi(b) = b^p$. As a consequence of Lemma C.2 the functional m_φ is thus Lipschitz continuous on $D_{cdf}([0, b])$ with constant b^p.

Example C.5 (Variance)**.** Let $a < b$ be real numbers. Let $k = 2$ and set $\varphi(x_1, x_2) = .5(x_1 - x_2)^2$, i.e., we consider the variance

$$F \mapsto .5 \int_{[a,b]} \int_{[a,b]} (x_1 - x_2)^2 dF(x_1) dF(x_2) = \int_{[a,b]} \left[x_1 - \int_{[a,b]} x_2 dF(x_2) \right]^2 dF(x_1). \tag{83}$$

Note that φ is bounded on $[a, b]^2$, is symmetric, and φ satisfies the continuity condition in Lemma C.2. For every $x_2 \in [a, b]$ the total variation of $x \mapsto .5(x - x_2)^2$ is $\int_{[a,b]} |x - x_2| dx \leq 2 \max(a^2, b^2) + .5(b^2 - a^2)$. It follows from Lemma C.2 that the variance functional is Lipschitz continuous with constant $2 \max(a^2, b^2) + .5(b^2 - a^2)$.

Example C.6 (Gini-mean difference)**.** Let $a < b$ be real numbers, and let $\varphi(x_1, x_2) = |x_1 - x_2|$. This corresponds to the functional

$$F \mapsto \int_{[a,b]} \int_{[a,b]} |x_1 - x_2| dF(x_1) dF(x_2), \tag{84}$$

which constitutes the numerator of the Gini-index and is sometimes called the Gini-mean difference. Clearly, φ is bounded and symmetric, and satisfies the continuity condition in Lemma C.2. Furthermore, for every $x_2 \in [a, b]$ it holds that the total variation of $x \mapsto |x - x_2|$ equals $(b - a)$. It follows from Lemma C.2 that m_φ is Lipschitz continuous on $D_{cdf}([a, b])$ with constant $2(b - a)$.

The following lemma is sometimes useful, because it avoids the continuity condition of the integrand in Lemma C.2 by working with a monotonicity condition.

Lemma C.7. *Let $a < b$ be real numbers and let $\varphi : [a, b] \to \mathbb{R}$ be right-continuous, and be non-decreasing or non-increasing. Then, the functional*

$$F \mapsto \int_{[a,b]} \varphi(x) dF(x) \tag{85}$$

is Lipschitz continuous on $D_{cdf}([a, b])$ with constant $|\varphi(b) - \varphi(a)|$.

Proof. Note first that the functional under consideration is well defined on $D_{cdf}([a, b])$; and that we only need to consider the case where φ is non-decreasing. To this end let $F, G \in D_{cdf}([a, b])$ and note that, by the transformation theorem, we have

$$\int_{[a,b]} \varphi(x) dF(x) - \int_{[a,b]} \varphi(x) dG(x) = \int_{[\varphi(a),\varphi(b)]} x dF_\varphi(x) - \int_{[\varphi(a),\varphi(b)]} x dG_\varphi(x), \tag{86}$$

where $F_\varphi \in D_{cdf}([\varphi(a), \varphi(b)])$ denotes the cdf corresponding to the image measure $\mu_F \circ \varphi$, and $G_\varphi \in D_{cdf}([\varphi(a), \varphi(b)])$ is defined analogously. An application of Example C.3 thus shows that

$$\left| \int_{[a,b]} \varphi(x) dF(x) - \int_{[a,b]} \varphi(x) dG(x) \right| \leq [\varphi(b) - \varphi(a)] \|F_\varphi - G_\varphi\|_\infty. \tag{87}$$

It remains to observe that $\|F_\varphi - G_\varphi\|_\infty \leq \|F - G\|_\infty$, by Lemma C.8. □

Lemma C.8. *Let F and G be cdfs, and let $\varphi : \mathbb{R} \to \mathbb{R}$ be right-continuous, and be non-decreasing or non-increasing. Then, $\|F_\varphi - G_\varphi\|_\infty \leq \|F - G\|_\infty$ holds, where F_φ denotes the cdf corresponding to the image measure $\mu_F \circ \varphi$, and G_φ is defined analogously.*

Proof. First of all, note that $\|F_\varphi - G_\varphi\|_\infty = \sup_{z \in C(F,G)} |F_\varphi(z) - G_\varphi(z)|$, where $C(F, G) \subseteq \mathbb{R}$ is defined as the subset of points at which both F_φ and G_φ are continuous. Next, let $z \in C(F, G)$ and define $\varphi^-(x) := \inf\{z \in \mathbb{R} : \varphi(z) \geq x\}$, i.e., a generalized inverse of φ. Part (5) of Proposition 1 in Embrechts and Hofert (2013) shows that for every $z \in \mathbb{R}$ we have

$$A(z) := \{x \in \mathbb{R} : \varphi(x) < z\} = \{x \in \mathbb{R} : x < \varphi^-(z)\}. \tag{88}$$

Using this expression for $A(z)$, we can for every $z \in C(F,G)$ rewrite $|F_\varphi(z) - G_\varphi(z)|$ as

$$\begin{aligned}|\mu_{F_\varphi}((-\infty, z)) - \mu_{G_\varphi}((-\infty, z))| &= |\mu_F(A(z)) - \mu_G(A(z))| \\ &= |\mu_F(\{x \in \mathbb{R} : x < \varphi^-(z)\}) - \mu_G(\{x \in \mathbb{R} : x < \varphi^-(z)\})|.\end{aligned}$$

On the one hand, the expression to the far right in the previous display equals $0 \leq \|F - G\|_\infty$ in case $\varphi^-(z) \in \{-\infty, +\infty\}$. On the other hand, if $\varphi^-(z) \in \mathbb{R}$, the same expression is seen to equal $|F(\varphi^-(z)-) - G(\varphi^-(z)-)| \leq \|F - G\|_\infty$. Since this argument goes through for every $z \in C(F,G)$, we are done. ☐

C.4 Quantiles, quantile functions, L-functionals, Lorenz curve, and truncation

In the present subsection we provide some results concerning Lipschitz continuity of quantiles-based functionals. For $\alpha \in [0,1]$ we define the α-quantile of a cdf F as usual via $q_\alpha(F) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\}$. Note that for $\alpha = 0$ we have $q_\alpha(F) = -\infty$, and that (by monotonicity) the quantile function $\alpha \mapsto q_\alpha(F)$ is $\mathcal{B}([0,1]) - \mathcal{B}(\bar{\mathbb{R}})$ measurable. The first result is as follows:

Lemma C.9. *Let $\alpha \in (0,1]$ and let $F \in D_{cdf}([a,b])$ for real numbers $a < b$. Suppose $F(q_\alpha(F)) = \alpha$ and that there exists a positive real number r so that*

$$\begin{aligned} &F(q_\alpha(F) - x) - \alpha \leq -rx \qquad \text{if} \quad x > 0 \text{ and } q_\alpha(F) - x \geq a. \\ &F(q_\alpha(F) + x) - \alpha \geq rx \qquad \text{if} \quad x > 0 \text{ and } q_\alpha(F) + x \leq b. \end{aligned} \tag{89}$$

Then, for every $G \in D_{cdf}([a,b])$ it holds that $|q_\alpha(F) - q_\alpha(G)| \leq r^{-1}\|F - G\|_\infty$. Consequently, denoting by $\mathscr{D}$ the set of all cdfs that satisfy the conditions imposed on F above, it follows that q_α satisfies Assumption 2.1 with $a, b, \mathscr{D}$ and constant $C = r^{-1}$.

Proof. Let G be an element of $D_{cdf}([a,b])$. The claim is trivial if $F = G$. Thus, we assume that $F \neq G$. Note that $F(x) = 0 \leq \alpha$ for every $x < a$, and $F(x) = 1 > \alpha$ for every $x \geq b$ implies $q_\alpha(F) \in [a,b]$; and that, by the same reasoning, $q_\alpha(G) \in [a,b]$. Now, on the one hand, if $q_\alpha(F) - r^{-1}\|G - F\|_\infty < a$, then $q_\alpha(G) \geq q_\alpha(F) - r^{-1}\|G - F\|_\infty$. If, on the other hand, $q_\alpha(F) - r^{-1}\|G - F\|_\infty \geq a$, then from the first line in (89) with $x = r^{-1}\|G - F\|_\infty$ one obtains $\alpha \geq F(q_\alpha(F) - r^{-1}\|G - F\|_\infty) + \|G - F\|_\infty$, thus $\alpha \geq G(q_\alpha(F) - r^{-1}\|G - F\|_\infty)$ and hence, again, $q_\alpha(G) \geq q_\alpha(F) - r^{-1}\|G - F\|_\infty$. Similarly, on the one hand, if $q_\alpha(F) + r^{-1}\|G - F\|_\infty > b$, then $q_\alpha(G) \leq q_\alpha(F) + r^{-1}\|G - F\|_\infty$. If, on the other hand $q_\alpha(F) + r^{-1}\|G - F\|_\infty \leq b$, then the second line in (89) with $x = r^{-1}\|G - F\|_\infty$ shows that $F(q_\alpha(F) + r^{-1}\|G - F\|_\infty) - \|G - F\|_\infty \geq \alpha$, thus $G(q_\alpha(F) + r^{-1}\|G - F\|_\infty) \geq \alpha$, and hence, again, $q_\alpha(G) \leq q_\alpha(F) + r^{-1}\|G - F\|_\infty$. Summarizing yields $|q_\alpha(F) - q_\alpha(G)| \leq r^{-1}\|F - G\|_\infty$. The last statement is trivial. ☐

Example C.10 (Median)**.** The median of a distribution F is defined as its $\alpha = 1/2$ quantile $q_{1/2}(F)$. Let $a < b$ and $r > 0$ be real numbers, and denote by $\mathscr{D}$ the set of cdfs F so that $F(q_{1/2}(F)) = 1/2$, and so that Equation (89) is satisfied for $\alpha = 1/2$ (Lemma C.12 provides a sufficient condition for $F \in \mathscr{D}$). Then, the functional $F \mapsto q_{1/2}(F)$ satisfies Assumption 2.1 with a, b and $\mathscr{D}$ with constant $C = r^{-1}$.

The second result is auxiliary, and concerns not a single quantile, but the Lipschitz continuity of the quantile *function* $F \mapsto q_{\cdot}(F)$ on certain subsets of $[0,1]$.

Lemma C.11. *Let $F \in D_{cdf}([a,b])$ for real numbers $a < b$, and let $\alpha_* < \alpha^*$ for α_* and α^* in $(0,1]$. Suppose $F(q_\alpha(F)) = \alpha$ holds for every $\alpha \in [\alpha_*, \alpha^*]$, and that there exists a positive real number r so that Equation (89) is satisfied for every $\alpha \in [\alpha_*, \alpha^*]$. Then, for every $G \in D_{cdf}([a,b])$ it holds that*

$$\sup_{\alpha \in [\alpha_*, \alpha^*]} |q_\alpha(F) - q_\alpha(G)| \leq r^{-1} \|F - G\|_\infty.$$

Proof. The statement follows immediately from Lemma C.9. □

A simple sufficient condition for the assumption on F in Lemma C.11 (and hence also for the assumption on F in Lemma C.9) is that F admits a density that is bounded from below (on the support of F):

Lemma C.12. *Let $a < b$ be real numbers and let $F \in D_{cdf}([a,b])$. Suppose F is continuous, and is right-sided differentiable on (a,b) with right-sided derivative F^+, which furthermore satisfies $F^+(x) \geq r$ for every $x \in (a,b)$ for some $r > 0$. Then, $F(q_\alpha(F)) = \alpha$ and Equation (89) holds for every $\alpha \in (0,1]$.*

Proof. The condition $F^+(x) \geq r$ for every $x \in (a,b)$ for a $r > 0$ implies that F is strictly increasing on $[a,b]$, which (together with continuity of F) implies $F(q_\alpha(F)) = \alpha$ for every $\alpha \in (0,1]$. The second claim follows from the mean-value theorem for right-differentiable functions in Minassian (2007) (noting that $q_\alpha(F) \in [a,b]$ for every $\alpha \in (0,1]$, cf. the proof of Lemma C.9). □

The next result in this section concerns population versions of L-statistics.

Lemma C.13. *Let ν be a measure on the Borel sets of $[0,1]$, and let $J : [0,1] \to \mathbb{R}$ be such that $\int_{[0,1]} |J(\alpha)| d\nu(\alpha) = c < \infty$. Assume further that $\nu(0) = 0$. Let $d \in \mathbb{N} \cup \{0\}$, let $0 < p_1 < \ldots < p_d \leq 1$, and let $v_1, \ldots, v_d$ be real numbers. Let $a < b$ be real numbers and define on $D_{cdf}([a,b])$ the functional*

$$\mathsf{T}(F) = \int_{[0,1]} q_\alpha(F) J(\alpha) d\nu(\alpha) + \sum_{j=1}^{d} v_i q_{p_i}(F), \tag{90}$$

the sum to the right being interpreted as 0 if $d = 0$. Let $F \in D_{cdf}([a,b])$ satisfy $F(q_\alpha(F)) = \alpha$ for every $\alpha \in (0,1]$, and suppose there is a positive real number r so that Equation (89) holds for every $\alpha \in (0,1]$. Then, for every $G \in D_{cdf}([a,b])$ it holds that

$$|\mathsf{T}(F) - \mathsf{T}(G)| \leq r^{-1} \left[c + \sum_{i=1}^{d} |v_i| \right] \|F - G\|_\infty.$$

Consequently, denoting by $\mathscr{D}$ the set of all cdfs that satisfy the conditions imposed on F above, it follows that T defined in Equation (90) satisfies Assumption 2.1 with $a, b, \mathscr{D}$ and constant $C = r^{-1}[c + \sum_{i=1}^{d} |v_i|]$.

Proof. We first verify that $\mathsf{T}(F)$ is well defined for every $F \in D_{cdf}([a,b])$. To this end, we show that the integral $\int_{[0,1]} q_\alpha(F)J(\alpha)d\nu(\alpha)$ exists for every $F \in D_{cdf}([a,b])$. Note that the ($\mathcal{B}([0,1])$ - $\mathcal{B}(\bar{\mathbb{R}})$ measurable) function $g(\alpha) : [0,1] \to \bar{\mathbb{R}}$ defined via $\alpha \mapsto |q_\alpha(F)J(\alpha)|$ coincides ν-almost everywhere with the ($\mathcal{B}([0,1])$ - $\mathcal{B}(\mathbb{R})$ measurable) function

$$g^*(\alpha) := \begin{cases} g(\alpha) & \text{if } \alpha \in (0,1] \\ 0 & \text{if } \alpha = 0. \end{cases} \tag{91}$$

Note further that the function $\alpha \mapsto q_\alpha(F)$ is well defined on $(0,1]$ and its range is contained in $[a,b]$ (cf. the argument in the beginning of the proof of Lemma C.9). It thus follows that $|g^*(\alpha)| \leq \max(|a|,|b|)|J(\alpha)|$, and the integrability condition on J shows that g^* (and thus g) is integrable. Now, let F and G be as in the statement of the lemma. Consider $|\mathsf{T}(F) - \mathsf{T}(G)|$. Clearly, by the triangle inequality and Lemma C.11, it suffices to verify the statement for the case $d = 0$. Then, $|\mathsf{T}(F) - \mathsf{T}(G)|$ is not greater than

$$\int_{(0,1]} |q_\alpha(F) - q_\alpha(G)||J(\alpha)|d\nu(\alpha). \tag{92}$$

Note that the function $\alpha \mapsto |q_\alpha(F) - q_\alpha(G)|$ is bounded on $(0,1]$. By the monotonic convergence theorem for $\varepsilon \searrow 0$ the integral $\int_{[\varepsilon,1]} |q_\alpha(F) - q_\alpha(G)||J(\alpha)|d\nu(\alpha)$ converges to the integral in (92). But $\int_{[\varepsilon,1]} |q_\alpha(F) - q_\alpha(G)||J(\alpha)|d\nu(\alpha) \leq r^{-1}c\|F - G\|_\infty$ by Lemma C.11. The last statement is trivial. ☐

One particularly important application concerns the so-called Lorenz curve associated with a cdf F which is defined (cf. Gastwirth (1971)) below in Equation (94).

Lemma C.14. *Let $a < b$ be real numbers and define on $D_{cdf}([a,b])$ the family of functionals indexed by $u \in [0,1]$ and defined by*

$$Q(F,u) := \int_{[0,u]} q_\alpha(F)d\alpha; \tag{93}$$

furthermore, if $a > 0$, define the family of functionals indexed by $u \in [0,1]$ via

$$L(F,u) := \mu(F)^{-1} \int_{[0,u]} q_\alpha(F)d\alpha \tag{94}$$

Let $F \in D_{cdf}([a,b])$ satisfy $F(q_\alpha(F)) = \alpha$ for every $\alpha \in (0,1]$, and suppose there is a positive real number r so that Equation (89) holds for every $\alpha \in (0,1]$. Then, for every $G \in D_{cdf}([a,b])$ it holds that

$$|Q(F,u) - Q(G,u)| \leq r^{-1}u\|F - G\|_\infty \leq r^{-1}\|F - G\|_\infty.$$

Consequently, denoting by $\mathscr{D}$ the set of all cdfs that satisfy the conditions imposed on F above, it follows that $T = Q(.,u)$ satisfies Assumption 2.1 with $a, b, \mathscr{D}$ and constant $C = r^{-1}u$. Furthermore, if $a > 0$, then

$$|L(F,u) - L(G,u)| \leq a^{-1}(r^{-1}u + (b-a)a^{-1})\|F - G\|_\infty,$$

and it follows that $T = L(.,u)$ satisfies Assumption 2.1 with $a, b, \mathscr{D}$ and constant $C = a^{-1}(r^{-1} + (b - a)a^{-1})$.

Proof. For the first claim, we just apply Lemma C.13 with ν equal to Lebesgue measure, $J = \mathbf{1}_{[0,u]}$, which satisfies the integrability condition with $c = u \leq 1$.

For the second claim, note that $L(.,u)$ is well defined on $D_{cdf}([a,b])$ because $a > 0$. Next, observe that for F and G as in the statement of the lemma we can bound $|L(F,u) - L(G,u)|$ from above by

$$\mu(F)^{-1}\left\{\left|Q(F,u) - Q(G,u)\right| + |1 - \mu(F)/\mu(G)| \int_{[0,u]} q_\alpha(G) d\alpha\right\}. \tag{95}$$

Since $\mu(F) \geq a$, since $q_\alpha(G) \leq b$ for $\alpha \in (0,u]$, and because we already know that

$$\left|Q(F,u) - Q(G,u)\right| \leq r^{-1}u\|F - G\|_\infty,$$

it remains to observe that

$$|1 - (\mu(F)/\mu(G))| \leq (b-a)\|F - G\|_\infty/\mu(G) \leq (b-a)a^{-1}\|F - G\|_\infty \tag{96}$$

to conclude that the expression in (95) is not greater than $a^{-1}\left\{r^{-1}u + (b-a)a^{-1}\right\}\|F - G\|_\infty$. □

The final result in this section concerns trimmed generalized-mean functionals. We consider one-sidedly trimmed functionals, the trimming affecting the lower or upper tail. Two-sided trimming can be dealt with similarly. We abstain from spelling out the details.

Lemma C.15. *Let $a < b$ be real numbers, let $\varphi : \mathbb{R} \to \mathbb{R}$, let φ restricted to $[a,b]$ be continuous, let the total variation of φ on $[a,b]$ be not greater than C, and let $|\varphi(x)| \leq u$ hold for all $x \in [a,b]$. Furthermore, let $\alpha \in (0,1)$. For $F \in D_{cdf}([a,b])$ define*

$$\mathsf{m}^{t-}_{\varphi;\alpha}(F) := \int_{[a,q_\alpha(F)]} \varphi(x)dF(x) \quad \text{and} \quad \mathsf{m}^{t+}_{\varphi;\alpha}(F) := \int_{[q_\alpha(F),b]} \varphi(x)dF(x). \tag{97}$$

Let $F \in D_{cdf}([a,b])$, assume that F is continuous, and right-sided differentiable on (a,b), with right-sided derivative F^+ satisfying $r \leq F^+(x) \leq \kappa$ for every $x \in (a,b)$, and for positive real numbers κ and r. Then, for every $G \in D_{cdf}([a,b])$ it holds that

$$|\mathsf{m}^{t-}_{\varphi;\alpha}(F) - \mathsf{m}^{t-}_{\varphi;\alpha}(G)| \leq [C + u(1 + \kappa r^{-1})]\|F - G\|_\infty,$$

and

$$|\mathsf{m}^{t+}_{\varphi;\alpha}(F) - \mathsf{m}^{t+}_{\varphi;\alpha}(G)| \leq [C + u(1 + \kappa r^{-1})]\|F - G\|_\infty, \tag{98}$$

Consequently, denoting by $\mathscr{D}$ the set of all cdfs that satisfy the conditions imposed on F above, it follows that $m^{t-}_{\varphi;\alpha}$ and $m^{t+}_{\varphi;\alpha}$ satisfy Assumption 2.1 with $a, b, \mathscr{D}$ and constant $C + u(1 + \kappa r^{-1})$.

Proof. We only provide an argument for the first claimed inequality, the second is obtained analogously. Furthermore, throughout the proof we write $\mathsf{m}^{t}_{\varphi;\alpha}$ instead of $\mathsf{m}^{t-}_{\varphi;\alpha}$. First, note that the functional $\mathsf{m}^{t}_{\varphi;\alpha}(F)$ is indeed well defined for every $F \in D_{cdf}([a,b])$. This follows from $q_\alpha(F) \in [a,b]$, and since φ is bounded on $[a,b]$. Next, let F be as in the statement of the lemma and satisfy the conditions imposed, and let $G \in D_{cdf}([a,b])$. Note first that $q_\alpha(F), q_\alpha(G) \in [a,b]$ (cf. the argument in the beginning of the proof of Lemma C.9). By the triangle inequality, $|\mathsf{m}^t_{\varphi;\alpha}(F) - \mathsf{m}^t_{\varphi;\alpha}(G)| \leq A + B$, where (using the notation introduced in Equation (72))

$$A := \left|\mathsf{m}_{\varphi;a,q_\alpha(G)}(F) - \mathsf{m}_{\varphi;a,q_\alpha(G)}(G)\right| \leq (C+u)\|F-G\|_\infty,$$

the upper bound following from Lemma C.2, and

$$B := \int g(x)|\varphi(x)|dF(x) \leq u \int g(x)dF(x), \tag{99}$$

where $g(x) = \left|\mathbf{1}_{[a,q_\alpha(F)]}(x) - \mathbf{1}_{[a,q_\alpha(G)]}(x)\right|$. By continuity of F:

$$\int g(x)dF(x) \leq |F(q_\alpha(G)) - \alpha|. \tag{100}$$

which, by the assumed behavior of the right-derivative of F and a mean-value theorem for right-differentiable functions (for example the one by Minassian (2007)), is not greater than

$$\kappa|q_\alpha(G) - q_\alpha(F)| \leq \kappa r^{-1}\|F-G\|_\infty \tag{101}$$

the last inequality following from Lemma C.9 together with Lemma C.12. This proves the claim. The last statement is trivial. □

D Proofs of results in Sections 3, 4 and 5

Throughout the appendix, $KL(\cdot,\cdot)$ denotes the Kullback-Leibler (KL) divergence between two probability measures (on a Borel σ-algebra clear from the context) or, if applicable, a version of their densities.

D.1 Proofs of results in Section 3

D.1.1 Proof of Theorem 3.1

To establish Equation (5) we need to show that for every K-tuple $(F^1,\ldots,F^K)$ with $F^i \in \mathscr{D}$, $i = 1,\ldots,K$, we have $\mathbb{E}(R_N(\hat{\pi})) \leq c\sqrt{Kn\overline{\log}(n)}$. Note that this inequality trivially holds in case $\mathsf{T}(F^1) = \ldots = \mathsf{T}(F^K)$. In particular, if $\mathsf{T}(F^1) = \ldots = \mathsf{T}(F^K)$ holds for every $F^1,\ldots,F^K$ such that $F_i \in \mathscr{D}$ for $i = 1,\ldots,K$ there is nothing to prove. We thus assume without loss of generality that the K-tuple $F^1,\ldots,F^K$ is such that $\mathsf{T}(F^i)$ is not constant in $i \in \{1,\ldots,K\}$; also implying that the constant C from Assumption 2.1 must satisfy $C > 0$.

We now claim that for every i with $\Delta_i > 0$ we have

$$\mathbb{E}[S_i(N)] \leq \frac{2C^2\beta\log(n)}{\Delta_i^2} + \frac{\beta+2}{\beta-2}, \tag{102}$$

where we recall that $n = \mathbb{E}(N)$. Before proving this claim, note that Assumption 2.1 implies $\Delta_i \leq C$, and that Equation (2) shows that

$$\mathbb{E}[R_N(\hat{\pi})] = \sum_{i:\Delta_i>0} \Delta_i\mathbb{E}[S_i(N)], \tag{103}$$

which together with the claim in Equation (102) yield

$$\begin{aligned}\mathbb{E}[R_N(\hat{\pi})] &= \sum_{i:\Delta_i>0} \sqrt{\Delta_i^2 \mathbb{E}[S_i(N)]}\sqrt{\mathbb{E}[S_i(N)]} \\ &\le \sqrt{2C^2\beta\log(n) + C^2(\beta+2)/(\beta-2)} \sum_{i:\Delta_i>0} \sqrt{\mathbb{E}[S_i(N)]} \\ &\le \sqrt{2C^2\beta\log(n) + C^2(\beta+2)/(\beta-2)}\sqrt{Kn},\end{aligned}$$

where the last line follows from the Cauchy-Schwarz inequality and $\sum_{i:\Delta_i>0}\mathbb{E}[S_i(N)] \le \mathbb{E}[N] = n$. Upon choosing $c = C\sqrt{2\beta + (\beta+2)/(\beta-2)}$ we would thus obtain Equation (5). Therefore, it remains to prove the statement in Equation (102). Before doing that, we note for later use that Equations (102) and (103) also give the regret bound

$$\mathbb{E}[R_N(\hat{\pi})] \le \sum_{i:\Delta_i>0} \left(\frac{2C^2\beta\log(n)}{\Delta_i} + \frac{\beta+2}{\beta-2}\Delta_i \right). \tag{104}$$

Now, to prove Equation (102), note that by Tonelli's theorem

$$\mathbb{E}(S_i(N)) = \mathbb{E}\left(\sum_{t=1}^{\infty} 1_{\{N=t\}} S_i(N) \right) = \sum_{t=1}^{\infty} \mathbb{E}\left(1_{\{N=t\}} S_i(t) \right), \tag{105}$$

where we used that $\mathbb{P}(N \in \mathbb{N}) = 1$, a consequence of $\mathbb{E}(N) = n \in \mathbb{N}$. Denote the σ-algebra generated by $Y_1, \ldots, Y_t$ by $\mathcal{A}_t$. By assumption, $\sigma(N)$ and $\mathcal{A}_t$ are independent for every $t \in \mathbb{N}$. Note furthermore that $S_i(t)$ is $\mathcal{A}_t$ measurable for every $t \in \mathbb{N}$. Hence, for every $t \in \mathbb{N}$ we have

$$\mathbb{E}\left(1_{\{N=t\}} S_i(t) \right) = \mathbb{E}\left(S_i(t) \mathbb{E}\left(1_{\{N=t\}} | \mathcal{A}_t \right) \right) = \mathbb{P}\left(N=t\right) \mathbb{E}\left(S_i(t) \right), \tag{106}$$

from which it follows that

$$\mathbb{E}(S_i(N)) = \sum_{t=1}^{\infty} \mathbb{P}\left(N=t\right) \mathbb{E}\left(S_i(t) \right). \tag{107}$$

If Equation (102) were already known to be true for any N that coincides with its expectation with probability one, we could apply this to any $N \equiv t$ (implying that the corresponding expectation equals t), which together with the previous display would deliver that

$$\mathbb{E}(S_i(N)) \le \frac{2C^2\beta \sum_{t=1}^{\infty} \mathbb{P}\left(N=t\right)\log(t)}{\Delta_i^2} + \frac{\beta+2}{\beta-2} \le \frac{2C^2\log(n)}{\Delta_i^2} + \frac{\beta+2}{\beta-2}, \tag{108}$$

where we used Jensen's inequality to obtain the last inequality. Therefore, it remains to establish Equation (102) for N such that $\mathbb{P}(N = \mathbb{E}(N)) = 1$.

Let N satisfy $\mathbb{P}(N = \mathbb{E}(N)) = 1$. Note that without loss of generality we can assume that $N = n$ holds everywhere. If $n \le K$, we have $S_i(n) \le 1$ and hence (102) is obviously satisfied. From now on, we therefore let $n \geq K+1$. Furthermore, we fix an i such that $\Delta_i > 0$. Now, for every $t \in \{K+1, \ldots, n\}$, we note that $\{\pi_t = i\} \subseteq A_t \cup B_{i,t} \cup C_{i,t}$, where

$$\begin{aligned}
A_t &:= \Big\{ \mathsf{T}(\hat{F}_{i^*,t-1}) + C\sqrt{\beta \log(t)/2S_{i^*}(t-1)} \leq \mathsf{T}(F^{i^*}) \Big\}, \\
B_{i,t} &:= \Big\{ \mathsf{T}(\hat{F}_{i,t-1}) > \mathsf{T}(F^i) + C\sqrt{\beta \log(t)/2S_i(t-1)} \Big\}, \\
C_{i,t} &:= \Big\{ S_i(t-1) < \frac{2\beta C^2 \log(n)}{\Delta_i^2} \Big\},
\end{aligned}$$

and where we define i^* as the smallest element of $\arg\max_{i=1,\ldots,K} \mathsf{T}(F_i)$. Indeed, on the complement of $A_t \cup B_{i,t} \cup C_{i,t}$ we have

$$\begin{aligned}
\mathsf{T}(\hat{F}_{i^*,t-1}) + C\sqrt{\beta \log(t)/2S_{i^*}(t-1)} &> \mathsf{T}(F^{i^*}) \\
&\geq \mathsf{T}(F^i) + 2C\sqrt{\beta \log(n)/2S_i(t-1)} \\
&\geq \mathsf{T}(F^i) + 2C\sqrt{\beta \log(t)/2S_i(t-1)} \\
&\geq \mathsf{T}(\hat{F}_{i,t-1}) + C\sqrt{\beta \log(t)/2S_i(t-1)},
\end{aligned}$$

implying $\hat{\pi}_t = i^*$, which contradicts $\hat{\pi}_t = i$, because $i \neq i^*$ as $\Delta_i > 0$. Using $\{\pi_t = i\} \subseteq A_t \cup B_{i,t} \cup C_{i,t}$ and setting $u := \left\lceil \frac{2C^2\beta\log(n)}{\Delta_i^2} \right\rceil$, we now obtain (recall that $n \geq K+1$) that

$$\begin{aligned}
S_i(n) = \sum_{t=1}^{K} 1_{\{\pi_t = i\}} + \sum_{t=K+1}^{n} 1_{\{\pi_t = i\}} = 1 + \sum_{t=K+1}^{n} 1_{\{\pi_t = i\}} = 1 + \sum_{t=K+1}^{n} 1_{\{\pi_t = i\} \cap C_{i,t}} + \sum_{t=K+1}^{n} 1_{\{\pi_t = i\} \cap C_{i,t}^c} \\
\leq u + \sum_{t=K+1}^{n} 1_{A_t \cup B_{i,t}},
\end{aligned}$$

where we also used $1 + \sum_{t=K+1}^{n} 1_{\{\pi_t = i\} \cap C_{i,t}} \leq u$. From the upper bound in the previous display we get

$$\mathbb{E}[S_i(n)] \leq u + \sum_{t=K+1}^{n} \mathbb{P}(A_t) + \mathbb{P}(B_{i,t}).$$

We will show further below that:

$$\begin{aligned}
\mathbb{P}(A_t) &\leq \sum_{s=1}^{t} \mathbb{P}\big(\mathsf{T}(F_{i^*,s}) + C\sqrt{\beta \log(t)/2s} \leq \mathsf{T}(F^{i^*})\big) \\
\mathbb{P}(B_{i,t}) &\leq \sum_{s=1}^{t} \mathbb{P}\big(\mathsf{T}(F_{i,s}) > \mathsf{T}(F^i) + C\sqrt{\beta \log(t)/2s}\big),
\end{aligned} \tag{109}$$

where for every $s \in \{1, \ldots, t\}$ and every $l \in \{i, i^*\}$ we define $F_{l,s} := s^{-1} \sum_{j=1}^{s} 1_{\{Y_{l,j} \leq \cdot\}}$. From Equation (109), Assumption 2.1 and the Dvoretzky-Kiefer-Wolfowitz (DKW) inequality in the form established in Corollary 1 in Massart (1990) (note that Equation (1.5) in Massart (1990) obviously remains valid if ">" is replaced by "≥") we then obtain

$$\mathbb{P}(A_t) \leq \sum_{s=1}^{t} \mathbb{P}(||F_{i^*,s} - F^{i^*}||_\infty \geq \sqrt{\beta \log(t)/2s}) \leq 2 \sum_{s=1}^{t} \frac{1}{t^\beta} = \frac{2}{t^{\beta-1}}$$
$$\mathbb{P}(B_{i,t}) \leq \sum_{s=1}^{t} \mathbb{P}(||F_{i,s} - F^{i}||_\infty > \sqrt{\beta \log(t)/2s}) \leq 2 \sum_{s=1}^{t} \frac{1}{t^\beta} = \frac{2}{t^{\beta-1}}.$$

The identity

$$\sum_{t=K+1}^{n} \frac{1}{t^{\beta-1}} \leq \int_K^\infty \frac{1}{x^{\beta-1}} dx = \frac{1}{(\beta-2)K^{\beta-2}} \leq \frac{1}{\beta-2} \tag{110}$$

combined with $u \leq 1 + 2C^2 \beta \log(n)/\Delta_i^2$ now establishes (102).

It remains to verify the two inequalities in Equation (109). To this end we need some more notation: For every $l \in \{1, \ldots, K\}$, every $r \in \mathbb{N}$ and every $\omega \in \Omega$ let

$$t_{l,r}(\omega) := \inf\{s \in \mathbb{N} : \sum_{j=1}^{s} \mathbb{1}_{\{\hat{\pi}_j(Z_{j-1}(\omega))=l\}} = r\}. \tag{111}$$

Lemma D.1. *For every $l \in \{1, \ldots, K\}$, every $r \in \mathbb{N}$ and every $\omega \in \Omega$ it holds that $t_{l,r}(\omega) \in \mathbb{N}$.*

Proof. Suppose there exists a triple l, r, ω such that $\sum_{j=1}^{s} \mathbb{1}_{\{\hat{\pi}_j(Z_t(\omega))=l\}} < r$ holds for every $s \in \mathbb{N}$, implying in particular that

$$1 \leq \sum_{j=1}^{\infty} \mathbb{1}_{\{\hat{\pi}_j(Z_{j-1}(\omega))=l\}} =: \kappa(\omega) < r, \tag{112}$$

where we used $t_{l,1} = l$. Let $t \geq K+1$. From the definition of $\hat{\pi}$ it follows that

$$\hat{\pi}_t(Z_{t-1}(\omega)) \in \arg\max_{j \in \mathcal{I}} \left\{ \mathsf{T}(\hat{F}_{j,t-1}(.)(\omega)) + C\sqrt{\beta \log(t)/2S_j(t-1)(\omega)} \right\}, \tag{113}$$

where we now evaluate all random variables at ω which we emphasize in the preceding display. In particular, $\mathsf{T}(\hat{F}_{j,t-1}(.)(\omega))$ and $S_j(t-1)(\omega)$ are sequences of real numbers. For convenience, we shall write $\hat{F}_{j,t-1}(\omega)$ instead of $\hat{F}_{j,t-1}(.)(\omega)$ in what follows. From Equation (112) and the previous display it follows that eventually

$$\mathsf{T}(\hat{F}_{\hat{\pi}_t(Z_{t-1}(\omega)),t-1}(\omega)) + C\sqrt{\beta \log(t)/2S_{\hat{\pi}_t(Z_{t-1}(\omega))}(t-1)(\omega)} \geq \mathsf{T}(\hat{F}_{l,t-1}(\omega)) + C\sqrt{\beta \log(t)/2S_l(t-1)(\omega)}, \tag{114}$$

which is equivalent to (recall that $C > 0$ from the discussion in the first paragraph of the present subsection)

$$a_t \left[\mathsf{T}(\hat{F}_{\hat{\pi}_t(Z_{t-1}(\omega)),t-1}(\omega)) - \mathsf{T}(\hat{F}_{l,t-1}(\omega)) \right] \geq \left[[S_l(t-1)(\omega)]^{-1/2} - [S_{\hat{\pi}_t(Z_{t-1}(\omega))}(t-1)(\omega)]^{-1/2} \right], \tag{115}$$

where $a_t := [C\sqrt{\beta \log(t)/2}]^{-1} \to 0$ as $t \to \infty$. Consequently, the sequence on the left hand side of the previous inequality converges to 0 as $t \to \infty$. To see this, let $F \in \mathscr{D}$ and note that

$$|\mathsf{T}(\hat{F}_{\hat{\pi}_t(Z_{t-1}(\omega)),t-1}(\omega)) - \mathsf{T}(\hat{F}_{l,t-1}(\omega))| \leq |\mathsf{T}(\hat{F}_{\hat{\pi}_t(Z_{t-1}(\omega)),t-1}(\omega)) - \mathsf{T}(F)| + |\mathsf{T}(F) - \mathsf{T}(\hat{F}_{l,t-1}(\omega))| \leq 2C. \quad (116)$$

It thus follows that

$$\limsup_{t\to\infty} \left[[S_l(t-1)(\omega)]^{-1/2} - [S_{\hat{\pi}_t(Z_{t-1}(\omega))}(t-1)(\omega)]^{-1/2} \right] \leq 0, \quad (117)$$

or equivalently, noting that Equation (112) implies $\lim_{t\to\infty} S_l(t-1)(\omega) = \kappa(\omega)$, that

$$\limsup_{t\to\infty} S_{\hat{\pi}_t(Z_{t-1}(\omega))}(t-1)(\omega) \leq \kappa(\omega). \quad (118)$$

This, however, implies the contradiction that for every $j = 1, \ldots, K$ it must hold that $\lim_{t\to\infty} S_j(t-1)(\omega) < \infty$ (the limit existing due to monotonicity). To see the latter, suppose $\lim_{t\to\infty} S_j(t-1)(\omega) = \infty$ holds for treatment j. Define the subsequence $t' := \{t \in \mathbb{N} : \hat{\pi}_t(Z_{t-1}(\omega)) = j\}$ of $\mathbb{N}$, and note that $t' \to \infty$ due to our assumption $\lim_{t\to\infty} S_j(t-1)(\omega) = \infty$. Next, observe that $S_j(t'-1)(\omega) = S_{\hat{\pi}_{t'}(Z_{t'-1}(\omega))}(t'-1)(\omega)$, a contradiction to the previous display. □

Before we proceed, let $l \in \{1, \ldots, K\}$. Note that $t_{l,r} < t_{l,s}$ holds for all pairs of natural numbers $r < s$. Note also that for all pairs of natural numbers r and s the event $\{t_{l,s} = r\}$ is measurable w.r.t. the σ-algebra generated by $Y_1, \ldots, Y_{r-1}$, i.e., w.r.t. $\mathcal{A}_r := \sigma(Y_1, \ldots, Y_{r-1})$.

Lemma D.2. *For every $l \in \{1, \ldots, K\}$ and every $r \in \mathbb{N}$ the joint distribution of $Y_{l,1}, \ldots, Y_{l,r}$ coincides with the joint distribution of $Y_{l,t_{l,1}}, \ldots, Y_{l,t_{l,r}}$.*

Proof. Let $l \in \{1, \ldots, K\}$. Note that for any $r \in \mathbb{N}$ the random variables $Y_{l,t_{l,1}}, \ldots, Y_{l,t_{l,r}}$ are well defined as a consequence of Lemma D.1. To prove the statement, we use induction on r, and start with $r = 1$. In this case the statement is trivial, because $t_{l,1} = l$ implies $Y_{l,t_{l,1}} = Y_{l,l}$ which has the same distribution as $Y_{l,1}$. Next, assume that $r > 1$. By the induction hypothesis, we need to show that for $A_j \in \mathcal{B}(\mathbb{R})$ for $j = 1, \ldots, r$ we have

$$\mathbb{P}\left(Y_{l,t_{l,1}} \in A_1, \ldots, Y_{l,t_{l,r}} \in A_r\right) = \mathbb{P}\left(Y_{l,t_{l,1}} \in A_1, \ldots, Y_{l,t_{l,r-1}} \in A_{r-1}\right) \mathbb{P}(Y_{l,r} \in A_r). \quad (119)$$

Let $\mathbb{I}_r := \{A \subseteq \mathbb{N} : |A| = r\}$, and for any $I \in \mathbb{I}_r$, let $I_j,\ j = 1, ..., r$ denote the j-th element of I (I being ordered from smallest to largest). Observe that the random set $\{t_{l,1}, \ldots, t_{l,r}\}$ takes its values in I for some $I \subseteq \mathbb{I}_r$, implying that $\sum_{I\in\mathbb{I}_r} \prod_{k=1}^r 1_{\{t_{l,k}=I_k\}} = 1$. We can thus write

$$\mathbb{P}\left(Y_{l,t_{l,1}} \in A_1, \ldots, Y_{l,t_{l,r}} \in A_r\right) = \mathbb{E}\left(\prod_{j=1}^r 1_{A_j}(Y_{l,t_{l,j}})\right) = \mathbb{E}\left(\prod_{j=1}^r 1_{A_j}(Y_{l,t_{l,j}}) \sum_{I\in\mathbb{I}_r} \prod_{k=1}^r 1_{\{t_{l,k}=I_k\}}\right), \quad (120)$$

which can further be rewritten as

$$\sum_{I\in\mathbb{I}_r} \mathbb{E}\left(\prod_{j=1}^r 1_{A_j}(Y_{l,t_{l,j}}) 1_{\{t_{l,j}=I_j\}}\right) = \sum_{I\in\mathbb{I}_r} \mathbb{E}\left(\prod_{j=1}^r 1_{A_j}(Y_{l,I_j}) 1_{\{t_{l,j}=I_j\}}\right). \quad (121)$$

Next, using that $\{t_{l,j} = I_j\} \in \mathcal{A}_{I_r-1}$ holds for every $j = 1, \ldots, r$, we write the expectation to the far right in the previous display as

$$\mathbb{E}\left[\mathbb{E}\left(\prod_{j=1}^{r} 1_{A_j}(Y_{l,I_j}) 1_{\{t_{l,j}=I_j\}} \middle| \mathcal{A}_{I_r-1}\right)\right] = \mathbb{E}\left[\prod_{j=1}^{r-1} 1_{A_j}(Y_{l,I_j}) 1_{\{t_{l,j}=I_j\}} 1_{\{t_{l,r}=I_r\}} \mathbb{E}\left(1_{A_r}(Y_{l,I_r}) | \mathcal{A}_{I_r-1}\right)\right]. \quad (122)$$

Clearly Y_{l,I_r} is independent of $Y_1, \ldots, Y_{I_r-1}$, and thus the inner conditional expectation coincides with $\mathbb{E}\left(1_{A_r}(Y_{l,I_r})\right) = \mathbb{P}(Y_{l,I_r} \in A_r) = \mathbb{P}(Y_{l,r} \in A_r)$. To prove (119) it hence remains to verify that

$$\sum_{I \in \mathbb{I}_r} \mathbb{E}\left[\prod_{j=1}^{r-1} 1_{A_j}(Y_{l,I_j}) 1_{\{t_{l,j}=I_j\}} 1_{\{t_{l,r}=I_r\}}\right] = \mathbb{P}(Y_{l,t_{l,1}} \in A_1, \ldots, Y_{l,t_{l,r-1}} \in A_{r-1}). \quad (123)$$

To see this, write

$$\begin{aligned}\sum_{I \in \mathbb{I}_r} \mathbb{E}\left[\prod_{j=1}^{r-1} 1_{A_j}(Y_{l,I_j}) 1_{\{t_{l,j}=I_j\}} 1_{\{t_{l,r}=I_r\}}\right] &= \sum_{I \in \mathbb{I}_{r-1}} \sum_{k > I_{r-1}} \mathbb{E}\left[\prod_{j=1}^{r-1} 1_{A_j}(Y_{l,I_j}) 1_{\{t_{l,j}=I_j\}} 1_{\{t_{l,r}=k\}}\right] \\ &= \sum_{I \in \mathbb{I}_{r-1}} \mathbb{E}\left[\prod_{j=1}^{r-1} 1_{\{t_{l,j}=I_j\}} 1_{A_j}(Y_{l,I_j}) \sum_{k > I_{r-1}} 1_{\{t_{l,r}=k\}}\right],\end{aligned}$$

and note that $t_{l,r} > t_{l,r-1}$ implies that $\prod_{j=1}^{r-1} 1_{\{t_{l,j}=I_j\}} \sum_{k>I_{r-1}} 1_{\{t_{l,r}=k\}} = \prod_{j=1}^{r-1} 1_{\{t_{l,j}=I_j\}}$, which together with $\sum_{I \in \mathbb{I}_{r-1}} \prod_{k=1}^{r-1} 1_{\{t_{l,k}=I_k\}} = 1$ shows that the expression to the right in the previous display equals

$$\sum_{I \in \mathbb{I}_{r-1}} \mathbb{E}\left[\prod_{j=1}^{r-1} 1_{\{t_{l,j}=I_j\}} 1_{A_j}(Y_{l,I_j})\right] = \sum_{I \in \mathbb{I}_{r-1}} \mathbb{E}\left[\prod_{j=1}^{r-1} 1_{A_j}(Y_{l,t_{l,j}}) \prod_{k=1}^{r-1} 1_{\{t_{l,k}=I_k\}}\right] = \mathbb{E}\left(\prod_{j=1}^{r-1} 1_{A_j}(Y_{l,t_{l,j}})\right),$$

the latter being equal to $\mathbb{P}(Y_{l,t_{l,1}} \in A_1, \ldots, Y_{l,t_{l,r-1}} \in A_{r-1})$. □

Finally, to obtain the upper bounds claimed in Equation (109), note first that

$$\begin{aligned}\mathbb{P}(B_{i,t}) &= \mathbb{P}(\mathsf{T}(\hat{F}_{i,t-1}) > \mathsf{T}(F^i) + C\sqrt{\beta \log(t)/2S_i(t-1)}) \\ &= \sum_{s=1}^{t} \mathbb{P}(\mathsf{T}(\hat{F}_{i,t-1}) > \mathsf{T}(F^i) + C\sqrt{\beta \log(t)/2s}, S_i(t-1) = s).\end{aligned}$$

Note further that for every index s, on the event $\{S_i(t-1) = s\}$ the empirical cdf $\hat{F}_{i,t-1}$ coincides by definition with $s^{-1}\sum_{j=1}^{s} 1_{\{Y_{i,t_{i,j}} \leq \cdot\}}$. Hence, the sum in the second line of the previous display is not greater than

$$\sum_{s=1}^{t} \mathbb{P}\left[\mathsf{T}\left(s^{-1}\sum_{j=1}^{s} 1_{\{Y_{i,t_{i,j}} \leq \cdot\}}\right) > \mathsf{T}(F^i) + C\sqrt{\beta \log(t)/2s}\right].$$

By Lemma D.2 the joint distribution of $Y_{i,t_{i,1}}, \ldots, Y_{i,t_{i,s}}$ coincides with the joint distribution of $Y_{i,1}, \ldots, Y_{i,s}$. It thus follows that we can replace $Y_{i,t_{i,1}}, \ldots, Y_{i,t_{i,s}}$ by $Y_{i,1}, \ldots, Y_{i,s}$ in the previous display without changing the probabilities. In other words, we can replace $s^{-1}\sum_{j=1}^{s} 1_{\{Y_{i,t_{i,j}} \leq \cdot\}}$ by $F_{i,s}$ in the previous display, from which the upper bound on $\mathbb{P}(B_{i,t})$ in Equation (109) follows. The upper bound on $\mathbb{P}(A_t)$ is obtained analogously.

D.1.2 Proof of Theorem 3.6

We begin with a lemma that provides an upper bound on the KL divergence between members of $\mathcal{H}$ as defined in Definition 3.2. In particular, the KL-divergence between two elements $\mathbb{P}_{h_a}$ and $\mathbb{P}_{h_b}$ of $\mathcal{H}$ is *sub-quadratic* in the distance between a and b.

Lemma D.3. *The KL divergence between any two elements of $\mathcal{H}$ (cf. Definition 3.2) satisfies:*

$$KL(h_a, h_b) = \int_0^1 \log\left(\frac{h_a(y)}{h_b(y)}\right) h_a(y)dy \leq \frac{1}{(1+b)(1+a)}(b-a)^2$$

Proof. By simple calculus we obtain that

$$\begin{aligned}
\int_0^1 \log\left(\frac{h_a(y)}{h_b(y)}\right) h_a(y)dy &= \int_0^1 \log\left(\frac{(1+a)y^a}{(1+b)y^b}\right)(1+a)y^a dy \\
&= \int_0^1 \left[\log\left(\frac{1+a}{1+b}\right) + (a-b)\log(y)\right](1+a)y^a dy \\
&\leq \frac{a-b}{1+b} + (a-b)(1+a)\int_0^1 \log(y)y^a dy \\
&= \frac{a-b}{1+b} - \frac{(a-b)(1+a)}{(1+a)^2} \\
&= \frac{a-b}{1+b} - \frac{(a-b)}{(1+a)} = \frac{(a-b)^2}{(1+b)(1+a)}.
\end{aligned}$$

□

Proof of Theorem 3.6. Throughout the proof we fix a policy π and assume without loss of generality that $\mathsf{T}(H_{a_2}) - \mathsf{T}(H_{a_2}) \geq c(a_2 - a_1)$ for all $a_1, a_2 \in [\bar{a} - \delta, \bar{a} + \delta] \subseteq (-1, \infty)$ and $a_2 \geq a_1$. The case where $a \mapsto \mathsf{T}(H_a)$ is locally uniformly decreasing follows analogously.

Let treatment 1 have distribution $\mathbb{P}_{\bar{a}}$ with cdf $H_{\bar{a}}$ and treatment 2 have distribution $\mathbb{P}_{\bar{a}-\varepsilon}$ with cdf $H_{\bar{a}-\varepsilon}$ or $\mathbb{P}_{\bar{a}+\varepsilon}$ with cdf $H_{\bar{a}+\varepsilon}$ for some $\varepsilon > 0$. It suffices to show that the maximal regret incurred over the two two-tuples $(H_{\bar{a}}, H_{\bar{a}-\varepsilon})$ and $(H_{\bar{a}}, H_{\bar{a}+\varepsilon})$ is greater than $\underline{c}\sqrt{n}$ for some $\underline{c} > 0$. Denote by $\mathbb{P}^t_{\pi,-\varepsilon}$

and $\mathbb{P}^t_{\pi,\varepsilon}$, respectively, the distribution of $(Y_{\pi_t(Z_{t-1}),t}, ..., Y_{\pi_1,1})$ under the relevant tuple. Since

$$\begin{aligned}\sup_{j\in\{-\varepsilon,\varepsilon\}} \mathbb{E}^n_{\pi,j} R_n(\pi) &\geq \frac{1}{2}\left(\mathbb{E}^n_{\pi,-\varepsilon} R_n(\pi) + \mathbb{E}^n_{\pi,\varepsilon} R_n(\pi)\right) \\ &= \frac{1}{2}\left(\sum_{t=1}^n \sum_{i=1}^2 \Delta_i \mathbb{E}^n_{\pi,-\varepsilon} 1_{\{\pi_t(Z_{t-1})=i\}} + \sum_{t=1}^n \sum_{i=1}^2 \Delta_i \mathbb{E}^n_{\pi,\varepsilon} 1_{\{\pi_t(Z_{t-1})=i\}}\right) \\ &\geq \frac{c\varepsilon}{2}\left(\sum_{t=1}^n \mathbb{E}^n_{\pi,-\varepsilon} 1_{\{\pi_t(Z_{t-1})=2\}} + \sum_{t=1}^n \mathbb{E}^n_{\pi,\varepsilon} 1_{\{\pi_t(Z_{t-1})=1\}}\right).\end{aligned}$$

where the third estimate used that $\mathsf{T}(H_{\bar{a}+\varepsilon}) - \mathsf{T}(H_{\bar{a}})$ and $\mathsf{T}(H_{\bar{a}}) - \mathsf{T}(H_{\bar{a}-\varepsilon})$ are bounded from below by $c\varepsilon$ for $\varepsilon \leq \delta$. Next, note that $\mathbb{E}^n_{\pi,-\varepsilon} 1_{\{\pi_t(Z_{t-1})=2\}} + \mathbb{E}^n_{\pi,\varepsilon} 1_{\{\pi_t(Z_{t-1})=1\}} = \mathbb{E}^n_{\pi,-\varepsilon} 1_{\{\pi_t(Z_{t-1})=2\}} + 1 - \mathbb{E}^n_{\pi,\varepsilon} 1_{\{\pi_t(Z_{t-1})=2\}}$ is the sum of type 1 and 2 errors for the testing problem $H_0 : \mathbb{P} = \mathbb{P}^n_{\pi,-\varepsilon}$ vs $H_a : \mathbb{P} = \mathbb{P}^n_{\pi,\varepsilon}$ for the test $1_{\{\pi_t(Z_{t-1})=2\}}$. Thus, using Theorem 2.2(iii) of Tsybakov (2009), we get for $t = 2, ..., n$

$$\mathbb{E}^n_{\pi,-\varepsilon} 1_{\{\pi_t(Z_{t-1})=2\}} + \mathbb{E}^n_{\pi,\varepsilon} 1_{\{\pi_t(Z_{t-1})=1\}} \geq \frac{1}{4}\exp\left(-KL(\mathbb{P}^n_{\pi,-\varepsilon}, \mathbb{P}^n_{\pi,\varepsilon})\right).$$

Using the chain rule for Kullback-Leibler divergence, cf. Theorem 2.5.3 of Cover and Thomas (2012),

$$KL(\mathbb{P}^n_{\pi,-\varepsilon}, \mathbb{P}^n_{\pi,\varepsilon}) = KL(\mathbb{P}^{n-1}_{\pi,-\varepsilon}, \mathbb{P}^{n-1}_{\pi,\varepsilon}) + \mathbb{E}^{n-1}_{\pi,-\varepsilon} KL(\mathbb{P}_{\pi,-\varepsilon,n}, \mathbb{P}_{\pi,\varepsilon,n})$$

where $\mathbb{P}_{\pi,j,n},\ j \in \{-\varepsilon, \varepsilon\}$ is the conditional distribution of $Y_{\pi_n(Z_{n-1}),n}$ given Z_{n-1} under the policy π and distribution $\mathbb{P}^n_{\pi,j}$. Since $Y_{1,n}$ and $Y_{2,n}$ are independent of Z_{n-1}, we observe

$$\mathbb{P}_{\pi,j,n} = \mathbb{P}_{\bar{a}} 1_{\{\pi_n(Z_{n-1})=1\}} + \mathbb{P}_{\bar{a}+j} 1_{\{\pi_n(Z_{n-1})=2\}}.$$

Hence, by Lemma D.3,

$$KL(\mathbb{P}_{\pi,-\varepsilon,n}, \mathbb{P}_{\pi,\varepsilon,n}) \leq KL(H_{\bar{a}-\varepsilon}, H_{\bar{a}+\varepsilon}) 1_{\{\pi_n(Z_{n-1})=2\}} \leq \frac{4\varepsilon^2}{(1+\bar{a}-\delta)^2} 1_{\{\pi_n(Z_{n-1})=2\}}.$$

Thus, by induction, we observe that

$$KL(\mathbb{P}^n_{\pi,-\varepsilon}, \mathbb{P}^n_{\pi,\varepsilon}) \leq \frac{4\varepsilon^2}{(1+\bar{a}-\delta)^2} N_\pi = \tilde{c}\varepsilon^2 N_\pi,$$

with $\tilde{c} = \frac{4}{(1+\bar{a}-\delta)^2}$ and $N_\pi = \mathbb{E}^n_{\pi,-\varepsilon} \sum_{t=1}^n 1_{\{\pi_t(Z_{t-1})=2\}}$. But since we also have $\sup_{j\in\{-\varepsilon,\varepsilon\}} \mathbb{E}^n_{\pi,j} R_n(\pi) \geq \frac{c\varepsilon}{2} N_\pi$ we conclude that

$$\begin{aligned}\sup_{j\in\{-\varepsilon,\varepsilon\}} \mathbb{E}^n_{\pi,j} R_n(\pi) &\geq \frac{c\varepsilon}{2} \max\left(\frac{n}{4}\exp(-\tilde{c}\varepsilon^2 N_\pi), N_\pi\right) \\ &\geq \frac{c\varepsilon}{4}\left(\frac{n}{4}\exp(-\tilde{c}\varepsilon^2 N_\pi) + N_\pi\right) \\ &\geq \frac{c\varepsilon}{4} \inf_{z\geq 0}\left(\frac{n}{4}\exp(-\tilde{c}\varepsilon^2 z) + z\right).\end{aligned}$$

Note that

$$z^* := \operatorname{argmin}\left(\frac{n}{4}\exp(-\tilde{c}\varepsilon^2 z) + z\right) = \log\left(\frac{\tilde{c}\varepsilon^2 n}{4}\right)/(\tilde{c}\varepsilon^2),$$

is positive if $\varepsilon^2 > \frac{4}{n\tilde{c}}$. Thus, choosing $\varepsilon = \sqrt{\frac{8}{n\tilde{c}}}$ (which is less than δ for $n \geq n_0 = \lceil \frac{8}{\delta^2 \tilde{c}} \rceil$) shows

$$\sup_{j \in \{-\varepsilon, \varepsilon\}} \mathbb{E}^n_{\pi,j} R_n(\pi) \geq \frac{c}{\sqrt{\tilde{c}}} \frac{\log(2)}{\sqrt{128}} \cdot \sqrt{n}.$$

The first n_0 terms are handled by using $\sup_{j\in\{-\varepsilon,\varepsilon\}} \mathbb{E}^n_{\pi,j} R_n(\pi) \geq \frac{c\varepsilon}{4} \inf_{z\geq 0}\left(\frac{n}{4}\exp(-\tilde{c}\varepsilon^2 z) + z\right)$ with $\varepsilon = \delta/2$ and by choosing the constant in the statement of the theorem small enough. □

D.2 Proofs of results in Section 4

First, we provide an auxiliary result that will be useful in the proofs of Theorems 4.4 and 4.7. In the local treatment problem for individuals with covariates in B_j, $*$ denotes the index of a treatment in the set $\arg\max \mathsf{T}(F^i_j)$. That is, $\mathsf{T}(F^*_j) = \max_{i\in\mathcal{I}} \mathsf{T}(F^i_j)$. To save on notation, we shall write $\pi_t(X_t)$ instead of $\pi_t(X_t, Z_{t-1})$ throughout this section.

Lemma D.4. *Suppose that Assumption 2.1 and 4.2 are satisfied and a grouping is characterised by* $\{V_1, ..., V_F\}$ *and* $\{\bar{B}_1, ..., \bar{B}_F\}$. *Then, for any* $i \in \mathcal{I}$, $j \in \{1, \ldots, F\}$ *and* $x, \tilde{x} \in B_j$, *we obtain that*

$$\begin{aligned}
|\mathsf{T}(F^i(\cdot, x)) - \mathsf{T}(F^i(\cdot, \tilde{x}))| &\leq CLV_j^\gamma, \\
|\mathsf{T}\big(F^{\pi^\star(x)}(\cdot, x)\big) - \mathsf{T}\big(F^{\pi^\star(\tilde{x})}(\cdot, \tilde{x})\big)| &\leq CLV_j^\gamma, \\
|\mathsf{T}(F^i_j) - \mathsf{T}(F^i(\cdot, x))| &\leq CLV_j^\gamma, \\
|\mathsf{T}(F^\star(\cdot, x)) - \mathsf{T}(F^*_j)| &\leq CLV_j^\gamma.
\end{aligned}$$

Proof. Fix i, j and $x, \tilde{x} \in B_j$. Assumption 4.2 implies that

$$||F^i(\cdot, x) - F^i(\cdot, \tilde{x})||_\infty \leq L||x - \tilde{x}||^\gamma \leq LV_j^\gamma.$$

Then, the first statement follows immediately from Assumption 2.1. Using this result, we also obtain the second statement via

$$\begin{aligned}
|\mathsf{T}\big(F^{\pi^\star(x)}(\cdot, x)\big) - \mathsf{T}\big(F^{\pi^\star(\tilde{x})}(\cdot, \tilde{x})\big)| &= |\max_{i\in\mathcal{I}} \mathsf{T}(F^i(\cdot, x)) - \max_{i\in\mathcal{I}} \mathsf{T}(F^i(\cdot, \tilde{x}))| \\
&\leq \max_{i\in\mathcal{I}} |\mathsf{T}(F^i(\cdot, x)) - \mathsf{T}(F^i(\cdot, \tilde{x}))| \leq CLV_j^\gamma.
\end{aligned}$$

Now, we move to the third part. For any s and x, we have that

$$F^i_j(y) - F^i(y, x) = \frac{1}{\mathbb{P}_X(B_j)} \int_{B_j} (F^i(y, s) - F^i(y, x))\mathbb{P}_X(ds).$$

As $x, s \in B_j$, Assumption 4.2 leads to $|F^i(y,s) - F^i(y,x)| \le LV_j^\gamma$ and hence $||F_j^i - F^i(\cdot,x)||_\infty \le LV_j^\gamma$ for each $x \in B_j$. Then, the third claim is the direct consequence of Assumption 2.1 on T.

Concerning the last statement, we observe that

$$\begin{aligned}|\mathsf{T}(F^\star(\cdot,x)) - \mathsf{T}(F_j^*)| &= \big|\max_{i\in\mathcal{I}} \mathsf{T}(F^i(\cdot,x)) - \max_{i\in\mathcal{I}} \mathsf{T}(F_j^i)\big| \\ &\le \max_{i\in\mathcal{I}} |\mathsf{T}(F^i(\cdot,x)) - \mathsf{T}(F_j^i)| \le CLV_j^\gamma,\end{aligned}$$

which finishes the proof. □

Now, we are ready to deal with the proofs of main results.

Proof of Theorem 4.4. First, we write $R_N(\bar\pi) = \sum_{j=1}^F \tilde R_j(\bar\pi)$ with

$$\tilde R_j(\bar\pi) := \sum_{t=1}^N \big[\mathsf{T}\big(F^{\pi^\star(X_t)}(\cdot,X_t)\big) - \mathsf{T}\big(F^{\bar\pi_t(X_t)}(\cdot,X_t)\big)\big]1_{\{X_t\in B_j\}}.$$

Fix $j \in \{1,\dots,F\}$. Recalling the definition of F_j^i in (9), each summand in the previous display can be written as

$$\left[\mathsf{T}\left(F^{\pi^\star(X_t)}(\cdot,X_t)\right) - \mathsf{T}(F_j^*) + \mathsf{T}(F_j^*) - \mathsf{T}(F_j^{\bar\pi_t(X_t)}) + \mathsf{T}(F_j^{\bar\pi_t(X_t)}) - \mathsf{T}\left(F^{\bar\pi_t(X_t)}(\cdot,X_t)\right)\right]1_{\{X_t\in B_j\}}, \quad (124)$$

which by Lemma D.4 is not greater than $\mathsf{T}(F_j^*) - \mathsf{T}(F_j^{\bar\pi_t(X_t)}) + 2CLV_j^\gamma$. Therefore, we obtain

$$\tilde R_j(\bar\pi) \le \sum_{t=1}^N \left[\mathsf{T}\big(F_j^*\big) - \mathsf{T}\big(F_j^{\bar\pi_t(X_t)}\big)\right]1_{\{X_t\in B_j\}} + 2CLV_j^\gamma \sum_{t=1}^N 1_{\{X_t\in B_j\}}. \quad (125)$$

By Wald's identity $\mathbb{E}(\sum_{t=1}^N 1_{\{X_t\in B_j\}}) \le n\bar c\bar B_j$. Hence, to prove the theorem, it remains to show that for some c (which in fact will be the same $c(\beta, C)$ as in Theorem 3.1) it holds that

$$\mathbb{E}\left(\sum_{t=1}^N \left[\mathsf{T}\big(F_j^*\big) - \mathsf{T}\big(F_j^{\bar\pi_t(X_t)}\big)\right]1_{\{X_t\in B_j\}}\right) \le c\sqrt{K\bar c\bar B_j n\overline{\log}(\bar c\bar B_j n)}. \quad (126)$$

To this end, for every $m \in \mathbb{N}$ and every $v = (v_1,\dots,v_m) \in \{1,\dots,F\}^m$, we define the event

$$\Omega(m,v) := \{N = m, X_1 \in B_{v_1},\dots,X_m \in B_{v_m}\} \subseteq \Omega. \quad (127)$$

Note that $\{\Omega(m,v) : m \in \mathbb{N}, v \in \{1,\dots,F\}^m\}$ defines a partition of Ω. For the sake of brevity, set $f := \sum_{t=1}^N [\mathsf{T}\big(F_j^*\big) - \mathsf{T}\big(F_j^{\bar\pi_t(X_t)}\big)]1_{\{X_t\in B_j\}}$. Then, note that by Tonelli's theorem we can write

$$\mathbb{E}(f) = \sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}} \mathbb{E}(1_{\Omega(m,v)}f) = \sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}} \mathbb{P}(\Omega(m,v))\mathbb{E}(f|\Omega(m,v)), \quad (128)$$

where we define $\mathbb{E}(f|\Omega(m,v)) := \mathbb{P}^{-1}(\Omega(m,v))\mathbb{E}(1_{\Omega(m,v)}f)$ in case $\mathbb{P}(\Omega(m,v)) > 0$ and $\mathbb{E}(f|\Omega(m,v)) := 0$ else. Fix m and v, and assume that $\{v_s : s \in \{1,\ldots,m\}, v_s = j\}$ is not empty. Denote the elements of the latter set by $t_1,\ldots,t_{\bar{m}}$, ordered from smallest to largest. On the event $\Omega(m,v)$ (i.e., for every $\omega \in \Omega(m,v)$) we can use the definition of $\bar{\pi}$ to rewrite (recall the definition of the FSA policy $\hat{\pi}$ in the no-covariate case)

$$f = \sum_{s=1}^{\bar{m}} \left[\mathsf{T}(F_j^*) - \mathsf{T}\left(F_j^{\hat{\pi}_s(Z_{s-1})}\right)\right], \tag{129}$$

where for every $s > 1$ we define $Z_{s-1} = (Y_{\hat{\pi}_{s-1},t_{s-1}},\ldots,Y_{\hat{\pi}_1,t_1})$, and for $s = 1$ we recall from the definition of the FSA policy that $\hat{\pi}_1 := 1$, which is deterministic. The previous display shows that on the event $\Omega(m,v)$ we can write f as a function of $(Y_{t_1},\ldots,Y_{t_{\bar{m}}})$, i.e., as $H(Y_{t_1},\ldots,Y_{t_{\bar{m}}})$, say. We conclude that

$$\mathbb{E}(f|\Omega(m,v)) = \mathbb{E}\left(\sum_{s=1}^{\bar{m}} \left[\mathsf{T}(F_j^*) - \mathsf{T}\left(F_j^{\hat{\pi}_s(Z_{s-1})}\right)\right]\Bigg|\Omega(m,v)\right) = \mathbb{E}\left(H(Y_{t_1},\ldots,Y_{t_{\bar{m}}})|\Omega(m,v)\right). \tag{130}$$

The quantity to the right equals $\mathbb{E}^*(H(Y_{t_1},\ldots,Y_{t_{\bar{m}}}))$, where the probability measure $\mathbb{P}^*$ corresponding to $\mathbb{E}^*$ is defined as the $\mathbb{P}$-measure with density $\mathbb{P}^{-1}(\Omega(m,v))1_{\Omega(m,v)}$. Note that for $A_i \in \mathcal{B}(\mathbb{R}^K)$ for $i = 1,\ldots,m$ we have that $\mathbb{P}^*(Y_{t_1} \in A_1,\ldots,Y_{t_{\bar{m}}} \in A_{\bar{m}})$ equals (using various independence properties of the observations and N)

$$\mathbb{P}^{-1}(\Omega(m,v))\mathbb{P}\left(Y_{t_1} \in A_1,\ldots,Y_{t_{\bar{m}}} \in A_{\bar{m}},\Omega(m,v)\right) = \prod_{s=1}^{\bar{m}} \frac{\mathbb{P}(Y_{t_s} \in A_s, X_{t_s} \in B_j)}{\mathbb{P}(X_{t_s} \in B_j)} \tag{131}$$

$$= \prod_{s=1}^{\bar{m}} \mathbb{P}(Y_{t_s} \in A_s|\{X_{t_s} \in B_j\}). \tag{132}$$

We thus see that $\mathbb{P}^*$ is the $\bar{m}$-fold product of $\mathbb{Q}(.) := \mathbb{P}(Y_1 \in .|\{X_1 \in B_j\})$. For i.i.d. random K-vectors $Y_1^*,\ldots,Y_{\bar{m}}^*$, say, each with distribution $\mathbb{Q}$ (which exist possibly after enlarging the underlying probability space), it hence follows from the definition of H that

$$\mathbb{E}(H(Y_{t_1},\ldots,Y_{t_m})|\Omega(m,v)) = \mathbb{E}(H(Y_1^*,\ldots,Y_{\bar{m}}^*)) = \mathbb{E}\left(\sum_{s=1}^{\bar{m}} \left[\mathsf{T}(F_j^*) - \mathsf{T}\left(F_j^{\hat{\pi}_s(Z_{s-1}^*)}\right)\right]\right) \tag{133}$$

where $Z_{s-1}^* = (Y_{\hat{\pi}_{s-1},s-1}^*,\ldots,Y_{\hat{\pi}_1,1}^*)$ for $s > 1$. Noting that the r-th marginal of $\mathbb{Q}$ has cdf F_j^r, it now follows from Theorem 3.1 applied with marginal distribution $\mathbb{Q}$ and (constant) $N = \bar{m}$ that the quantity in the previous display, and thus $\mathbb{E}(f|\Omega(m,v))$ is not greater than $c\sqrt{K\bar{m}\overline{\log}(\bar{m})}$. It thus follows from (128) (noting that f vanishes on those exceptional sets $\Omega(m,v)$ for which the set $\{v_s : s \in \{1,\ldots,m\}, v_s = j\}$ is empty, and to which the just derived upper bound does not apply) that

$$\mathbb{E}(f) \leq c \sum_{\substack{m\in\mathbb{N} \\ v\in\{1,\ldots,F\}^m}} \mathbb{P}(\Omega(m,v))\sqrt{K\bar{m}\overline{\log}(\bar{m})}. \tag{134}$$

Recall, that $\bar{m} = |\{v_s : s \in \{1,\ldots,m\}, v_s = j\}|$. Hence, we can interpret $\bar{m}$ as a random variable on the set of all tuples (m,v), over which the sum in the previous display extends, equipped with the probability measure $\mathbb{P}(\Omega(m,v))$. It remains to observe that the function h defined via $x \mapsto \sqrt{Kx\overline{\log}(x)}$ is concave on $[0,\infty)$, which allows us to apply Jensen's inequality to upper bound the right hand side in the previous display by $c\sqrt{Kx\overline{\log}(x)}$ with

$$x = \sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}} \mathbb{P}(\Omega(m,v))\bar{m} = \mathbb{E}\big(\sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}} 1_{\Omega(m,v)}\bar{m}\big) = \mathbb{E}\big(\sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}} \big[1_{\Omega(m,v)}\sum_{s=1}^{N} 1_{X_s\in B_j}\big]\big) \tag{135}$$

$$= \mathbb{E}\big(\sum_{s=1}^{N}\sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}} 1_{\Omega(m,v)}1_{X_s\in B_j}\big) = \mathbb{E}\big(\sum_{s=1}^{N} 1_{X_s\in B_j}\sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}} 1_{\Omega(m,v)}\big) = \mathbb{E}\big(\sum_{s=1}^{N} 1_{X_s\in B_j}\big). \tag{136}$$

We used Tonelli's theorem in the second equality. We know already that $x \le \bar{c}\bar{B}_j n$. Since the function h is also monotonically increasing, it follows that $\mathbb{E}(f) \le ch(\bar{c}\bar{B}_j n)$, which is the statement in Equation (126). □

Proof of Corollary 4.5. The given choice of groups results in $F = P^d$, $\bar{B}_j = P^{-d}$ and $V_j = \sqrt{d}P^{-1}$. Hence, Theorem 4.4 and the choice $P = \lceil n^{\frac{1}{2\gamma+d}}\rceil$ yields

$$\begin{aligned}\mathbb{E}[R_N(\bar{\pi})] &\le cP^d\big(\sqrt{KnP^{-d}\overline{\log}(nP^{-d})} + nP^{-\gamma-d}\big)\\ &\le c\sqrt{K\overline{\log}(n)}P^d\big(\sqrt{nP^{-d}} + nP^{-\gamma-d}\big)\\ &\le c\sqrt{K\overline{\log}(n)}n^{1-\frac{\gamma}{2\gamma+d}}\end{aligned}$$

which is the claimed result. □

Proof of Theorem 4.7. Define $c_1 := 4CLd^{\gamma/2} + 1$. Recalling $P = \lceil n^{1/(2\gamma+d)}\rceil$, we shall assume without loss of generality that n is large enough ($n \ge n_0$, say) such that $c_1 P^{-\gamma} \le 1$ holds (this will allow us to apply Assumption 4.6 with $\delta = c_1 P^{-\gamma}$ in the arguments below). Note that by Assumption 2.1 for $n < n_0$ it holds that $\mathbb{E}[R_N(\pi)] \le Cn_0$. Hence, c in the statement of Theorem 4.7 can be chosen large enough to deal with the initial terms smaller than n_0. Throughout the proof the bins are sorted in the lexicographic order and we shall write $B_1, ..., B_{P^d}$ for the P^d bins. The proof is divided into several steps:

Step 1: Decomposition of bins into different types. To obtain the desired upper bound, we shall treat three types of bins separately. This division of bins was also used in Perchet and Rigollet (2013) to establish the properties of their successive elimination algorithm in a classic bandit problem targeting the distribution with the highest (conditional) mean. The bins are split into

$$\begin{aligned}\mathcal{J} &:= \Big\{j \in \{1,\dots,P^d\} : \exists\, \bar{x} \in B_j, \mathsf{T}(F^{\pi^\star(\bar{x})}(\cdot,\bar{x})) - \mathsf{T}(F^{\pi^\sharp(\bar{x})}(\cdot,\bar{x})) > c_1 P^{-\gamma}\Big\},\\ \mathcal{J}_s &:= \Big\{j \in \{1,\dots,P^d\} : \exists\, \bar{x} \in B_j, \mathsf{T}(F^{\pi^\star(\bar{x})}(\cdot,\bar{x})) = \mathsf{T}(F^{\pi^\sharp(\bar{x})}(\cdot,\bar{x}))\Big\},\\ \mathcal{J}_w &:= \Big\{j \in \{1,\dots,P^d\} : 0 < \mathsf{T}(F^{\pi^\star(x)}(\cdot,x)) - \mathsf{T}(F^{\pi^\sharp(x)}(\cdot,x)) \le c_1 P^{-\gamma} \text{ for all } x \in B_j\Big\}.\end{aligned} \tag{137}$$

The bins corresponding to indices in $\mathcal{J}$, $\mathcal{J}_s$, and $\mathcal{J}_w$ will be referred to as "well-behaved", "strongly ill-behaved" and "weakly ill-behaved" bins, respectively. Note that $\mathcal{J}_w$ and $\mathcal{J} \cup \mathcal{J}_s$ are clearly disjoint. That $\mathcal{J}$ and $\mathcal{J}_s$ are disjoint is shown in Step 2 below. Hence, the sets of bins corresponding to indices in $\mathcal{J}$, $\mathcal{J}_s$, $\mathcal{J}_w$ constitute a partition of the set of all P^d bins B_j, and we can thus write

$$\mathbb{E}(R_N(\bar{\pi})) = \sum_{j\in\mathcal{J}_s} \mathbb{E}(\tilde{R}_j(\bar{\pi})) + \sum_{j\in\mathcal{J}_w} \mathbb{E}(\tilde{R}_j(\bar{\pi})) + \sum_{j\in\mathcal{J}} \mathbb{E}(\tilde{R}_j(\bar{\pi})), \tag{138}$$

where

$$\tilde{R}_j(\bar{\pi}) := \sum_{t=1}^{N} \Big[\mathsf{T}\big(F^{\pi^\star(X_t)}(\cdot, X_t)\big) - \mathsf{T}\big(F^{\bar{\pi}_t(X_t)}(\cdot, X_t)\big) \Big] 1_{\{X_t \in B_j\}}. \tag{139}$$

Step 2: Strongly ill-behaved bins. For every $j \in \mathcal{J}_s$, by definition, there exists a $\bar{x} \in B_j$ such that $\mathsf{T}\big(F^{\pi^\star(\bar{x})}(\cdot, \bar{x})\big) = \mathsf{T}\big(F^{\pi^\sharp(\bar{x})}(\cdot, \bar{x})\big)$. From the definition of $\pi^\sharp$ it thus follows that $\mathsf{T}\big(F^{\pi^\star(\bar{x})}(\cdot, \bar{x})\big) = \mathsf{T}\big(F^{i}(\cdot, \bar{x})\big)$ for every $i \in \mathcal{I}$. Therefore, for every $x \in B_j$ and every $i \in \mathcal{I}$, Lemma D.4 yields

$$\begin{aligned} \mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^{i}(\cdot, x)) &= \mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^{i}(\cdot, x)) - [\mathsf{T}(F^{\pi^\star(\bar{x})}(\cdot, \bar{x})) - \mathsf{T}(F^{i}(\cdot, \bar{x}))] && (140)\\ &\le 2CLd^{\gamma/2}P^{-\gamma} \le c_1 P^{-\gamma}. && (141) \end{aligned}$$

First of all, this shows that $\mathcal{J}$ and $\mathcal{J}_s$ are disjoint. Furthermore, from Equations (139) and (140), we obtain

$$\begin{aligned} \sum_{j\in\mathcal{J}_s} \tilde{R}_j(\bar{\pi}) &\le c_1 P^{-\gamma} \sum_{j\in\mathcal{J}_s} \sum_{t=1}^{N} 1_{\{X_t\in B_j\}} 1_{\{0<\mathsf{T}(F^{\pi^\star(X_t)}(\cdot,X_t))-\mathsf{T}(F^{\pi^\sharp(X_t)}(\cdot,X_t))\}} && (142)\\ &\le c_1 P^{-\gamma} \sum_{t=1}^{N} 1_{\{0<\mathsf{T}(F^{\pi^\star(X_t)}(\cdot,X_t))-\mathsf{T}(F^{\pi^\sharp(X_t)}(\cdot,X_t))\le c_1P^{-\gamma}\}}. && (143) \end{aligned}$$

From Condition 4.6 we hence obtain:

$$\sum_{j\in\mathcal{J}_s} \mathbb{E}[\tilde{R}_j(\bar{\pi})] \le c_1 n P^{-\gamma} \mathbb{P}_X\big(0 < \mathsf{T}\big(F^{\pi^\star(X)}(\cdot, X)\big) - \mathsf{T}\big(F^{\pi^\sharp(X)}(\cdot, X) \le c_1 P^{-\gamma}\big) \le C_0 c_1^{1+\alpha} n P^{-\gamma(1+\alpha)}. \tag{144}$$

Step 3: Weakly ill-behaved bins. Since $\{X_t \in B_j\}$ for $j \in \mathcal{J}_w$ are disjoint subsets of

$$\{0 < \mathsf{T}(F^{\pi^\star(X_t)}(\cdot, X_t)) - \mathsf{T}(F^{\pi^\sharp(X_t)}(\cdot, X_t)) \le c_1 P^{-\gamma}\},$$

we obtain from Condition 4.6, recall that $\mathbb{P}(X_t \in B_j) \ge \frac{\underline{c}}{P^d}$, that

$$|\mathcal{J}_w| \frac{\underline{c}}{P^d} \le \sum_{j\in\mathcal{J}_w} \mathbb{P}(X_t \in B_j) \le \mathbb{P}_X\big(0 < \mathsf{T}\big(F^{\pi^\star(X)}(\cdot, X)\big) - \mathsf{T}\big(F^{\pi^\sharp(X)}(\cdot, X)\big) \le c_1 P^{-\gamma}\big) \le C_0 c_1^{\alpha} P^{-\gamma\alpha}, \tag{145}$$

which yields $|\mathcal{J}_w| \le (C_0 c_1^\alpha / \underline{c}) P^{d-\gamma\alpha}$. Using (125) and (126) with $V_j = \sqrt{d}P^{-1}$ and $\bar{B}_j = P^{-d}$ we obtain

$$\mathbb{E}[\tilde{R}_j(\bar{\pi})] \le c' \left(\sqrt{Kn\overline{\log}(n)} P^{-d/2} + nP^{-\gamma-d} \right), \tag{146}$$

where c' depends on $d, L, \gamma, \bar{c}, C, \beta$, but *not* on n. Combining this with (146) leads to

$$\sum_{j\in\mathcal{J}_w} \mathbb{E}[\tilde{R}_j(\bar{\pi})] \le c''\big(\sqrt{Kn\overline{\log}(n)}P^{d/2-\gamma\alpha} + nP^{-\gamma(1+\alpha)}\big), \tag{147}$$

where c'' depends on $d, L, \gamma, \underline{c}, \bar{c}, C, C_0, \alpha, \beta$, but *not* on n.

Step 4: Well-behaved bins. For every $j \in \mathcal{J}$ let $x_j \in B_j$ be such that

$$\mathsf{T}(F^{\pi^\star(x_j)}(\cdot, x_j)) - \mathsf{T}(F^{\pi^\sharp(x_j)}(\cdot, x_j)) > c_1 P^{-\gamma}. \tag{148}$$

Next, define the following sets of indices ("corresponding to the optimal and suboptimal arms given x_j"):

$$I_j^\star := \{i \in \mathcal{I} : \mathsf{T}\big(F^{\pi^\star(x_j)}(\cdot, x_j)\big) = \mathsf{T}(F^i(\cdot, x_j))\},$$
$$I_j^0 := \{i \in \mathcal{I} : \mathsf{T}\big(F^{\pi^\star(x_j)}(\cdot, x_j)\big) - \mathsf{T}(F^i(\cdot, x_j)) > c_1 P^{-\gamma}\}.$$

Clearly $\pi^\star(x_j) \in I_j^\star$ and $\pi^\sharp(x_j) \in I_j^0$ (cf. (148)). Hence $I_j^\star$ and I_j^0 define a nontrivial partition of $\mathcal{I}$. For every $j \in \mathcal{J}$ we can thus decompose $\tilde{R}_j(\bar{\pi})$ defined in Equation (139) as the sum of

$$\begin{aligned} \tilde{R}_{j,I_j^\star}(\bar{\pi}) &:= \sum_{i\in I_j^\star}\sum_{t=1}^{N} \Big[\mathsf{T}\big(F^{\pi^\star(X_t)}(\cdot, X_t)\big) - \mathsf{T}\big(F^i(\cdot, X_t)\big)\Big] \mathbb{1}_{\{X_t\in B_j\}}\mathbb{1}_{\{\bar{\pi}_t(X_t)=i\}}, \\ \tilde{R}_{j,I_j^0}(\bar{\pi}) &:= \sum_{i\in I_j^0}\sum_{t=1}^{N} \Big[\mathsf{T}\big(F^{\pi^\star(X_t)}(\cdot, X_t)\big) - \mathsf{T}\big(F^i(\cdot, X_t)\big)\Big] \mathbb{1}_{\{X_t\in B_j\}}\mathbb{1}_{\{\bar{\pi}_t(X_t)=i\}}. \end{aligned} \tag{149}$$

Step 4a: A bound for $\mathbb{E}(\tilde{R}_{j,I_j^\star}(\bar{\pi}))$. For any $i \in I_j^\star$ and every $x \in B_j$ satisfying $\mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) \neq \mathsf{T}(F^i(\cdot, x))$, the triangle inequality, the definition of $\pi^\sharp$, and Lemma D.4 yield

$$\begin{aligned} 0 < \mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^{\pi^\sharp(x)}(\cdot, x)) &\le \mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^i(\cdot, x)) \\ &= \mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^{\pi^\star(x_j)}(\cdot, x_j)) + \mathsf{T}(F^i(\cdot, x_j)) - \mathsf{T}(F^i(\cdot, x)) \\ &\le 2CLd^{\gamma/2}P^{-\gamma} \le c_1 P^{-\gamma}, \end{aligned}$$

the last inequality following from $c_1 = 4CLd^{\gamma/2} + 1$. But this means that for any $i \in I_j^\star$ and every $x \in B_j$

$$\mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^i(\cdot, x)) \le c_1 P^{-\gamma}\mathbb{1}_{\{0<\mathsf{T}(F^{\pi^\star(x)}(\cdot,x))-\mathsf{T}(F^{\pi^\sharp(x)}(\cdot,x))\le c_1P^{-\gamma}\}}. \tag{150}$$

We deduce

$$\mathbb{E}[\tilde{R}_{j,I_j^\star}(\bar{\pi})] \le \mathbb{E}\sum_{t=1}^{N} c_1 P^{-\gamma}\mathbb{1}_{\{0<\mathsf{T}(F^{\pi^\star(X_t)}(\cdot,X_t))-\mathsf{T}(F^{\pi^\sharp(X_t)}(\cdot,X_t))\le c_1P^{-\gamma}\}}\mathbb{1}_{\{X_t\in B_j\}} \le nc_1P^{-\gamma}q_j, \tag{151}$$

where $q_j := \mathbb{P}(0 < \mathsf{T}(F^{\pi^\star(X_t)}(\cdot, X_t)) - \mathsf{T}(F^{\pi^\sharp(X_t)}(\cdot, X_t)) \le c_1 P^{-\gamma}, X_t \in B_j)$, which is independent of t due to the X_t being identically distributed.

Step 4b: A bound for $\mathbb{E}(\tilde{R}_{j,I_j^0}(\bar{\pi}))$. By Lemma D.4 for every $x \in B_j$ and every $i \in I_j^0$ we have

$$\mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^i(\cdot, x)) \leq \left[\mathsf{T}(F_j^*) - \mathsf{T}(F_j^i)\right] + c_1 P^{-\gamma}, \tag{152}$$

from which it follows that

$$\begin{aligned}\mathbb{E}[\tilde{R}_{j,I_j^0}(\bar{\pi})] &\leq \mathbb{E}\sum_{i\in I_j^0}\sum_{t=1}^{N}\left[\mathsf{T}\big(F_j^*\big) - \mathsf{T}\big(F_j^i\big)\right] 1_{\{X_t\in B_j\}}1_{\{\bar{\pi}_t=i\}} + c_1 P^{-\gamma}\mathbb{E}\sum_{i\in I_j^0}\sum_{t=1}^{N} 1_{\{X_t\in B_j\}}1_{\{\bar{\pi}_t=i\}}, \\ &= \sum_{i\in I_j^0}\Delta_j^i \mathbb{E}S(i,N,j) + c_1 P^{-\gamma}\sum_{i\in I_j^0}\mathbb{E}S(i,N,j),\end{aligned} \tag{153}$$

where for every $i \in I_j^0$ the sum $S(i,N,j) := \sum_{t=1}^N 1_{\{X_t\in B_j\}}1_{\{\bar{\pi}_t=i\}}$, and where $\Delta_j^i := \mathsf{T}(F_j^*) - \mathsf{T}(F_j^i)$. We now claim that (this claim will be verified before moving to Step 4c below)

$$\mathbb{E}S(i,N,j) \leq \frac{2C^2\beta\log(\bar{c}nP^{-d})}{[\Delta_j^i]^2} + \frac{\beta+2}{\beta-2}. \tag{154}$$

Defining $\underline{\Delta}_j := \min_{i\in I_j^0}\Delta_j^i$, noting that $\max_{i\in I_j^0}\Delta_j^i \leq 2C$ by Assumption 2.1, and combining Equations (153) and (154) we obtain the bound

$$\mathbb{E}[\tilde{R}_{j,I_j^0}(\bar{\pi})] \leq K\frac{2C^2\beta\log(\bar{c}nP^{-d})}{\underline{\Delta}_j}\left(1 + \frac{c_1P^{-\gamma}}{\underline{\Delta}_j}\right) + (c_1+2C)K\frac{\beta+2}{\beta-2}. \tag{155}$$

Now, it remains to prove the claim in Equation (154). To this end we apply a conditioning argument as in the proof of Theorem 4.4. We shall now use some quantities (in particular the sets $\Omega(m,v)$) that were defined in that proof: note that

$$\mathbb{E}S(i,N,j) = \sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}}\mathbb{P}(\Omega(m,v))\mathbb{E}(S(i,N,j)|\Omega(m,v)). \tag{156}$$

Arguing as in the proof of Theorem 4.4, it is now easy to see that $\mathbb{E}(S(i,N,j)|\Omega(m,v))$ can be written as the expected number of times arm i is selected in running the FSA policy $\hat{\pi}$ (without covariates) in a problem with $\bar{m}$ (fixed) i.i.d. inputs with distribution $\mathbb{Q}$. We can hence apply the bound in Equation (102), to the just mentioned problem, to obtain that

$$\mathbb{E}(S(i,N,j)|\Omega(m,v)) \leq \frac{2C^2\beta\log(\bar{m})}{[\Delta_j^i]^2} + \frac{\beta+2}{\beta-2}. \tag{157}$$

We can now combine the obtained inequality with Equation (156) to see that

$$\mathbb{E}S(i,N,j) \leq \sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\dots,F\}^m}}\mathbb{P}(\Omega(m,v))\frac{2C^2\beta\log(\bar{m})}{[\Delta_j^i]^2} + \frac{\beta+2}{\beta-2}. \tag{158}$$

The claim in (154) now follows from Jensen's inequality, and (cf. the end of the proof of Theorem 4.4)

$$\sum_{\substack{m\in\mathbb{N}\\ v\in\{1,\ldots,F\}^m}} \mathbb{P}(\Omega(m,v))\bar{m} \le \bar{c}\bar{B}_j n = \bar{c}nP^{-d}. \tag{159}$$

Step 4c: A bound for $\mathbb{E}(\tilde{R}_j(\bar{\pi}))$ with $j \in \mathcal{J}$. For all $i \in I_j^0$ and all $x \in B_j$ the triangle inequality and Lemma D.4 with $V_j = \sqrt{d}P^{-1}$ yield

$$\begin{aligned} c_1 P^{-\gamma} <& |\mathsf{T}(F^{\pi^\star(x_j)}(\cdot,x_j)) - \mathsf{T}(F^i(\cdot,x_j))| \\ \le& |\mathsf{T}(F^{\pi^\star(x_j)}(\cdot,x_j)) - \mathsf{T}(F^{\pi^\star(x)}(\cdot,x))| + |\mathsf{T}(F^{\pi^\star(x)}(\cdot,x)) - \mathsf{T}(F^i(\cdot,x))| + |\mathsf{T}(F^i(\cdot,x)) - \mathsf{T}(F^i(\cdot,x_j))| \\ \le& 2CLd^{\gamma/2}P^{-\gamma} + |\mathsf{T}(F^{\pi^\star(x)}(\cdot,x)) - \mathsf{T}(F^i(\cdot,x))|. \end{aligned}$$

Recalling that $c_1 = 4CLd^{\gamma/2} + 1$, we obtain

$$\mathsf{T}\big(F^{\pi^\star(x)}(\cdot,x)\big) - \mathsf{T}(F^i(\cdot,x)) > (1 + 2CLd^{\gamma/2})P^{-\gamma}. \tag{160}$$

[In particular, since $I_j^0 \neq \emptyset$ holds, $0 < \mathsf{T}(F^{\pi^\star(x)}(\cdot,x)) - \mathsf{T}(F^{\pi^\sharp(x)}(\cdot,x))$ for all $x \in B_j$ if $j \in \mathcal{J}$, an observation we shall need later in Step 4d.] For each $i \in I_j^0$ and every $x \in B_j$, (160) and Lemma D.4 imply

$$\Delta_j^i = \mathsf{T}(F_j^*) - \mathsf{T}(F_j^i) \ge \mathsf{T}(F^{\pi^\star(x)}(\cdot,x)) - \mathsf{T}(F^i(\cdot,x)) - 2CLd^{\gamma/2}P^{-\gamma} > P^{-\gamma}; \tag{161}$$

in particular for any $j \in \mathcal{J}$ and any $i \in I_j^0$ we have $\underline{\Delta}_j = \min_{i\in I_j^0} \Delta_j^i > P^{-\gamma}$. Recalling that $\tilde{R}_j(\bar{\pi}) = \tilde{R}_{j,I_j^*}(\bar{\pi}) + \tilde{R}_{j,I_j^0}(\bar{\pi})$, we combine (151) and (155) (with the just observed $\underline{\Delta}_j > P^{-\gamma}$) to see that for any $j \in \mathcal{J}$

$$\mathbb{E}[\tilde{R}_j(\bar{\pi})] \le nc_1P^{-\gamma}q_j + \frac{2C^2(c_1+1)K\beta\log(\bar{c}nP^{-d})}{\underline{\Delta}_j} + (c_1+2C)K\frac{\beta+2}{\beta-2}. \tag{162}$$

Step 4d: A bound for $\sum_{j\in\mathcal{J}}\mathbb{E}[\tilde{R}_j(\bar{\pi})]$. Using Equation (162) and $|\mathcal{J}| \le P^d$ we obtain

$$\sum_{j\in\mathcal{J}}\mathbb{E}[\tilde{R}_j(\bar{\pi})] \le (c_1+2C)K\frac{\beta+2}{\beta-2}P^d + nc_1P^{-\gamma}\sum_{j\in\mathcal{J}}q_j + \sum_{j\in\mathcal{J}}\frac{2C^2(c_1+1)K\beta\log(\bar{c}nP^{-d})}{\underline{\Delta}_j}. \tag{163}$$

Since the B_j are disjoint, recalling the definition of q_j after Equation (151) we obtain

$$nc_1P^{-\gamma}\sum_{j\in\mathcal{J}}q_j \le c_1nP^{-\gamma}\mathbb{P}_X\big(0 < \mathsf{T}(F^{\pi^\star(X)}(\cdot,X)) - \mathsf{T}(F^{\pi^\sharp(X)}(\cdot,X)) < c_1P^{-\gamma}\big) \le C_0c_1^{1+\alpha}nP^{-\gamma(1+\alpha)}, \tag{164}$$

where we used Assumption 4.6 to obtain the last inequality. To deal with the last term in (163), we need a better lower bound on the $\underline{\Delta}_j$-s than the already available $P^{-\gamma}$. For notational simplicity, let's suppose that the well-behaved bins are indexed as $\mathcal{J} = \{1,2,\ldots,j_1\}$ such that $0 < \underline{\Delta}_1 \le \underline{\Delta}_2 \le \ldots \le \underline{\Delta}_{j_1}$

(cf. Equation (161) and the ensuing discussion for $0 < \underline{\Delta}_1$). Fix $j \in \mathcal{J}$. Then, for any $k = 1, \dots, j$, we claim that:

$$B_k \subseteq \left\{ x : 0 < \mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^{\pi^\sharp(x)}(\cdot, x)) < \underline{\Delta}_j + 2CLd^{\gamma/2}P^{-\gamma} \right\}. \tag{165}$$

To see (165), note that, by definition, there exists an $i \in \mathcal{I}_k^0$ such that $\underline{\Delta}_k = \mathsf{T}(F_k^*) - \mathsf{T}(F_k^i)$. For $x \in B_k$ Lemma D.4 yields (the first inequality following from the observation after Equation (160))

$$\begin{aligned} 0 < \mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^{\pi^\sharp(x)}(\cdot, x)) &\leq \mathsf{T}(F^{\pi^\star(x)}(\cdot, x)) - \mathsf{T}(F^i(\cdot, x)) \\ &\leq \underline{\Delta}_k + 2CLd^{\gamma/2}P^{-\gamma} \\ &\leq \underline{\Delta}_j + 2CLd^{\gamma/2}P^{-\gamma}, \end{aligned}$$

and thus x is an element of the set on the right-hand-side of (165). Since all bins B_k are disjoint and $\underline{\Delta}_j + 2CLd^{\gamma/2}P^{-\gamma} \leq c_1\underline{\Delta}_j$ (obtained by recalling $c_1 = 4CLd^{\gamma/2} + 1$, and using the observation $\underline{\Delta}_j > P^{-\gamma}$ made directly after Equation (161)), the inclusion (165) yields that for any $j \in \mathcal{J}$:

$$\mathbb{P}_X\big(0 < \mathsf{T}(F^{\pi^\star(X)}(\cdot, X)) - \mathsf{T}(F^{\pi^\sharp(X)}(\cdot, X)) < c_1\underline{\Delta}_j\big) \geq \sum_{k=1}^{j} \mathbb{P}_X(B_k) \geq \frac{\underline{c}j}{P^d}. \tag{166}$$

Let's denote $j_2 := \max\{j \in \mathcal{J} : \underline{\Delta}_j \leq 1/c_1\}$ (here interpreting the maximum of an empty set as 0). Then, for each $j \in \{1, \dots, j_2\}$ by Assumption 4.6 :

$$\mathbb{P}_X\big(0 < \mathsf{T}(F^{\pi^\star(X)}(\cdot, X)) - \mathsf{T}(F^{\pi^\sharp(X)}(\cdot, X)) < c_1\underline{\Delta}_j\big) \leq C_0(c_1\underline{\Delta}_j)^\alpha. \tag{167}$$

Combining (166), (167), and $\underline{\Delta}_j > P^{-\gamma}$, for any $j \in \{1, \dots, j_2\}$ we get $\underline{\Delta}_j \geq \max\left(c_*\big(jP^{-d}\big)^{1/\alpha}, P^{-\gamma}\right)$, with constant $c_* := c_1^{-1}\underline{c}^{1/\alpha}C^{-1/\alpha}$. Combining this with the identity $\underline{\Delta}_j > 1/c_1$ for $j > j_2$, we obtain that

$$\sum_{j \in \mathcal{J}} \frac{1}{\underline{\Delta}_j} \leq \sum_{j=1}^{j_2} \min\left(c_*^{-1}\big(P^d/j\big)^{1/\alpha}, P^\gamma\right) + \sum_{j=j_2+1}^{j_1} c_1 \leq \sum_{j=1}^{P^d} \min\left(c_*^{-1}\big(P^d/j\big)^{1/\alpha}, P^\gamma\right) + c_1P^d.$$

For $\tilde{P} := \lceil P^{d-\alpha\gamma} \rceil$ (in fact for *any* $\tilde{P} \in \{1, \dots, P^d\}$, and thus in particular for our particular choice) it holds that

$$\sum_{j=1}^{P^d} \min\left(c_*^{-1}\big(P^d/j\big)^{1/\alpha}, P^\gamma\right) \leq \sum_{j=1}^{\tilde{P}} P^\gamma + c_*^{-1}P^{d/\alpha} \sum_{j=\tilde{P}+1}^{\infty} j^{-1/\alpha} \leq c_{**}P^{d+\gamma(1-\alpha)},$$

for $c_{**} := [2 + c_*^{-1}(\alpha^{-1} - 1)^{-1}]$, where we used $\sum_{j=\tilde{P}+1}^{\infty} j^{-1/\alpha} \leq (\alpha^{-1} - 1)^{-1}\tilde{P}^{1-\alpha^{-1}}$ (cf. (110)). Hence, combining Equations (163), (164) with the bounds in the previous two displays we obtain

$$\sum_{j \in \mathcal{J}} \mathbb{E}[\tilde{R}_j(\bar{\pi})] \leq c''' \left(nP^{-\gamma(1+\alpha)} + K\overline{\log}(nP^{-d})P^d + K\overline{\log}(nP^{-d})P^{d+\gamma(1-\alpha)}\right), \tag{168}$$

for a constant c''', say, that depends on $d, L, \gamma, \underline{c}, C, C_0, \alpha$ and β, but *not* on n.

Step 5: Combining. From Equations (138), (144), (147) and (168) we obtain

$$\mathbb{E}[R_N(\bar{\pi})] \leq \frac{c''''}{4}\left(nP^{-\gamma(1+\alpha)} + \sqrt{Kn\overline{\log}(n)}P^{d/2-\gamma\alpha} + K\overline{\log}(nP^{-d})P^d + K\overline{\log}(nP^{-d})P^{d+\gamma(1-\alpha)}\right) \quad (169)$$

for a constant c'''' that depends on $d, L, \gamma, \underline{c}, C, C_0, \alpha$ and β, but *not* on n. From $P = \lceil n^{1/(2\gamma+d)} \rceil$ we get $n \leq P^{2\gamma+d}$, and obtain

$$\mathbb{E}[R_N(\bar{\pi})] \leq \frac{c''''}{4}K\overline{\log}(n)\left(nP^{-\gamma(1+\alpha)} + n^{1/2}P^{d/2-\gamma\alpha} + 2P^{d+\gamma(1-\alpha)}\right) \leq c''''K\overline{\log}(n)P^{d+\gamma(1-\alpha)}, \quad (170)$$

from which the conclusion follows.

□

The following lemma allows to upper bound the number of suboptimal assignments made by the FSA policy. It will also play a crucial role in providing a lower bound on regret in Section D.3 since it establishes that if the margin condition (Assumption 4.6) is in place, such a lower bound can be obtained by lower bounding the number of suboptimal assignments made.

Lemma D.5. *Let a functional T be given and assume that Assumption 4.6 is satisfied. Furthermore, N is independent of all covariates and has expectation n. Then, there exists a $\tilde{C} > 0$ such that for any policy π*

$$\mathbb{E}[R_N(\pi)] \geq \tilde{C}n^{-1/\alpha}\big(\mathbb{E}[S_N(\pi)]\big)^{1+1/\alpha}. \quad (171)$$

Proof. Choose $D_0 \geq 2$ such that $1/(C_0D_0)^{1/\alpha} < 1$. We show that

$$\mathbb{E}[R_N(\pi)] \geq \tilde{C}n^{-1/\alpha}\big(\mathbb{E}[S_N(\pi)]\big)^{1+1/\alpha}. \quad (172)$$

for $\tilde{C} = \tilde{C}(\alpha) = (1 - 1/D_0)/(C_0D_0)^{1/\alpha}$. If $\mathbb{E}[S_N(\pi)] = 0$, (172) is trivially valid. Thus, suppose that $\mathbb{E}[S_N(\pi)] > 0$. Note that for any $\delta > 0$,

$$\begin{aligned}
R_N(\pi) &\geq \delta \sum_{t=1}^{N} \mathbb{1}_{\{\mathsf{T}(F^{\pi^\star(X_t)}(\cdot,X_t)) - \mathsf{T}(F^{\sharp(X_t)}(\cdot,X_t)) > \delta\}} \mathbb{1}_{\left\{\pi_t(X_t,Z_{t-1}) \notin \arg\max\{\mathsf{T}(F^i(\cdot,X_t)),\ i=1,\ldots,K\}\right\}} \\
&= \delta S_N(\pi) - \delta \sum_{t=1}^{N} \mathbb{1}_{\{\mathsf{T}(F^{\pi^\star(X_t)}(\cdot,X_t)) - \mathsf{T}(F^{\sharp(X_t)}(\cdot,X_t)) \leq \delta\}} \mathbb{1}_{\left\{\pi_t(X_t,Z_{t-1}) \notin \arg\max\{\mathsf{T}(F^i(\cdot,X_t)),\ i=1,\ldots,K\}\right\}} \\
&= \delta S_N(\pi) - \delta \sum_{t=1}^{N} \mathbb{1}_{\{0 < \mathsf{T}(F^{\pi^\star(X_t)}(\cdot,X_t)) - \mathsf{T}(F^{\sharp(X_t)}(\cdot,X_t)) \leq \delta\}} \mathbb{1}_{\left\{\pi_t(X_t,Z_{t-1}) \notin \arg\max\{\mathsf{T}(F^i(\cdot,X_t)),\ i=1,\ldots,K\}\right\}} \\
&\geq \delta S_N(\pi) - \delta \sum_{t=1}^{N} \mathbb{1}_{\{0 < \mathsf{T}(F^{\pi^\star(X_t)}(\cdot,X_t)) - \mathsf{T}(F^{\sharp(X_t)}(\cdot,X_t)) \leq \delta\}},
\end{aligned}$$

where the second equality used that if $\pi_t(X_t, Z_{t-1}) \notin \arg\max\{\mathsf{T}(F^i(\cdot, X_t)),\ i = 1, ..., K\}$, then $0 < \mathsf{T}(F^{\pi^\star(X_t)}(\cdot, X_t)) - \mathsf{T}(F^{\sharp(X_t)}(\cdot, X_t))$. Choosing $\delta := (\mathbb{E}[S_N(\pi)]/(nC_0D_0))^{1/\alpha} \leq 1/(C_0D_0)^{1/\alpha} < 1$ (the inequality following from $\mathbb{E}[S_N(\pi)] \leq \mathbb{E}(N) \leq n$), Assumption 4.6 yields

$$\mathbb{E}[R_N(\pi)] \geq \delta(\mathbb{E}[S_N(\pi)] - C_0 n\delta^\alpha) = \delta(1 - 1/D_0)\mathbb{E}[S_N(\pi)] = \tilde{C}n^{-1/\alpha}\big(\mathbb{E}[S_N(\pi)]\big)^{1+1/\alpha}, \tag{173}$$

which proves (172). □

Proof of Theorem 4.8. Combine Theorem 4.7 and Lemma D.5. □

D.3 Proofs for Section 5

We begin with an auxiliary lemma bounding the Kolmogorov distance between any two members of $\mathcal{H}$ as defined in Definition 3.2.

Lemma D.6. *For all $a_1 < a_2$ in $(-1, \infty)$ it holds that*

$$\|H_{a_1} - H_{a_2}\|_\infty = \left[\frac{a_1+1}{a_2+1}\right]^{(a_1+1)/(a_2-a_1)} \frac{a_2 - a_1}{a_2+1} \leq \frac{a_2 - a_1}{a_2+1}. \tag{174}$$

Proof. Let $a_1 < a_2$ be elements of $(-1, \infty)$. By definition of the $\|.\|_\infty$-norm and the cdf H_a it holds that

$$\|H_{a_1} - H_{a_2}\|_\infty = \sup_{x\in[0,1]} |x^{a_1+1} - x^{a_2+1}|. \tag{175}$$

For every $x \in [0, 1]$ the function $a \mapsto x^{a+1}$ is strictly decreasing on $(-1, \infty)$. Hence, using $a_1 < a_2$, it follows that the supremum to the right in the previous display equals

$$\sup_{x\in[0,1]} \left(x^{a_1+1} - x^{a_2+1}\right) = \max_{x\in(0,1)} \left(x^{a_1+1} - x^{a_2+1}\right), \tag{176}$$

the equality being trivial. It is elementary to verify (e.g., by checking the first and second order conditions for a maximum) that the maximum in the previous display is attained at

$$x_* := \left[\frac{a_1+1}{a_2+1}\right]^{1/(a_2-a_1)}, \tag{177}$$

from which it follows that

$$\|H_{a_1} - H_{a_2}\|_\infty = x_*^{a_1+1} - x_*^{a_2+1} = x_*^{a_1+1}\left(1 - x_*^{a_2-a_1}\right) = x_*^{a_1+1}\frac{a_2 - a_1}{a_2+1}, \tag{178}$$

which proves the claimed equality, the inequality being a trivial consequence of $x_* \in (0, 1)$. □

Lemma D.7. *Suppose the functional T satisfies Assumptions 2.1 and 3.3. With $\bar{a}$ and δ as in Assumption 3.3, denote the image of T over $\mathbf{A}(\delta) := \{H_a : a \in [\bar{a} - \delta, \bar{a} + \delta]\}$ by $\mathsf{T}(\mathbf{A}(\delta))$. Then, $\mathsf{T}(\mathbf{A}(\delta))$ is a non-empty compact interval and*

1. *the function from* $[\bar{a}-\delta, \bar{a}+\delta]$ *to* $\mathsf{T}(\mathbf{A}(\delta))$ *defined via* $a \mapsto \mathsf{T}(H_a)$ *is Lipschitz continuous and possesses an inverse that is Lipschitz continuous.*

2. *for any non-empty compact interval* $I \subseteq \mathsf{T}(\mathbf{A}(\delta))$, *there exists a Lipschitz continuous function* $A : \mathsf{T}(\mathbf{A}(\delta)) \to [\bar{a}-\delta, \bar{a}+\delta]$ *such that for any function* $f : [0,1]^d \to I$ *it holds that*

$$\mathsf{T}(H_{A(f(x))}) = f(x) \quad \textit{for every } x \in [0,1]^d. \tag{179}$$

Proof. Consider part 1. first. It suffices to verify the statement under the condition in Equation (6) (the other statement can be obtained from this one upon passing from T to $-\mathsf{T}$). Under Equation (6), one has $\mathsf{T}(\mathbf{A}(\delta)) = [\mathsf{T}(H_{\bar{a}-\delta}), \mathsf{T}(H_{\bar{a}+\delta})]$ which is a non-empty compact interval. Furthermore, it suffices to verify that there exists a constant $L > 0$, say, such that for every pair $a_1 \neq a_2$ in $[\bar{a}-\delta, \bar{a}+\delta]$ it holds that

$$L^{-1}|a_1 - a_2| \leq |\mathsf{T}(H_{a_1}) - \mathsf{T}(H_{a_2})| \leq L|a_1 - a_2|. \tag{180}$$

For the upper bound in the previous display, we use Assumption 2.1 together with Lemma D.6 to obtain $\|\mathsf{T}(H_{a_1}) - \mathsf{T}(H_{a_2})\|_\infty \leq C\|H_{a_1} - H_{a_2}\|_\infty \leq C(\min(a_1, a_2) + 1)^{-1}|a_2 - a_1|$. Since $\min(a_1, a_2) \geq \bar{a} - \delta > -1$ the second inequality in the previous display follows with $L = (\bar{a} - \delta + 1)^{-1}C$. Next, observe that the assumption in Equation (6) implies the first inequality in the previous display with constant c (instead of L^{-1}). Increasing L, if necessary, proves the first part of the lemma. Denoting the inverse of $a \mapsto \mathsf{T}(H_a)$ by A one has in particular that

$$\mathsf{T}(H_{A(z)}) = z \quad \text{for every } z \in \mathsf{T}(\mathbf{A}(\delta))$$

which yields the second part of the lemma upon using that $f(x) \in I \subseteq \mathsf{T}(\mathbf{A}(\delta))$ for any $x \in [0,1]^d$. □

In proving Theorem 5.1 it will be useful to make the dependence of regret

$$R_n(\pi) = R_n(\pi, F^1, F^2) = \sum_{t=1}^{n} \left|\mathsf{T}\big(F^1(\cdot, X_t)\big) - \mathsf{T}\big(F^2(\cdot, X_t)\big)\right| \mathbb{1}_{\{\pi^\star(X_t) \neq \pi_t(X_t, Z_{t-1})\}}. \tag{181}$$

and and the number of suboptimal assignments

$$S_n(\pi) = S_n(\pi, F^1, F^2) = \sum_{t=1}^{n} \mathbb{1}_{\{\mathsf{T}(F^1(\cdot, X_t)) \neq \mathsf{T}(F^2(\cdot, X_t)),\ \pi^\star(X_t) \neq \pi_t(X_t, Z_{t-1})\}}.$$

on the conditional distributions F_1 and F_1 explicit for any policy π. We make the following remark on notation prior to proving Theorem 5.1

Proof of Theorem 5.1. Throughout the proof, fix a functional T satisfying Assumptions 2.1 and 3.3 as well as a policy π. It will be notationally convenient to label treatment 2 as -1. The idea of proving a lower bound on regret follows the general pattern of reducing the problem to a testing problem as in Chapter 2 of Tsybakov (2009).

The proof consists of four steps. First, we construct a certain set of Hölder continuous functions $\mathcal{C}$. Second, based on $\mathcal{C}$, we construct a set $\underline{\mathcal{S}}$ of 2-tuples of (conditional) distributions on the Borel sets of $[0,1]$. Third, these distributions are shown to satisfy Assumptions 4.2 and 4.6, i.e $\underline{\mathcal{S}} \subseteq \mathcal{S}$ such that the treatment problem falls under the assumptions of Theorem 4.7 Fourth, we show that for any policy π

$$\sup_{(F^1,F^{-1})\in\underline{\mathcal{S}}} \mathbb{E}[R_n(\pi, F^1, F^{-1})] \geq c_2 n^{1-\frac{\gamma(1+\alpha)}{2\gamma+d}}.$$

for some $c_2 > 0$ independent of π. To establish the above display, we use Lemma D.5 to conclude that it suffices to provide a lower bound on the number of suboptimal treatments made by π. This, in turn, is achieved by using standard techniques for obtaining minimax lower bounds (e.g., Chapter 2 of Tsybakov (2009)). In particular, by lower bounding the sum of Type 1 and Type 2 errors in a certain binary testing problem.

Step 1: Construction of the Hölder class $\mathcal{C}$.

For $P \geq 2$, let $B_1, B_2, \ldots, B_{P^d}$ be the hypercubes defined in (11) sorted in the lexicographic order. Let $q_i,\ i = 1,...,P^d$ be the center of B_i. Furthermore, set $m = \lceil 0.5P^{d-\alpha\gamma} \rceil$ and note that $m \leq P^d$ for $P \geq 2$. Next, set $\Omega_m = \{-1,1\}^m$ and observe that $|\Omega_m| = 2^m$. For each $\omega \in \Omega_m$, we define the function f_ω on $[0,1]^d$ as

$$f_\omega(x) = \frac{1}{2} + \sum_{j=1}^{m} \omega_j \varphi_j(x),$$

where $\varphi_j(x) = 4^{-1}P^{-\gamma}\phi(2P(x-q_j))1_{B_j}(x)$ and $\phi(x) = (1-||x||_\infty)^\gamma$, and we let $\|x\|_\infty = \max_{1\leq i\leq d}|x_i|$ for $x \in \mathbb{R}^d$. With this notation in place, we can define the class of functions

$$\mathcal{C}_m := \mathcal{C} = \{f_\omega : \omega \in \Omega_m\}.$$

We now show that each element of $\mathcal{C}$ is Hölder continuous. More precisely, for each $f_\omega \in \mathcal{C}$, we show that for any pair $x_1, x_2 \in [0,1]^d$

$$|f_\omega(x_1) - f_\omega(x_2)| \leq 2^{-1}||x_1 - x_2||^\gamma,$$

where $\|.\|$ denotes the Euclidean norm. Observe that for any pair $x_1, x_2 \in [0,1]^d$ one has $|\phi(x_1)-\phi(x_2)| \leq ||x_1-x_2||_\infty^\gamma \leq ||x_1-x_2||^\gamma$ where the first inequality uses i) for $a, b \geq 0$ one has $a^\gamma \leq b^\gamma + |b-a|^\gamma$ for $\gamma \leq 1$ and ii) the reverse triangle inequality for $\|\cdot\|_\infty$.

If $x_1, x_2 \in B_j$ for some $j \in \{1,...,P^d\}$, the definition of f_ω and $|\phi(x_1)-\phi(x_2)| \leq ||x_1-x_2||^\gamma$ lead to

$$|f_\omega(x_1) - f_\omega(x_2)| \leq |\varphi_j(x_1) - \varphi_j(x_2)| \leq 4^{-1}P^{-\gamma}(2P)^\gamma ||x_1 - q_j - (x_2 - q_j)||^\gamma \leq 2^{-1}||x_1 - x_2||^\gamma. \quad (182)$$

Suppose instead that $x_1 \in B_j, x_2 \in B_k$ for some $j \neq k$. Let $S = \{\theta x_1 + (1-\theta)x_2 : \theta \in [0,1]\}$ be the line connecting x_1 and x_2. Define $y_1 = \operatorname{argmin}_{z\in S\cap\bar{B}_j}||z - x_2||$ and $y_2 = \operatorname{argmin}_{z\in S\cap\bar{B}_k}||z - x_1||$. Noting that $\varphi_j(y_1) = \varphi_k(y_2) = 0$ we obtain from (182) that

$$\begin{aligned}
|f_\omega(x_1) - f_\omega(x_2)| &= |\omega_j\varphi_j(x_1) - \omega_k\varphi_k(x_2)| \\
&\leq |\omega_j\varphi_j(x_1) - \omega_j\varphi_j(y_1)| + |\omega_j\varphi_j(y_1) - \omega_k\varphi_k(y_2)| + |\omega_k\varphi_k(y_2) - \omega_k\varphi_k(x_2)| \\
&\leq 4^{-1}2^\gamma||x_1 - y_1||^\gamma + 4^{-1}2^\gamma||y_2 - x_2||^\gamma \\
&\leq 2^{-1}||x_1 - x_2||^\gamma
\end{aligned}$$

where we exploited $||x_1 - x_2|| = ||x_1 - y_1|| + ||y_1 - y_2|| + ||y_2 - x_2|| \geq ||x_1 - y_1|| + ||y_2 - x_2||$ combined with the inequality $(a^\gamma + b^\gamma) \leq 2^{1-\gamma}(a+b)^\gamma$ for $\gamma \leq 1$.

Step 2: Construction of $\underline{\mathcal{S}}$.

Observe that the range of each $f \in \mathcal{C}$ is contained in $[1/4, 3/4]$. If necessary, a linear transformation of T ensures that $[1/4, 3/4] \subseteq \mathsf{T}(\mathbf{A}(\delta))$ (with $\mathsf{T}(\mathbf{A}(\delta))$ as in Lemma D.7) without affecting the rates of the lower bound on maximal regret. Part 2. of Lemma D.7 now implies for each $f \in \mathcal{C}$ there exists a $g_f : [0,1]^d \to [\bar{a} - \delta, \bar{a} + \delta]$ with $g_f(x) = A(f(x))$ for all $x \in [0,1]^d$ and A is Lipschitz continuous such that $\mathsf{T}(H_{g_f(x)}) = f(x)$. Let $\mathbb{P}_{f,1}(\cdot, x)$ be the distribution on the Borel sets of $[0,1]$ with cdf $H_{g_f(x)}(y) = y^{g_f(x)+1}$. Borel measurability of $x \mapsto \mathbb{P}_{f,1}(A, x)$ for each $A \in \mathcal{B}([0,1])$ follows from continuity of g_f and Scheffé's Lemma. Let the joint distribution $\mathbb{P}_f$ of $(Y_{1,t}, X_t)$ on $\mathcal{B}([0,1]^{1+d})$ be defined as $\mathbb{P}_f(A \times B) = \int_B \mathbb{P}_{f,1}(A, x)\mathbb{P}_X(dx)$ for $A \in \mathcal{B}([0,1])$ and $B \in \mathcal{B}([0,1]^d)$.

To finish the construction of the distribution of $(Y_{-1,t}, Y_{1,t}, X_t)$ on $\mathcal{B}([0,1]^{2+d})$, denoted $\mathbb{P}_{f,1/2}$, let $\mathbb{P}_{1/2,-1}(\cdot)$ be the distribution on the Borel sets of $[0,1]$ with cdf $H_{g_{1/2}}(y) = y^{g_{1/2}+1}$ where $g_{1/2} = A(1/2)$ is a constant (function constant in x). Finally, let $\mathbb{P}_{f,1/2}(A_1 \times A_2 \times B) = \int_{A_2 \times B} \mathbb{P}_{1/2,-1}(A_1) d\mathbb{P}_f(y,x) = \mathbb{P}_{1/2,-1}(A_1) \cdot \mathbb{P}_f(A_2 \times B)$ for $A_1, A_2 \in \mathcal{B}([0,1])$ and $B \in \mathcal{B}([0,1]^d)$. Thus, $Y_{-1,t}$ is independent of $(Y_{1,t}, X_t)$. This independence is merely chosen for concreteness as the important ingredients in Steps 3 and 4 below are the conditional distributions of $\mathbb{P}_{f,1}$ (of $Y_{1,t}$ given X_t) and $\mathbb{P}_{1/2,-1}$ (of $Y_{-1,t}$ given X_t), respectively [4].

With these definitions in place, for each $f \in \mathcal{C}$, let $\mathbb{P}^t_{\pi,f}$ be the distribution of Z_t on the Borel sets of $\mathbb{R}^{(d+1)t}$ with corresponding expectation $\mathbb{E}^t_{\pi,f}$. Define $\mathbb{P}_{\pi,f,1,t}(A, (x,z)) := \mathbb{P}_{f,1}(A, x)$ for every $A \in \mathcal{B}([0,1])$ and $(x,z) \in \mathbb{R}^d \times \mathbb{R}^{(d+1)(t-1)}$. By independence of $(Y_{1,t}, X_t)$ and Z_{t-1} the Markov kernel $\mathbb{P}_{\pi,f,1,t}$ defines a regular conditional probability of $Y_{1,t}$ given (X_t, Z_{t-1}). Similarly, $\mathbb{P}_{\pi,f,-1,t}(A, (x,z)) := \mathbb{P}_{1/2,-1}(A)$ defines a regular conditional distribution of $Y_{-1,t}$ given (X_t, Z_{t-1}).

To exhibit $\mathbb{P}^t_{\pi,f}$ explicitly, let $u_1 = x_1 \in \mathbb{R}^d$ and $z_1 = (y_1, x_1) \in \mathbb{R} \times \mathbb{R}^d$. For $s = 2, ..., t$ let $z_s = (y_s, x_s, z_{s-1}) \in \mathbb{R}^{(d+1)s}$. Then, for any $A \in \mathcal{B}([0,1])$, $u_s = (x_s, z_{s-1}) \in \mathbb{R}^d \times \mathbb{R}^{(s-1)(d+1)}$ and $f \in \mathcal{C}$ define

$$\begin{aligned}
\mathbb{P}_{\pi,f,s}(A, u_s) &:= \mathbb{P}_{\pi,f,1,s}(A, u_s) 1_{\{\pi_s(u_s)=1\}} + \mathbb{P}_{\pi,f,-1,s}(A, u_s) 1_{\{\pi_s(u_s)=-1\}} \\
&= \mathbb{P}_{f,1}(A, x_s) 1_{\{\pi_s(u_s)=1\}} + \mathbb{P}_{1/2,-1}(A) 1_{\{\pi_s(u_s)=-1\}} \qquad (183) \\
&= \int_A \left[(1 + g_f(x_s)) y_s^{g_f(x_s)} 1_{\{\pi_s(u_s)=1\}} + (1 + g_{1/2}) y_s^{g_{1/2}} 1_{\{\pi_s(u_s)=-1\}} \right] dy_s. \qquad (184)
\end{aligned}$$

Note that $\mathbb{P}_{\pi,f,s}$ defines a regular conditional distribution of $Y_{\pi_t(X_t, Z_{t-1}),t}$ given (X_t, Z_{t-1}). Thus, setting $d_s(z_s) = (1 + g_f(x_s)) y_s^{g_f(x_s)} 1_{\{\pi_s(u_s)=1\}} + (1 + g_{1/2}) y_s^{g_{1/2}} 1_{\{\pi_s(u_s)=-1\}}$, we observe by (184) for $A_s \in \mathcal{B}([0,1]), B_s \in \mathcal{B}([0,1]^d),\ s = 1, ..., t$

$$\begin{aligned}
\mathbb{P}^t_{\pi,f}(\times_{s=1}^t (A_s \times B_s)) &= \int_{\times_{s=1}^{t-1}(A_s \times B_s)} \int_{B_t} \mathbb{P}_{\pi,f,t}(A_t, (x_t, z_{t-1})) \mathbb{P}_X(dx_t) \mathbb{P}^{t-1}_{\pi,f}(dz_{t-1}) \\
&= \int_{\times_{s=1}^{t-1}(A_s \times B_s)} \int_{B_t} \int_{A_t} d_t(z_t) dy_t dx_t \mathbb{P}^{t-1}_{\pi,f}(dz_{t-1})
\end{aligned}$$

[4]Recall that the marginal distribution of $Y_{-1,t}$, $\mathbb{P}_{1/2,-1}$, is also the conditional distribution of $Y_{-1,t}$ given X_t by the independence of $Y_{-1,t}$ and X_t.

which can be used as the induction step to show (the induction start is trivial) that $\mathbb{P}^t_{\pi,f}$ is absolutely continuous with density $d(z_t) = \prod_{s=1}^t d_s(z_s)$ with respect to the $t(d+1)$-dimensional Lebesgue measure (restricted to $[0,1]^{t(d+1)}$).

Now, set $\underline{\mathcal{S}} = \{(H_{g_f(x)}, H_{g_{1/2}}) : \ f \in \mathcal{C}\}$.

Step 3: Verifying that $\underline{\mathcal{S}} \subseteq \mathcal{S}$.

To verify that $\underline{\mathcal{S}} \subseteq \mathcal{S}$ we show that for every $f \in \mathcal{C}$: i) one has $||H_{g_f(x_1)} - H_{g_f(x_2)}||_\infty \leq L||x_1 - x_2||^\gamma$ for some $L > 0$ (Assumption 4.2), and ii) the margin condition (Assumption 4.6) is satisfied.

Verifying Assumption 4.2:

We begin by verifying that for each $f \in \mathcal{C}$ one has $||H_{g_f(x_1)} - H_{g_f(x_2)}||_\infty \leq L||x_1 - x_2||^\gamma$ for some $L > 0$. Note that by Lemma D.6

$$||H_{g_f(x_1)}(y) - H_{g_f(x_2)}(y)||_\infty \leq \frac{|g_f(x_1) - g_f(x_2)|}{\bar{a} - \delta + 1}, \quad \bar{a} - \delta > -1$$

such that the conclusion follows upon recalling that $g_f(x) = A(f(x))$ with Lipschitz continuous A and Hölder continuous f. Denoting by c_1 the Lipschitz constant of A, we can choose $L = \frac{c_1}{2(\bar{a}-\delta+1)}$. Since $H_{g_{1/2}}(y)$ does not depend on x it is Hölder continuous as well.

Verifying Assumption 4.6: We next verify that each tuple in $\underline{\mathcal{S}}$ satisfies the margin condition. To be precise, we shall show that for every $f \in \mathcal{C}$

$$\mathbb{P}_X\big(0 < |\mathsf{T}(H_{g_f(X)}) - \mathsf{T}(H_{g_{1/2}(X)})| \leq \delta\big) \leq 8d\delta^\alpha \text{ for all } \delta \in [0,1], \tag{185}$$

which will verify the margin condition with $C_0 = 8d$ since there are only two treatments. To this end, we note that $\mathsf{T}(H_{g_f(X)}) - \mathsf{T}(H_{g_{1/2}}) = f_\omega(X) - 1/2$ for some $\omega \in \Omega_m$. Since $\mathbb{P}_X$ is the uniform distribution on $[0,1]^d$ and recalling $\phi(x) = (1 - ||x||_\infty)^\gamma$, for any $\omega \in \Omega_m$, the substitution $u = 2Px - 2Pq_1$ yields

$$\begin{aligned}
\mathbb{P}_X(0 < |f_\omega(X) - 1/2| \leq \delta) &= \sum_{j=1}^m \mathbb{P}_X(0 < |f_\omega(X) - 1/2| \leq \delta, X \in B_j) \\
&= m\mathbb{P}_X(0 < \phi(2P(X - q_1)) \leq 4P^\gamma\delta, X \in B_1) \\
&= m(2P)^{-d} \int_{2PB_1 - 2Pq_1} 1_{\{\phi(x) \leq 4P^\gamma\delta\}} dx \\
&= m(2P)^{-d} \int_{[-1,1]^d} 1_{\{\phi(x) \leq 4P^\gamma\delta\}} dx \\
&= mP^{-d} \int_{[0,1]^d} 1_{\{\phi(x) \leq 4P^\gamma\delta\}} dx,
\end{aligned}$$

where the last equality follows from $\phi(x)$ being invariant to changing the signs of the coordinates of x. To bound the last line of the above display consider two cases. If $4P^\gamma\delta > 1$, then

$$\mathbb{P}_X(0 < |f_\omega(X) - 1/2| \leq \delta) = mP^{-d} \int_{[0,1]^d} 1_{\{\phi(x) \leq 4P^\gamma\delta\}} dx = mP^{-d} \leq 2P^{-\gamma\alpha} \leq 8\delta^\alpha,$$

where we used $m = \lceil 0.5P^{d-\gamma\alpha} \rceil \leq 0.5P^{d-\gamma\alpha} + 1 \leq 2P^{d-\gamma\alpha}$ and $\alpha \in (0,1)$.

On the other hand, if $4P^\gamma\delta \leq 1$, we obtain that

$$\begin{aligned}
\mathbb{P}_X(0 < |f_\omega(X) - 1/2| \leq \delta) &= mP^{-d} \int_{[0,1]^d} 1_{\{\phi(x) \leq 4P^\gamma\delta\}} dx \\
&= mP^{-d} - mP^{-d} \int_{[0,1]^d} 1_{\{||x||_\infty < 1 - 4^{1/\gamma}\delta^{1/\gamma}P\}} dx \\
&= mP^{-d}[1 - (1 - 4^{1/\gamma}\delta^{1/\gamma}P)^d] \\
&\leq mP^{-d} d 4^{1/\gamma}\delta^{1/\gamma}P \\
&\leq 2dP^{1-\alpha\gamma}4^{1/\gamma}\delta^{1/\gamma} \\
&\leq 2d(4\delta)^\alpha \leq 8d\delta^\alpha,
\end{aligned}$$

which establishes (185).

Step 4: Lower bounding $\sup_{(F^1,F^{-1})\in\underline{\mathcal{S}}} \mathbb{E}(R_n(\pi, F^1, F^{-1}))$**.**

By Lemma D.5 it suffices to show that

$$\sup_{(F^1,F^{-1})\in\underline{\mathcal{S}}} \mathbb{E}(S_n(\pi, F^1, F^{-1})) \geq c_3 n^{1-\frac{\alpha\gamma}{d+2\gamma}} \tag{186}$$

for some $c_3 > 0$ independent of π. Note also that the left hand side of (186) is equal to $\sup_{f\in\mathcal{C}} \mathbb{E}^n_{\pi,f}\left[S_n(\pi)\right]$ which we shall now lower bound. Since X_t is independent of Z_{t-1} and $\pi^\star(x) = \operatorname{sign}(f_\omega(x) - 1/2)$,

$$\begin{aligned}
\sup_{f\in\mathcal{C}} \mathbb{E}^n_{\pi,f}[S_n(\pi)] &= \sup_{\omega\in\Omega_m} \sum_{t=1}^n \mathbb{E}^{t-1}_{\pi,f_\omega}\left[\mathbb{P}_X\left(\pi_t(X_t, Z_{t-1}) \neq \operatorname{sign}(f_\omega(X_t) - 1/2),\ f_\omega(X_t) \neq 1/2\right)\right] \\
&\geq \sup_{\omega\in\Omega_m} \sum_{j=1}^m \sum_{t=1}^n \mathbb{E}^{t-1}_{\pi,f_\omega}[\mathbb{P}_X(\pi_t(X_t, Z_{t-1}) \neq \omega_j, X_t \in B_j)] \\
&\geq \frac{1}{2^m} \sum_{j=1}^m \sum_{t=1}^n \sum_{\omega\in\Omega_m} \mathbb{E}^{t-1}_{\pi,f_\omega}[\mathbb{P}_X(\pi_t(X_t, Z_{t-1}) \neq \omega_j, X_t \in B_j)].
\end{aligned} \tag{187}$$

Note that for every $j \in \{1, ..., m\}$ and $t \in \{1, ..., n\}$,

$$Q^j_t := \sum_{\omega\in\Omega_m} \mathbb{E}^{t-1}_{\pi,f_\omega}[\mathbb{P}_X(\pi_t(X_t, Z_{t-1}) \neq \omega_j, X_t \in B_j)] = \sum_{\omega_{-j}\in\Omega_{m-1}} \sum_{i\in\{-1,1\}} \mathbb{E}^{t-1}_{\pi,f_{\omega^i_{-j}}}[\mathbb{P}_X(\pi_t(X_t, Z_{t-1}) \neq i, X_t \in B_j)]$$

where $\omega_{-j} = (\omega_1, ..., \omega_{j-1}, \omega_{j+1}, ..., \omega_m)$ and $\omega^i_{-j} = (\omega_1, ..., \omega_{j-1}, i, \omega_{j+1}, ..., \omega_m)$ for $i \in \{-1, 1\}$. Note that for any $u \in \mathbb{R}^{(t-1)(d+1)}$, $\mathbb{P}_X(\pi_t(X_t, u) \neq i, X_t \in B_j) = \mathbb{P}^j_X(\pi_t(X_t, u) \neq i)/P^d$ with $\mathbb{P}^j_X(A) = \mathbb{P}_X(A|X_t \in B_j)$ for any $A \in \mathcal{B}([0,1]^d)$. Expectations with respect to $\mathbb{P}^j_X$ are denoted by $\mathbb{E}^j_X$. Hence, for every $\omega_{-j} \in \Omega_{m-1}$,

$$\sum_{i\in\{-1,1\}} \mathbb{E}^{t-1}_{\pi,f_{\omega^i_{-j}}}[\mathbb{P}_X(\pi_t(X_t, Z_{t-1}) \neq i, X_t \in B_j)] = \frac{1}{P^d} \sum_{i\in\{-1,1\}} \mathbb{E}^{t-1}_{\pi,f_{\omega^i_{-j}}}[\mathbb{P}^j_X(\pi_t(X_t, Z_{t-1}) \neq i)].$$

Here $\sum_{i\in\{-1,1\}}\mathbb{E}^{t-1}_{\pi,f_{\omega^i_{-j}}}[\mathbb{P}^j_X(\pi_t(X_t,Z_{t-1})\neq i)] = \mathbb{E}^{t-1}_{\pi,f_{\omega^{-1}_{-j}}}\mathbb{E}^j_X 1_{\{\pi_t(X_t,Z_{t-1})=1\}} + 1 - \mathbb{E}^{t-1}_{\pi,f_{\omega^1_{-j}}}\mathbb{E}^j_X 1_{\{\pi_t(X_t,Z_{t-1})=1\}}$ is the sum of Type 1 and Type 2 errors for the testing problem $H_0 : \mathbb{P}^{t-1}_{\pi,f}\otimes\mathbb{P}^j_X = \mathbb{P}^{t-1}_{\pi,f_{\omega^{-1}_{-j}}}\otimes\mathbb{P}^j_X$ vs $H_a : \mathbb{P}^{t-1}_{\pi,f}\otimes\mathbb{P}^j_X = \mathbb{P}^{t-1}_{\pi,f_{\omega^1_{-j}}}\otimes\mathbb{P}^j_X$ for the test $1_{\{\pi_t(X_t,Z_{t-1})=1\}}$. For any test π_t this sum can be bounded from below, using Theorem 2.2(iii) of Tsybakov (2009), by

$$\begin{aligned}\sum_{i\in\{-1,1\}}\mathbb{E}^{t-1}_{\pi,f_{\omega^i_{-j}}}[\mathbb{P}^j_X(\pi_t(X_t,Z_{t-1})\neq i)] &\geq \frac{1}{4}\exp\left[-KL\left(\mathbb{P}^{t-1}_{\pi,f_{\omega^{-1}_{-j}}}\otimes\mathbb{P}^j_X,\mathbb{P}^{t-1}_{\pi,f_{\omega^1_{-j}}}\otimes\mathbb{P}^j_X\right)\right]\\ &= \frac{1}{4}\exp\left[-KL\left(\mathbb{P}^{t-1}_{\pi,f_{\omega^{-1}_{-j}}},\mathbb{P}^{t-1}_{\pi,f_{\omega^1_{-j}}}\right)\right]\end{aligned}$$

Thus, for every $\omega_{-j}\in\Omega_{m-1}$,

$$\sum_{i\in\{-1,1\}}\mathbb{E}^{t-1}_{\pi,f_{\omega^i_{-j}}}[\mathbb{P}_X(\pi_t(X_t,Z_{t-1})\neq i, X_t\in B_j)] \geq \frac{1}{4P^d}\exp\left[-KL\left(\mathbb{P}^{t-1}_{\pi,f_{\omega^{-1}_{-j}}},\mathbb{P}^{t-1}_{\pi,f_{\omega^1_{-j}}}\right)\right]$$

and we next bound $KL\left(\mathbb{P}^{t-1}_{\pi,f_{\omega^{-1}_{-j}}},\mathbb{P}^{t-1}_{\pi,f_{\omega^1_{-j}}}\right)$ from above. Using the chain rule for Kullback-Leibler divergence, cf. Theorem 2.5.3 of Cover and Thomas (2012) [5], it follows that

$$\begin{aligned}KL\left(\mathbb{P}^{t-1}_{\pi,f_{\omega^{-1}_{-j}}},\mathbb{P}^{t-1}_{\pi,f_{\omega^1_{-j}}}\right) &= KL\left(\mathbb{P}^{t-2}_{\pi,f_{\omega^{-1}_{-j}}}\otimes\mathbb{P}_X,\mathbb{P}^{t-2}_{\pi,f_{\omega^1_{-j}}}\otimes\mathbb{P}_X\right) + \mathbb{E}^{t-2}_{\pi,f_{\omega^{-1}_{-j}}}\mathbb{E}_X KL\left(\mathbb{P}_{\pi,f_{\omega^{-1}_{-j}},t-1},\mathbb{P}_{\pi,f_{\omega^1_{-j}},t-1}\right)\\ &= KL\left(\mathbb{P}^{t-2}_{\pi,f_{\omega^{-1}_{-j}}},\mathbb{P}^{t-2}_{\pi,f_{\omega^1_{-j}}}\right) + \mathbb{E}^{t-2}_{\pi,f_{\omega^{-1}_{-j}}}\mathbb{E}_X KL\left(\mathbb{P}_{\pi,f_{\omega^{-1}_{-j}},t-1},\mathbb{P}_{\pi,f_{\omega^1_{-j}},t-1}\right).\end{aligned}$$

To proceed, note that by (183) for any $s=1,...,t-1$, $u=(x,z)\in\mathbb{R}^d\times\mathbb{R}^{(s-1)(d+1)}$ (where $u=x\in\mathbb{R}^d$ for $s=1$) and $f\in\mathcal{C}$

$$\begin{aligned}KL\left(\mathbb{P}_{\pi,f_{\omega^{-1}_{-j}},s},\mathbb{P}_{\pi,f_{\omega^1_{-j}},s}\right) &= KL\left(\mathbb{P}_{f_{\omega^{-1}_{-j}},1},\mathbb{P}_{f_{\omega^1_{-j}},1}\right)1_{\{\pi_s(u)=1\}} + KL\left(\mathbb{P}_{1/2,-1},\mathbb{P}_{1/2,-1}\right)1_{\{\pi_s(u)=-1\}}\\ &= KL\left(\mathbb{P}_{f_{\omega^{-1}_{-j}},1},\mathbb{P}_{f_{\omega^1_{-j}},1}\right)1_{\{\pi_s(u)=1\}}. \qquad (188)\end{aligned}$$

Next, observe that the function $f_{\omega^{-1}_{-j}} - f_{\omega^1_{-j}}$ is γ-Hölder continuous with constant 2. Furthermore, it vanishes on the boundary of B_j. Using these observations along with (188) and $g_f(x)=A(f(x))$ for any $f\in\mathcal{C}$, one obtains for any $s=1,...,t-1$:

$$\begin{aligned}KL\left(\mathbb{P}_{\pi,f_{\omega^{-1}_{-j}},s},\mathbb{P}_{\pi,f_{\omega^1_{-j}},s}\right) &= KL\left(\mathbb{P}_{f_{\omega^{-1}_{-j}},1},\mathbb{P}_{f_{\omega^1_{-j}},1}\right)1_{\{\pi_s(u)=1\}}\\ &\leq \frac{1}{(1+\bar{a}-\delta)^2}\left(g_{f_{\omega^{-1}_{-j}}}(x)-g_{f_{\omega^1_{-j}}}(x)\right)^2 1_{\{\pi_s(u)=1,x\in B_j\}}\\ &\leq \frac{c_1^2}{(1+\bar{a}-\delta)^2}\left(f_{\omega^{-1}_{-j}}(x)-f_{\omega^1_{-j}}(x)\right)^2 1_{\{\pi_s(z)=1,x\in B_j\}}\\ &\leq \frac{4\tilde{c}^2}{P^{2\gamma}}1_{\{\pi_s(u)=1,x\in B_j\}},\end{aligned}$$

[5]While the proof in Cover and Thomas (2012) is for discrete measures, the same proof technique applies equally well to measures equivalent to a product of Lebesgue measures which is the case in our setting as observed at the end of Step 2.

for a constant $\tilde{c} = c_1 d^{\gamma/2}/(1+\bar{a}-\delta)$ (with c_1 being the Lipschitz constant of A). It thus follows by induction that for any $t = 1, ..., n$, $j = 1, ..., m$ and policy π

$$KL\left(\mathbb{P}^{t-1}_{\pi, f_{\omega^{-1}_{-j}}}, \mathbb{P}^{t-1}_{\pi, f_{\omega^{1}_{-j}}}\right) \leq \frac{4\tilde{c}^2}{P^{2\gamma}} N_{j,\pi},$$

where $N_{j,\pi} := \mathbb{E}^{n-1}_{\pi, f_{\omega^{-1}_{-j}}} \mathbb{E}_X \sum_{s=1}^n 1_{\{\pi_s(X_s, Z_{s-1})=1, X_s \in B_j\}}$. Thus,

$$\sum_{t=1}^n Q_t^j \geq n \frac{2^{m-1}}{4P^d} \exp\left(-\frac{4\tilde{c}^2}{P^{2\gamma}} N_{j,\pi}\right)$$

On the other hand, one also has

$$\begin{aligned}
\sum_{t=1}^n Q_t^j &= \sum_{t=1}^n \sum_{\omega_{-j} \in \Omega_{m-1}} \sum_{i \in \{-1,1\}} \mathbb{E}^{t-1}_{\pi, f_{\omega^{i}_{-j}}} [\mathbb{P}_X(\pi_t(X_t, Z_{t-1}) \neq i, X_t \in B_j)] \\
&\geq \sum_{t=1}^n \sum_{\omega_{-j} \in \Omega_{m-1}} \mathbb{E}^{t-1}_{\pi, f_{\omega^{-1}_{-j}}} [\mathbb{P}_X(\pi_t(X_t, Z_{t-1}) = 1, X_t \in B_j)] \\
&= 2^{m-1} N_{j,\pi}.
\end{aligned}$$

Using the above two displays in (187) yields

$$\begin{aligned}
\sup_{f \in \mathcal{C}} \mathbb{E}^n_{\pi,f}[S_n(\pi)] &\geq \frac{1}{2} \sum_{j=1}^m \max\left(\frac{n}{4P^d} \exp\left(-\frac{4\tilde{c}^2}{P^{2\gamma}} N_{j,\pi}\right), N_{j,\pi}\right) \\
&\geq \frac{1}{4} \sum_{j=1}^m \left(\frac{n}{4P^d} \exp\left(-\frac{4\tilde{c}^2}{P^{2\gamma}} N_{j,\pi}\right) + N_{j,\pi}\right) \\
&\geq \frac{m}{4} \inf_{z \geq 0} \left(\frac{n}{4P^d} \exp\left(-\frac{4\tilde{c}^2}{P^{2\gamma}} z\right) + z\right).
\end{aligned}$$

The unique

$$z^* = \operatorname*{argmin}_{z \geq 0} \left\{\frac{n}{4P^d} \exp\left(-\frac{4\tilde{c}^2}{P^{2\gamma}} z\right) + z\right\} = \frac{P^{2\gamma}}{4\tilde{c}^2} \log\left(\tilde{c}^2 n P^{-d-2\gamma}\right)$$

is strictly positive if and only if $P < \left(n\tilde{c}^2\right)^{1/(d+2\gamma)}$ in which case, upon choosing $P = \lceil 0.5 \left(n\tilde{c}^2\right)^{1/(d+2\gamma)} \rceil$ and recalling $m = \lceil 0.5 P^{d-\alpha\gamma} \rceil$, we get

$$\begin{aligned}
\sup_{f \in \mathcal{C}} \mathbb{E}^n_{\pi,f}[S_n(\pi)] &\geq \frac{m}{4} \frac{P^{2\gamma}}{4\tilde{c}^2} \ln\left(\tilde{c}^2 n P^{-d-2\gamma}\right) \geq \frac{P^{d+\gamma(2-\alpha)}}{32\tilde{c}^2} \ln(2^{d+2\gamma}) \\
&\geq \frac{0.5^{d+\gamma(2-\alpha)}}{32\tilde{c}^2} (n\tilde{c}^2)^{\frac{d+\gamma(2-\alpha)}{d+2\gamma}} \ln(2^{d+2\gamma}) \geq \frac{0.5^{d+2} d \ln(2)}{32} \tilde{c}^{-\frac{2\alpha\gamma}{d+2\gamma}} n^{1-\frac{\alpha\gamma}{d+2\gamma}} \\
&\geq \frac{0.5^{d+2} d \ln(2)}{32} \left(\frac{c_1 d}{1+\bar{a}-\delta} \vee 1\right)^{-\frac{2\alpha\gamma}{d+2\gamma}} n^{1-\frac{\alpha\gamma}{d+2\gamma}} \geq \frac{0.5^{d+2} d \ln(2)}{32} \left(\frac{c_1 d}{1+\bar{a}-\delta} \vee 1\right)^{-\frac{2}{d+2}} n^{1-\frac{\alpha\gamma}{d+2\gamma}} \\
&= c_3 n^{1-\frac{\alpha\gamma}{d+2\gamma}}
\end{aligned}$$

for a constant $c_3 = \frac{0.5^{d+2} d \ln(2)}{32} \left(\frac{c_1 d}{1+\bar{a}-\delta} \vee 1 \right)^{-\frac{2}{d+2}}$ depending on neither α nor γ. $\square$

Proof of Theorem 5.2. The proof of this theorem relies on Theorem 5.1 and the notation used is as in that theorem. Let

$$\mathcal{F}(\gamma) = \left\{ f : [0,1]^d \to [1/4, 3/4] \text{ such that } |f(x_1) - f(x_2)| \leq 1/2 ||x_1 - x_2||^\gamma \text{ for all } x_1, x_2 \in [0,1]^d \right\}.$$

Note that $\bigcup_{\gamma > 0} \mathcal{F}(\gamma) \subseteq C[0,1]^d$. Thus, since c_l in Theorem 5.1 does not depend on γ, we get

$$\sup_{f \in C_{[0,1]^d}} \mathbb{E}^n_{\pi,f}[R_n(\pi)] \geq \sup_{f \in \bigcup\limits_{\gamma>0} \mathcal{F}(\gamma)} \mathbb{E}^n_{\pi,f}[R_n(\pi)] \geq c_l n.$$

$\square$

Proof of Theorem 5.3. Fix a policy π and choose $\alpha \in (0,1)$ such that $\gamma\alpha/(2\gamma + d) < \varepsilon$. Observe that $\mathcal{S} \subseteq \mathcal{S}_0$ for all $\alpha > 0$. Furthermore, since $c_l = c_l(\alpha)$ in Theorem 5.1 equals $\tilde{C}(\alpha) c_3^{1+1/\alpha}$ with $\tilde{C}(\alpha)$ as in Lemma D.5 (the dependence on α is suppressed there) and c_3 as in the last line of the proof of Theorem 5.1 one has that

$$\sup_{(F^1,F^2) \in \mathcal{S}_0} \mathbb{E}[R_n(\pi, F^1, F^2)] \geq \sup_{(F^1,F^2) \in \mathcal{S}} \mathbb{E}[R_n(\pi, F^1, F^2)] \geq c_l(\alpha) n^{1 - \frac{\gamma(1+\alpha)}{2\gamma+d}} \geq c_l(\alpha) n^{1-\frac{\gamma}{2\gamma+d}} n^{-\varepsilon}.$$

Since α depends on ε, we write $c_l(\varepsilon)$ instead of $c_l(\alpha)$. $\square$

Proof of Theorem 5.4. Throughout this proof we shall use notation defined in the proof of Theorem 5.1. Fix a policy $\pi \in \tilde{\Pi}$ and let $m_1 = 2m$ with $m = \lceil 0.5 P^{d-\alpha\gamma} \rceil$. Note that $m_1 \leq P^d$ for $P \geq P_0$ for P_0 sufficiently large. Set $P = P_0$. Define $\omega^{(1)} := (\iota_m, -\iota_m)$, where ι_m is a row vector of ones of length m, and let $\omega^{(-1)} := -\omega^{(1)}$. Set $f_i = f_{\omega^{(i)}}$ for $i \in \{-1, 1\}$. Given $f \in \{f_1, f_{-1}\} \subseteq \mathcal{C}_{m_1}$, define $\mathbb{P}_{f,1}$, $\mathbb{P}_{1/2,1}$, $\mathbb{P}_{1/2,-1}$ and $\mathbb{P}^t_{\pi,f}$ as in Step 2 of the proof of Theorem 5.1. From the argument given in Step 3 of that proof it follows that

$$\mathcal{C}_1 := \{(H_{g_{f(x)}}, H_{g_{1/2}}) : f \in \{f_1, f_{-1}\}\} \subseteq \mathcal{S}. \tag{189}$$

Next, we define $\tilde{Z}_t = (Y_{\pi_t(Z_{t-1}),t}, ..., Y_{\pi_1,1})$, and we denote the distribution of $\tilde{Z}_t$ by $\tilde{\mathbb{P}}^t_{\pi,f}$. We claim that $\tilde{\mathbb{P}}^t_{\pi,f_1} = \tilde{\mathbb{P}}^t_{\pi,f_{-1}}$ for every $t \geq 1$. To this end, note first that from $\pi \in \tilde{\Pi}$ it follows that $\tilde{Z}_t = F_t(Y_t, \ldots, Y_1)$ for some measurable function F_t. Hence, since the Y_t are i.i.d., in order to prove the claim it is enough to verify that the distribution of the random vector Y_1 does not depend on $f \in \mathcal{C}_1$. Using the notation in Step 2 of the proof of Theorem 5.1 this is equivalent to: $\mathbb{P}_{f_1,1/2}(A_1 \times A_2 \times [0,1]^d) = \mathbb{P}_{f_{-1},1/2}(A_1 \times A_2 \times [0,1]^d)$ for all Borel sets A_1, A_2 in $[0,1]$. To verify this equivalent condition, we write $\mathbb{P}_{f_1,1/2}(A_1 \times A_2 \times [0,1]^d) = \mathbb{P}_{1/2,-1}(A_1) \int \mathbb{P}_{f_1,1}(A_2, x) \mathbb{P}_X(dx)$ as

$$\mathbb{P}_{1/2,-1}(A_1) \sum_{j=1}^{P^d} \int_{B_j} \mathbb{P}_{f_1,1}(A_2, x) \mathbb{P}_X(dx) = \mathbb{P}_{1/2,-1}(A_1) \sum_{j=1}^{P^d} \int_{B_j} \mathbb{P}_{f_{-1},1}(A_2, x) \mathbb{P}_X(dx), \tag{190}$$

the latter coinciding with $\mathbb{P}_{f_{-1},1/2}(A_1 \times A_2 \times [0,1]^d)$, which proves the claim. Here we have used that $\omega^{(1)} = -\omega^{(-1)}$ implies

$$\int_{B_j} \mathbb{P}_{f_1,1}(A_2, x) \mathbb{P}_X(dx) = \int_{B_{2m+1-j}} \mathbb{P}_{f_{-1},1}(A_2, x) \mathbb{P}_X(dx) \quad \text{for } j = 1, \ldots, 2m, \tag{191}$$

and that

$$\int_{B_j} \mathbb{P}_{f_1,1}(A_2,x)\mathbb{P}_X(dx) = \int_{B_j} \mathbb{P}_{f_{-1},1}(A_2,x)\mathbb{P}_X(dx) \quad \text{for } j = 2m+1,\dots,P^d. \tag{192}$$

Now, to prove the theorem, by Lemma D.5 it suffices to show that $\sup_{\mathcal{S}} \mathbb{E}[S_n(\pi)]$ increases linearly in n. Using Equation (189) we see that $\sup_{\mathcal{S}} \mathbb{E}[S_n(\pi)] \geq \sup_{\mathcal{C}_1} \mathbb{E}[S_n(\pi)] = \sup_{f_1,f_{-1}} \mathbb{E}^n_{\pi,f}[S_n(\pi)]$. Next, arguing as in Equation (187) and exploiting $\pi \in \tilde{\Pi}$ we obtain

$$\begin{aligned}
\sup_{f_1,f_{-1}} \mathbb{E}^n_{\pi,f}[S_n(\pi)] &\geq \sup_{i\in\{-1,1\}} \sum_{j=1}^{m}\sum_{t=1}^{n} \mathbb{E}^{t-1}_{\pi,f_i}\left[\mathbb{P}_X\left(\pi_t(\tilde{Z}_{t-1}) \neq i, X_t \in B_j\right)\right] \\
&= \sup_{i\in\{-1,1\}} \sum_{j=1}^{m}\sum_{t=1}^{n} \mathbb{P}^{t-1}_{\pi,f_i}\left(\pi_t(\tilde{Z}_{t-1}) \neq i\right)\mathbb{P}_X(B_j) \\
&= P^{-d} \sup_{i\in\{-1,1\}} \sum_{j=1}^{m}\sum_{t=1}^{n} \tilde{\mathbb{P}}^{t-1}_{\pi,f_i}\left(\pi_t(\tilde{Z}_{t-1}) \neq i\right) \\
&\geq (2P^d)^{-1} \sum_{j=1}^{m}\sum_{t=1}^{n}\left(\tilde{\mathbb{P}}^{t-1}_{\pi,f_1}\left(\pi_t(\tilde{Z}_{t-1}) \neq -1\right) + \tilde{\mathbb{P}}^{t-1}_{\pi,f_{-1}}\left(\pi_t(\tilde{Z}_{t-1}) \neq 1\right)\right) = \frac{mn}{2P^d} \geq \frac{1}{4P_0^{\alpha\gamma}}n,
\end{aligned}$$

where we used independence of X_t and $\tilde{Z}_{t-1}$ to obtain the first equality, that each summand in the last double sum equals one (recall that $\tilde{\mathbb{P}}^{t-1}_{\pi,f_1} = \tilde{\mathbb{P}}^{t-1}_{\pi,f_{-1}}$) to obtain the last equality, and the definition of m to obtain the final lower bound. □

References

Athey, S. and S. Wager (2017). Efficient policy learning. *arXiv preprint arXiv:1702.02896*.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory 2*(3), 244–263.

Audibert, J.-Y. and A. B. Tsybakov (2007). Fast learning rates for plug-in classifiers. *Annals of Statistics 35*(2), 608–633.

Auer, P., N. Cesa-Bianchi, and P. Fischer (2002). Finite-time analysis of the multiarmed bandit problem. *Machine Learning 47*(2-3), 235–256.

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

Bubeck, S. and N. Cesa-Bianchi (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning 5*(1), 1–122.

Cassel, A., S. Mannor, and A. Zeevi (2018). A general approach to multi-armed bandits under risk criteria. *arXiv preprint arXiv:1806.01380*.

Chakravarty, S. R. (1983). A new index of poverty. *Mathematical Social Sciences 6*(3), 307–313.

Chakravarty, S. R. (2009). *Inequality, polarization and poverty*. Springer.

Chamberlain, G. (2000). Econometrics and decision theory. *Journal of Econometrics 95*(2), 255–283.

Cover, T. M. and J. A. Thomas (2012). *Elements of information theory*. John Wiley & Sons.

Cowell, F. (2011). *Measuring inequality*. Oxford University Press.

Cowell, F. A. (1980). Generalized entropy and the measurement of distributional change. *European Economic Review 13*(1), 147–159.

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

Embrechts, P. and M. Hofert (2013). A note on generalized inverses. *Mathematical Methods of Operations Research 77*(3), 423–432.

Folland, G. B. (2013). *Real Analysis: Modern Techniques and their Applications*. Wiley.

Foster, J., J. Greer, and E. Thorbecke (1984). A class of decomposable poverty measures. *Econometrica 52*(3), 761–766.

Foster, J., J. Greer, and E. Thorbecke (2010). The Foster–Greer–Thorbecke (FGT) poverty measures: 25 years later. *The Journal of Economic Inequality 8*(4), 491–524.

Gastwirth, J. L. (1971). A general definition of the Lorenz curve. *Econometrica 39*(6), 1037–1039.

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

Kakwani, N. (1980). On a class of poverty measures. *Econometrica*, 437–446.

Kakwani, N. (1986). *Analyzing redistribution policies: a study using Australian data*. Cambridge University Press.

Kitagawa, T. and A. Tetenov (2017). Equality-minded treatment choice. *Working paper*.

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

Kock, A. B. and M. Thyrsgaard (2017). Optimal sequential treatment allocation. *arXiv preprint arXiv:1705.09952*.

Kolm, S.-C. (1976a). Unequal inequalities. I. *Journal of Economic Theory 12*(3), 416–442.

Kolm, S.-C. (1976b). Unequal inequalities. II. *Journal of Economic Theory 13*(1), 82–111.

Kundu, A. and T. E. Smith (1983). An impossibility theorem on poverty indices. *International Economic Review*, 423–434.

Lai, T. L. and H. Robbins (1985). Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics 6*(1), 4–22.

Lambert, P. J. (2001). *The distribution and redistribution of income*. Manchester University Press.

Mammen, E. and A. B. Tsybakov (1999). Smooth discrimination analysis. *Annals of Statistics 27*(6), 1808–1829.

Manski, C. F. (2004). Statistical treatment rules for heterogenous populations. *Econometrica*.

Massart, P. (1990). The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *Annals of Probability 18*(3), 1269–1283.

Mehran, F. (1976). Linear measures of income inequality. *Econometrica 44*(4), 805–809.

Minassian, D. (2007). A mean value theorem for one-sided derivatives. *The American Mathematical Monthly 114*(1), 28.

Perchet, V. and P. Rigollet (2013). The multi-armed bandit problem with covariates. *Annals of Statistics*, 693–721.

Rigollet, P. and A. Zeevi (2010). Nonparametric bandits with covariates. *Proceedings of COLT*.

Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of the Americal Mathematical Society 58*(5), 527–535.

Rothe, C. (2010). Nonparametric estimation of distributional policy effects. *Journal of Econometrics 155*(1), 56–70.

Rothe, C. (2012). Partial distributional policy effects. *Econometrica 80*(5), 2269–2301.

Sen, A. (1976). Poverty: an ordinal approach to measurement. *Econometrica*, 219–231.

Serfling, R. J. (2009). *Approximation Theorems of Mathematical Statistics*, Volume 162. Wiley.

Shalev-Shwartz, S. (2012). Online learning and online convex optimization. *Foundations and Trends® in Machine Learning 4*(2), 107–194.

Stoye, J. (2009). Minimax regret treatment choice with finite samples. *Journal of Econometrics 151*(1), 70–81.

Stoye, J. (2012). Minimax regret treatment choice with covariates or with limited validity of experiments. *Journal of Econometrics 166*(1), 138–156.

Tetenov, A. (2012). Statistical treatment choice based on asymmetric minimax regret criteria. *Journal of Econometrics 166*(1), 157–165.

Theil, H. (1967). *Economics and Information Theory*. North-Holland: Amsterdam.

Tsybakov, A. B. (2004). Optimal aggregation of classifiers in statistical learning. *Annals of Statistics 32*(1), 135–166.

Tsybakov, A. B. (2009). *Introduction to nonparametric estimation*. Springer.