Optimal Robust Policy for Feature-Based Newsvendor

Luhao Zhang, Jincheng Yang
Department of Mathematics, The University of Texas at Austin, luhaozhang,jincheng@utexas.edu

Rui Gao
Department of Information, Risk and Operations Management, The University of Texas at Austin, rui.gao@mccombs.utexas.edu

We study policy optimization for the feature-based newsvendor, which seeks an end-to-end policy that renders an explicit mapping from features to ordering decisions. Most existing works restrict the policies to some parametric class that may suffer from sub-optimality (such as affine class) or lack of interpretability (such as neural networks). Differently, we aim to optimize over all functions of features. In this case, the classical empirical risk minimization yields a policy that is not well-defined on unseen feature values. To avoid such degeneracy, we consider a Wasserstein distributionally robust framework. This leads to an adjustable robust optimization, whose optimal solutions are notoriously difficult to obtain except for a few notable cases. Perhaps surprisingly, we identify a new class of policies that are proven to be exactly optimal and can be computed efficiently. The optimal robust policy is obtained by extending an optimal robust in-sample policy to unobserved feature values in a particular way and can be interpreted as a Lipschitz regularized critical fractile of the empirical conditional demand distribution. We compare our method with several benchmarks using synthetic and real data and demonstrate its superior empirical performance.

Key words: Side information, Contextual decision-making, Inventory management, Adjustable robust optimization

1. Introduction

The newsvendor model is a classical and fundamental problem in operations management, but faces new challenges in the era of big data. More often than not, numerous feature information—temporal, spatial, social, or economical—are available prior to the decision-making and reveals partial information on the product demand. The feature information reduces the uncertainty and helps the decision-maker customize ordering decisions to each individual feature realization. It is crucial to involve the feature information in decision-making. Otherwise, the decision may be inconsistent, namely, not converging to the true optimal policy even with an infinite amount of data (Ban and Rudin 2019). The ordering decision is often made for a population of features but not for a single feature realization. For example, ordering decisions are made for multiple shelves at different locations in selected time windows, or for a number of customers with different demographics. In these cases, it would make sense to consider the average performance over the entire distribution of features. In this work, we are interested in the *policy optimization* decision-making problem. It seeks a policy (a.k.a. decision rule) that outputs an ordering decision for every feature value to minimize the overall average cost. We refer to the paragraph Related Literature for a review of other popular approaches for the feature-based newsvendor.

If the underlying distribution is known, then the true optimal policy would equal the conditional critical fractile of the demand distribution under each feature value. Unfortunately, many real-world problems comprise a potentially large set of feature values that historical data cannot exhaust. Thus, the true underlying conditional distribution of the demand under a new feature value is likely unknown. In this case, a natural way is to replace the unknown underlying distribution of demands and features with its empirical counterpart. However, the resulting empirical risk minimization problem produces a pathological policy that can take arbitrary values on unseen feature values; see more detailed discussion in Section 2.2.

The pathological behavior of empirical feature-based newsvendor motivates the development of methods to generalize ordering decisions to unseen feature values. The most common approach is parameterization, namely, restricting the search to a parametric policy class. For example, Section 2.3.1 of Ban and Rudin (2019) studies affine policies, which can be efficiently solved using linear or

convex optimization methods. The affine class can be restrictive and sub-optimal. Indeed, numerical experiments in Ban and Rudin (2019) show that affine policies are outperformed by their proposed kernel optimization method in the same paper. One possible remedy is to consider nonlinear transformations of features (basis functions). Thereby, one can enlarge the policy search space to an arbitrarily complex class with coefficients affinely dependent on the basis functions (Bertsimas and Koduri 2022). However, specifying nonlinear transformations with good interpretability is a fundamentally challenging question. Similarly, neural-network polices (Oroojlooyjadid et al. 2020, Meng et al. 2021) may have nice empirical performance but are often hard to interpret and data-demanding. In summary, in existing methods, there is a trade-off between the richness of the policy class and its interpretability/tractability. As such, the following question remains open: *Can we find an optimal policy without restricting onto a parametric family while still maintaining computational efficiency and interpretability*?

To answer this question, we consider a distributionally robust policy optimization framework. More specifically, it optimizes over all policies that are measurable functions of the features and does not parameterize the policy class. Moreover, it involves a minimax Wasserstein distributionally robust formulation (Kuhn et al. 2019) that hedges against data uncertainty on the demand and features and helps to resolve the pathological issue of the empirical feature-based newsvendor. We remark that most literature on Wasserstein distributionally robust optimization has been focusing on deriving tractable reformulations when the decision variable is a finite-dimensional vector. Our main challenge here, however, is on the infinite-dimensional policy optimization. This distinguishes our model from most existing works.

Our formulation belongs to the class of *adjustable robust optimization* (Yanıkoğlu et al. 2019). Computationally, this class of problems involves a challenging infinite-dimensional functional optimization over the space of policies, which is "typically severely computationally intractable" (Ben-Tal et al. 2009, Chapter 14.2.3). Without parameterization, the optimal solutions are generally unknown except for a few notable cases. Perhaps surprisingly, by utilizing the structure of the problem, we are able to identify a new class of policies that are proven to be optimal and can be computed efficiently. More specifically,

(I) We show that the infinite-dimensional distributionally robust policy optimization problem can be solved in two steps. First, we solve a finite-dimensional robust policy optimization on the observed (in-sample) feature values only. Then, we generalize this in-sample optimal policy to the full feature space via a specific interpolation technique (Theorem 1), which we term the *Shapley policy*. This provides a new class of optimal policies for adjustable optimization that may be of independent interest.

(II) We further show that the optimal robust policy can be interpreted as a regularized critical fractile that regularizes the variation (measured by its Lipschitz norm with respect to the features) of the policy; see Figure 1 as an illustration. Based on this connection, the optimal robust policy optimization can be solved by linear programming. We compare the out-of-sample cost of the Shapley policy with various benchmarks using synthetic and real data, which demonstrate its superior empirical performance.

1.1. Related Literature

On feature-based newsvendor. We consider a policy optimization formulation to minimize the expected newsvendor cost over a population. Once solved, it outputs an ordering decision for every new feature value without explicitly estimating any intermediate value, such as the conditional expected newsvendor cost or solving/tuning any additional optimization problems. As such, this approach contrasts with two other methods that have been considered in the literature: *predict-then-optimize* and *conditional stochastic optimization*.

Traditionally, a widely used approach to this problem is a two-step predict-then-optimize process that first predicts the conditional demand distribution given a new feature value and then optimizes

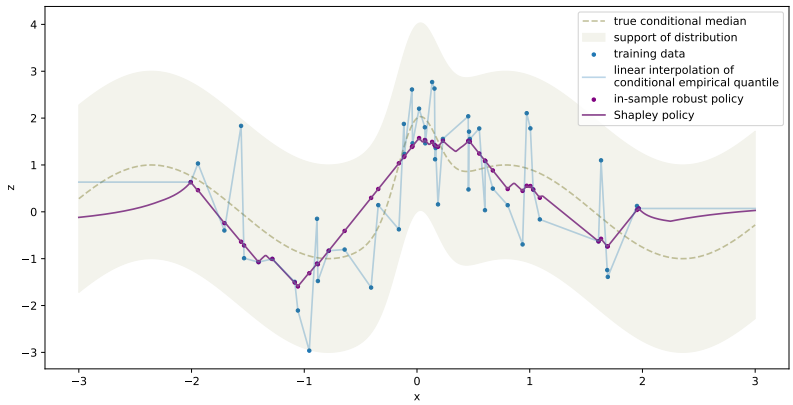


Figure 1 An illustration of policies where training data are generated from a two-dimensional continuous distribution of feature-demand pairs (x, z). Our proposed Shapley policy has a smaller variation (Lipschitz norm) compared to the linear interpolation of conditional quantiles of the empirical distribution.

for the ordering quantity (e.g., Toktay and Wein (2001), Zhu and Thonemann (2004)). There are some theoretical guarantees in this approach discussed in El Balghiti et al. (2019), Hu et al. (2022). However, as pointed out by Liyanage and Shanthikumar (2005), Ban and Rudin (2019) among others, for high-dimensional problems, the demand model specification may be challenging, and the first-step estimation error can be amplified in the second-step optimization.

To integrate prediction and decision optimization, a recently emerging approach, conditional stochastic optimization, directly minimizes the conditional expected newsvendor cost given the new feature value without explicitly estimating the conditional demand distribution. For example, the pioneer works of Ban and Rudin (2019, kernel optimization) and Bertsimas and Kallus (2020) estimated the conditional expected cost by a weighted average of the observed costs associated with historical observations of features and demand. They demonstrated the excellent generalization capability of this approach both theoretically and empirically. Nonetheless, as commented in Kallus and Mao (2022), these weights are chosen based on the prediction accuracy only but not tailored to the newsvendor cost structure, which leaves room for further integration of prediction and optimization. Indeed, Kallus and Mao (2022) proposed a learning scheme to construct random forests and weights directly targeted for the decision optimization objective and illustrated its effectiveness.

On joint prediction and optimization. Beyond the newsvendor model, there is a growing literature on the integration of prediction and optimization for general stochastic optimization. These works differ mainly in how the conditional expected cost is estimated, for example, based on Nadaraya-Watson kernel regression (Ban and Rudin 2019, Srivastava et al. 2021), trees and forests (Ban et al. 2019, Bertsimas and Kallus 2020, Kallus and Mao 2022), Dirichlet process (Hannah et al. 2010), local regression and classification (Bertsimas and Kallus 2020), smart prediction-then-optimization (Elmachtoub and Grigas 2022), robustness optimization and regularization (Zhu et al. 2022, Loke et al. 2020, Esteban-Pérez and Morales 2021), empirical residuals (Kannan et al. 2020a,b), etc. All these works aim at solving a conditional problem under every single feature value separately. In contrast, our framework marginalizes these conditional problems and provides an end-to-end framework aiming at finding the optimal policy in a minimax sense.

On robust newsvendor. To ensure good generalization capability, many existing works exploit robust formulation for inventory models by considering various uncertainty sets based on moments (Scarf 1958, Gallego and Moon 1993, Perakis and Roels 2008, Han et al. 2014, Xin and Goldberg 2021), precentiles (Gallego et al. 2001), shape information (Perakis and Roels 2008, Hanasusanto et al. 2015a, Natarajan et al. 2018), tail information (Das et al. 2021), temporal dependence (See and Sim 2010, Xin and

Goldberg 2022, Carrizosa et al. 2016), total variation distance (Rahimian et al. 2019a,b), phi-divergence (Ben-Tal et al. 2013, Wang et al. 2016, Bayraksan and Love 2015, Fu et al. 2021), Wasserstein distance (Lee et al. 2012, Mohajerin Esfahani and Kuhn 2018, Gao and Kleywegt 2022, Lee et al. 2020, Chen and Xie 2021), etc. Except for See and Sim (2010), most of these works do not consider feature information. In our analysis, we use the duality results for Wasserstein distributionally robust optimization (Gao and Kleywegt 2022) to obtain an equivalent reformulation of the worst-case newsvendor cost for a fixed policy. Nevertheless, we would like to emphasize that the main challenge and focus of this paper is on the policy optimization that is not studied by existing distributionally robust optimization literature.

On adjustable robust optimization. Various decision-rule approaches have been studied extensively in the literature, including affine families (Chen et al. 2008, Bertsimas et al. 2010, 2011, Bertsimas and Goyal 2012, Iancu et al. 2013, El Housni and Goyal 2021, Bertsimas et al. 2022, Georghiou et al. 2021), binary decision rules (Bertsimas and Georghiou 2015), k-adaptability (Hanasusanto et al. 2015b, 2016, Subramanyam et al. 2019), iterative splitting of uncertainty sets (Postek and Hertog 2016), non-parametric Markovian stopping rules (Sturt 2021), etc. Most of these works do not consider feature information, expect for Bertsimas et al. (2022) which considers dynamic decision-making with side information using affine decision rules. Unlike this work, we consider general decision rules in a static setting. Moreover, policy classes with provable zero sub-optimality gap are rare commodities (Bertsimas et al. 2010, Bertsimas and Goyal 2012, Iancu et al. 2013, Sturt 2021, Georghiou et al. 2021). Our result complements the literature by providing a new class of provably optimal policies for adjustable robust optimization.

The rest of the paper proceeds as follows. Our model setup and main results are stated in Section 2. We provide a proof sketch with geometric intuitions of our proposed `Shapley` policy in Section 3. In Section 4, we make a connection between our robust formulation and Lipschitz regularization and discuss its computational and statistical consequences. Finally, we present numerical results in Section 5 and conclude in Section 6. Proofs and additional results are deferred to Appendices.

2. Model Setup and Main Results

In this section, we present our model and state our main result.

2.1. Policy Optimization for Feature-based Newsvendor

Consider a company selling a perishable product who needs to decide the ordering quantity y before a random demand $Z \in \mathcal{Z} := \mathbb{R}_+$ is observed. Let h, b represent the unit holding cost and the unit backorder cost, respectively. The total cost Ψ is then computed as

$$\Psi(y, z) := h(y - z)_+ + b(z - y)_+,$$

where a_+ denotes the positive part of $a \in \mathbb{R}$. In the classical newsvendor problem with a known demand distribution $\mathbb{P}$, the optimal ordering quantity is well known as the critical fractile, i.e. the $\frac{b}{b+h}$-quantile of the demand distribution.

Suppose, prior to making the ordering decision, that the decision maker has access to some additional feature information which would help to make a better estimation of the demand or cost. We use a covariate $X \in \mathcal{X} \subset (\mathbb{R}^d, \|\cdot\|)$ to represent such feature information. For repeated sales, it is reasonable to find an ordering quantity y that minimizes the conditional expected cost upon observing a feature realization $X = x$:

$$\inf_{y \in \mathcal{Z}} \mathbb{E}[\Psi(y, Z) \mid X = x],$$

where the expectation is taken with respect to the conditional demand distribution of Z given $X = x$. Such an objective has been considered in the pioneer work of Ban and Rudin (2019) on the big data

newsvendor. If the true underlying demand distribution is known, the true optimal ordering quantity equals the critical fractile of the true conditional demand distribution.

Using the interchangeability principle (e.g., Shapiro et al. (2014, Theorem 7.92)), we have that

$$\mathbb{E}\left[\inf_{y\in\mathcal{Z}} \mathbb{E}[\Psi(y,Z) \mid X]\right] = \inf_{f:\mathcal{X}\to\mathcal{Z}} \mathbb{E}[\Psi(f(X),Z)]. \tag{1}$$

Here, on the left-hand side of (1), the outer expectation is taken over the marginal distribution of the feature X. For each feature realization, the corresponding y is chosen as the true optimal ordering quantity that minimizes the conditional expected cost. The value of the left-hand side of (1) is termed the *optimal true risk* in the literature (Ban and Rudin 2019). Whereas on the right-hand side of (1), the expectation is taken over the joint distribution of feature X and demand Z. It finds the optimal policy among the set of all measurable functions that map every feature X to an ordering quantity $f(X)$ so as to minimize the marginalized expected cost. Whenever the optimizers on both sides of (1) exist, it holds that for any x in the support of X, the optimal policy on the right-hand side takes a value as the conditional minimizer of the left-hand side when $X = x$. Note that the true risk on the left-hand side of (1) is the numerical performance measure considered by Ban and Rudin (2019), Kallus and Mao (2022) among other literature on decision-making with feature information.

2.2. Distributionally Robust Formulation

In practice, the true underlying distribution is often unknown. Instead, the decision maker often has historical data at disposal. Suppose the historical data contains n observations. We first group them into K groups according to distinct feature values $\widehat{x}_k$, $k = 1,\ldots,K$. Each $\widehat{x}_k$ is associated with demand observations $\widehat{z}_{ki}$, $i = 1,\ldots,n_k$, where $\sum_{k=1}^{K} n_k = n$. Thus, these observations formulate an empirical distribution of the form

$$\widehat{\mathbb{P}} = \frac{1}{n}\sum_{k=1}^{K}\sum_{i=1}^{n_k} \delta_{(\widehat{x}_k,\widehat{z}_{ki})},$$

where $\widehat{x}_k$'s are distinct feature values, $k = 1,\ldots,K$, and $\widehat{z}_{ki}$'s are not necessarily distinct. Let us denote by $[K]$ the set $\{1,2,\ldots,K\}$.

To solve the right-hand side of (1), a conventional wisdom is to consider the empirical risk minimization by replacing the true distribution with the empirical distribution $\widehat{\mathbb{P}}$

$$\inf_{f:\mathcal{X}\to\mathcal{Z}} \mathbb{E}_{(X,Z)\sim\widehat{\mathbb{P}}}[\Psi(f(X),Z)]. \tag{2}$$

Unfortunately, this would yield a degenerate solution that is only defined on the set of historical observations of features $\widehat{\mathcal{X}} := \{\widehat{x}_k : k \in [K]\}$, but can take arbitrary values elsewhere. In addition, suppose $\widehat{f}$ is the optimal policy of (2). Then for every $\widehat{x} \in \widehat{\mathcal{X}}$, $\widehat{f}(\widehat{x})$ is the critical fractile of the empirical conditional distribution. When the historical samples are generated from some continuous underlying distribution, then with probability one we have $n_k = 1$ for all k and $\widehat{f}(\widehat{x}_k) = \widehat{z}_{k1}$, which could be far away from the critical fractile of true conditional distribution.

Motivated by the above degeneracy of empirical risk minimization, we consider a minimax distributionally robust formulation which finds a decision hedging against a set $\mathfrak{M}$ of relevant probability distributions $\inf_{f\in\mathcal{F}} \sup_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)]$, where $\mathcal{F}$ is the set of all measurable functions on $\mathcal{X}$. In our formulation, we choose $\mathfrak{M}$ to be a ball of distributions that are within ρ Wasserstein distance to a nominal distribution ($\rho \geq 0$). It is a natural choice since such distributional uncertainty set is data-driven and incorporates distributions on unseen feature values (see, e.g., Kuhn et al. (2019)). Let $\|\cdot\|_*$ denote the dual norm of the norm $\|\cdot\|$ on $\mathcal{X}$. Let $\mathcal{P}_1(\mathcal{X}\times\mathcal{Z})$ be the set of probability distributions on $\mathcal{X}\times\mathcal{Z}$ with finite first moment. The *Wasserstein distance* (of order 1) is defined as

$$\mathcal{W}(\mathbb{P},\mathbb{Q}) := \inf_{\gamma\in\Gamma(\mathbb{P},\mathbb{Q})} \mathbb{E}_{(\tilde{X},\tilde{Z}),(X,Z)\sim\gamma}[\|\tilde{X}-X\| + |\tilde{Z}-Z|], \tag{3}$$

where $\Gamma(\mathbb{P},\mathbb{Q})$ denotes the set of probability distributions on $(\mathcal{X}\times\mathcal{Z})^2$ with marginals $\mathbb{P},\mathbb{Q}\in\mathcal{P}_1(\mathcal{X}\times\mathcal{Z})$. Let $\widehat{\mathbb{P}}_{\widehat{X}}$ be the x-marginal distribution of $\widehat{\mathbb{P}}$. Consider the following Wasserstein robust feature-based newsvendor problem

$$v_P := \inf_{f\in\mathcal{F}} \sup_{\mathbb{P}\in\mathcal{P}_1(\mathcal{X}\times\mathcal{Z})} \left\{\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)] : \mathcal{W}(\mathbb{P},\widehat{\mathbb{P}})\le\rho\right\}, \tag{P}$$

where ρ is the radius of the 1-Wasserstein ball where we use to construct the uncertainty set. Note that (P) can be viewed as a special case of policy optimization for two-stage Wasserstein distributionally robust optimization (Bertsimas et al. 2022). Throughout the paper, we assume $b>0$ and $0\le h\le b$. We remark that when $b<h$, all results in the paper still hold for sufficiently small ρ (see Remark EC.1 in Appendix D).

2.3. Main Results

In this section, we present the main result of this paper, which provides an explicit, tractable solution to the problem (P).

THEOREM 1. *Problem* (P) *can be solved in the following two steps:*

(I) *[In-sample robust policy] Solve the in-sample primal problem,*

$$v_{\hat{P}} := \min_{\widehat{f}:\widehat{\mathcal{X}}\to\mathcal{Z}} \sup_{\mathbb{P}\in\mathcal{P}_1(\widehat{\mathcal{X}}\times\mathcal{Z})} \left\{\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(\widehat{f}(X),Z)] : \mathcal{W}(\mathbb{P},\widehat{\mathbb{P}})\le\rho\right\}, \tag{$\widehat{\text{P}}$}$$

(II) *[Shapley Extension] With $\widehat{f}^*$ being the optimal solution to the linear programming problem above, an optimal policy f^* for problem* (P) *is defined by the following minimax matrix saddle point*

$$f^*(x) := \min_{j\in[K]}\max_{k\in[K]} A_{jk}(x) = \max_{k\in[K]}\min_{j\in[K]} A_{jk}(x).$$

where

$$A_{jk}(x) := \frac{\|x-\widehat{x}_k\|}{\|x-\widehat{x}_j\|+\|x-\widehat{x}_k\|}\widehat{f}^*(\widehat{x}_j) + \frac{\|x-\widehat{x}_j\|}{\|x-\widehat{x}_j\|+\|x-\widehat{x}_k\|}\widehat{f}^*(\widehat{x}_k).$$

REMARK 1. Theorem 1 shows that to solve the primal problem (P), it suffices to

(i) solve the problem that is solely based on the in-sample data. Here in-sample means that instead of considering the full feature space, we restrict our attention to historical observations of features only, which is a finite subset. Observe that the set of in-sample policies is $\widehat{\mathcal{F}} := \{\widehat{f}:\widehat{\mathcal{X}}\to\mathcal{Z}\}\subset\mathbb{R}_+^K$, hence the in-sample problem, is a finite-dimensional optimization problem;

(ii) extrapolate the optimal in-sample robust policy $\widehat{f}^*$ to the entire space $\mathcal{X}$, based on a novel extension defined in (II).

In Section 3, we will prove Theorem 1 through the in-sample dual problem ($\widehat{\text{D}}$)

$$v_{\hat{D}} := \min_{\widehat{f}:\widehat{\mathcal{X}}\to\mathcal{Z},\lambda\ge0} \left\{\lambda\rho + \mathbb{E}_{(\widehat{X},\widehat{Z})\sim\widehat{\mathbb{P}}}\left[\sup_{z\in\mathcal{Z}}\max_{x\in\widehat{\mathcal{X}}}\left\{\Psi(\widehat{f}(x),z)-\lambda\|x-\widehat{X}\|-\lambda|z-\widehat{Z}|\right\}\right]\right\}. \tag{$\widehat{\text{D}}$}$$

which is an equivalent reformulation of the *in-sample* robust primal problem ($\widehat{\text{P}}$), followed directly from duality on Wasserstein DRO (e.g., Mohajerin Esfahani and Kuhn (2018)).

In Section 4.1, we will show that the in-sample dual problem ($\widehat{\text{D}}$) is also equivalent to the following in-sample Lipschitz regularization problem

$$v_{\hat{R}} := \min_{\widehat{f}:\widehat{\mathcal{X}}\to\mathcal{Z}} \left\{(b\vee h)(1\vee\|\widehat{f}\|_{\text{Lip}})\rho + \mathbb{E}_{(\widehat{X},\widehat{Z})\sim\widehat{\mathbb{P}}}[\Psi(\widehat{f}(\widehat{X}),\widehat{Z})]\right\}, \tag{$\widehat{\text{R}}$}$$

which turns out to be equivalent to a finite-dimensional linear program that will be defined in Section 4.2.

In Section 3.2, we show that the matrix saddle point defined in Theorem 1(II) is the closed-form solution to the following Lipschitz constant minimization problem

$$f^*(x) = \arg\min_{y\in\mathbb{R}} \left\{ \max_{k\in[K]} \frac{|\widehat{f}^*(\widehat{x}_k) - y|}{\|\widehat{x}_k - x\|} \right\}. \tag{L}$$

This is a linear program with input x, $\{(\widehat{x}_k, \widehat{f}^*(\widehat{x}_k))\}_{k\in[K]}$ and an unknown decision variable y, which has better computational efficiency than the minimax expression in Theorem 1(II).

3. Proof of Main Results

In this section, we prove Theorem 1. Problem (P) is an infinite-dimensional optimization whose main difficulty is that, we need not only to assign an ordering quantity for every historical observation of features in $\widehat{\mathcal{X}}$ but also to each unseen feature values in $\mathcal{X} \setminus \widehat{\mathcal{X}}$. In Section 3.1, we prove that the Shapely extension from the in-sample robust problem ($\widehat{\text{P}}$) renders an optimal policy to the primal problem (P). In Section 3.2, we provide further intuition that drives behind the Shapely policy.

3.1. Shapley Policy and its Optimality

In this subsection, we show that the infinite-dimensional functional optimization (P) can be solved exactly by a novel extension of the solution to the finite-dimensional problem ($\widehat{\text{P}}$).

To begin with, applying strong duality for Wasserstein distributionally robust optimization (Lemma EC.1 in Appendix A) on the inner maximization of (P) and ($\widehat{\text{P}}$) respectively yields their strong dual problems

$$v_D := \inf_{f\in\mathcal{F}} \min_{\lambda\geq 0} \left\{ \lambda\rho + \mathbb{E}_{(\widehat{X},\widehat{Z})\sim\widehat{\mathbb{P}}} \left[\sup_{z\in\mathcal{Z}, x\in\mathcal{X}} \left\{ \Psi(f(x), z) - \lambda\|x - \widehat{X}\| - \lambda|z - \widehat{Z}| \right\} \right] \right\}, \tag{D}$$

$$v_{\widehat{D}} := \min_{\widehat{f}\in\widehat{\mathcal{F}}, \lambda\geq 0} \left\{ \lambda\rho + \mathbb{E}_{(\widehat{X},\widehat{Z})\sim\widehat{\mathbb{P}}} \left[\sup_{z\in\mathcal{Z}} \max_{x\in\widehat{\mathcal{X}}} \left\{ \Psi(\widehat{f}(x), z) - \lambda\|x - \widehat{X}\| - \lambda|z - \widehat{Z}| \right\} \right] \right\}. \tag{$\widehat{\text{D}}$}$$

Observe that ($\widehat{\text{D}}$) remains unchanged if one replaces the minimization over $\widehat{f}\in\widehat{\mathcal{F}}$ with minimization over $\widehat{f}\in\mathcal{F}$ because the objective value does not depend on the policy value outside $\widehat{\mathcal{X}}$. Thus, the main difference between the two problems above is on the set of x with respect to which the inner supremum is taken. It follows immediately that $v_D \geq v_{\widehat{D}}$ because the supremum in (D) is taken over a larger set. To show the other direction, it suffices to show that the minimizer $\widehat{f}^*$ of ($\widehat{\text{D}}$) admits an extension f^* such that for every $(\widehat{X}, \widehat{Z})$ in the support of $\widehat{\mathbb{P}}$,

$$\sup_{z\in\mathcal{Z}, x\in\mathcal{X}} \left\{ \Psi(f^*(x), z) - \lambda\|x - \widehat{X}\| - \lambda|z - \widehat{Z}| \right\} \leq \sup_{z\in\mathcal{Z}} \max_{x\in\widehat{\mathcal{X}}} \left\{ \Psi(\widehat{f}^*(x), z) - \lambda\|x - \widehat{X}\| - \lambda|z - \widehat{Z}| \right\}. \tag{4}$$

To this end, we establish the following key lemma.

Lemma 1 **(Shapley Extension)**. *For any function $\widehat{f}\in\widehat{\mathcal{F}}$, define its extension f as*

$$\begin{aligned} f(x) &:= \min_{1\leq k\leq K} \max_{1\leq j\leq K} A_{jk}(x) = \max_{1\leq j\leq K} \min_{1\leq k\leq K} A_{jk}(x), \quad \forall x\in\mathcal{X}, \\ \text{where } A_{jk}(x) &:= \frac{\|x - \widehat{x}_k\|}{\|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|} \widehat{f}(\widehat{x}_j) + \frac{\|x - \widehat{x}_j\|}{\|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|} \widehat{f}(\widehat{x}_k) \text{ when } j \neq k \\ \text{and } A_{kk}(x) &:= \widehat{f}(\widehat{x}_k). \end{aligned} \tag{S}$$

where the saddle point of the matrix $\{A_{jk}(x)\}_{jk}$ is guaranteed to exist. Then f satisfies

(i) *[Extension]* $f(\widehat{x}_k) = \widehat{f}(\widehat{x}_k)$ *for all* $k \in [K]$.

(ii) *[Optimality] For all* $k \in [K]$*, and for every convex function* $\Phi : \mathbb{R} \to \mathbb{R}$*,*

$$\sup_{x \in \mathcal{X}} \left\{ \Phi(f(x)) - \|x - \widehat{x}_k\| \right\} \le \max_{j=1,\ldots,K} \left\{ \Phi(\widehat{f}(\widehat{x}_j)) - \|\widehat{x}_j - \widehat{x}_k\| \right\}. \tag{5}$$

(iii) *[Boundedness]* $\min_{k \in [K]} \widehat{f}(\widehat{x}_k) \le f \le \max_{k \in [K]} \widehat{f}(\widehat{x}_k)$.

(iv) *[Lipschitzness] The Lipschitz norm of* f*, denoted as* $\|f\|_{\mathrm{Lip}}$*, is upper bounded by* $\max_{j \ne k} \dfrac{|\widehat{f}(\widehat{x}_j) - \widehat{f}(\widehat{x}_k)|}{\|\widehat{x}_j - \widehat{x}_k\|}$.

For any x on the line segment connecting $\widehat{x}_j$ and $\widehat{x}_k$, $A_{jk}(x)$ is simply a linear interpolation, elsewhere $A_{jk}(x)$ is a distance-dependent weighted average, which is equivalent to the inverse distance weighting with two points (Shepard 1968). The extension $f(x)$ is given by the saddle point (pure Nash equilibrium) of a matrix $A_{jk}(x)$, whose existence is due to Shapley's theorem (Lemma EC.2 in Appendix A), thus we call it the *Shapley policy*. The first property states that f and $\widehat{f}$ coincide on in-sample data, thus f is indeed an extension. The second property implies (4) and is the key to the proof of Theorem 1. The third and fourth properties indicate that the bound and Lipschitz norm of the extended policy is controlled by those of the in-sample policy.

As an illustration, in the two plots of Figure 2, we plot the Shapley extension when $K = 2, 3$ and when the feature space $\mathcal{X} = \mathbb{R}$. The horizontal axis represents the feature X and the vertical axis represents the policy value (ordering quantity). The points represent an in-sample robust optimal ordering policy. When $K = 2$, the extension $f^*(x) = A_{12}(x)$. On the line segment connecting two points $\widehat{x}_1$ and $\widehat{x}_2$, the interpolation is linear, and is curved elsewhere. As $|x| \to \infty$, the policy converges to a "non-informative" ordering quantity $(\widehat{f}^*(\widehat{x}_1) + \widehat{f}^*(\widehat{x}_2))/2$ which, intuitively, means that the historical observations of features provides little guidance on a faraway new feature value x and thus the policy simply takes the average of the two in-sample policy values. When $K = 3$, for each pair of three historical observations of features $\widehat{x}_1, \widehat{x}_2, \widehat{x}_3$, we plot three curves $A_{12}(x)$ (green), $A_{13}(x)$ (orange), and $A_{23}(x)$ (blue). By solving the minimax saddle point problem, the extended policy $f^*(x)$ would be the middle one among the three curves, as marked with the solid line. Thereby, the saddle point curve $f^*(x)$ is a balanced choice among all pairwise weighted averages. Generally, when x is close to $\widehat{x}_k$, $f^*(x)$ is close to $\widehat{f}^*(\widehat{x}_k)$. When x is away from all historical observations of features, $f^*(x)$ converges to $\frac{1}{2}\left(\min_{k \in [K]} \widehat{f}^*(\widehat{x}_k) + \max_{k \in [K]} \widehat{f}^*(\widehat{x}_k)\right)$. The case of two-dimensional feature space $\mathcal{X} = \mathbb{R}^2$ is similar, as shown in Figure EC.1 in the Appendix A.

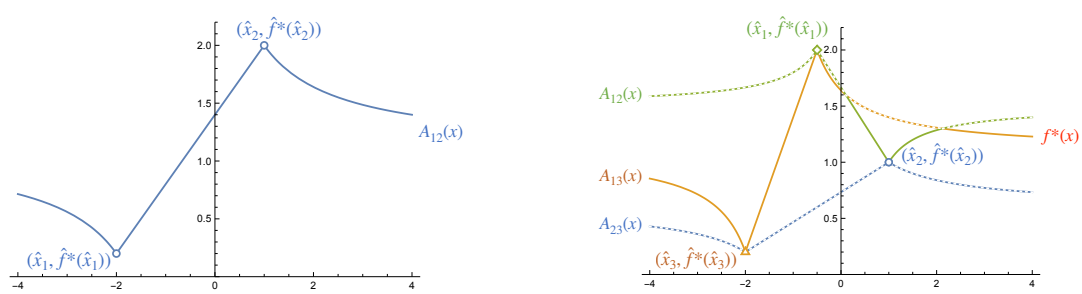


Figure 2 Graph of the Shapley policy $y = f(x)$ when $K = 2, 3$, $x \in \mathbb{R}$.

With Lemma 1, we can prove our main result easily by setting $\Phi(y) = \frac{1}{\lambda} \sup_{z \in \mathcal{Z}} \{\Psi(y, z) - \lambda |z - \widehat{Z}|\}$ (the degenerate case $\lambda = 0$ is trivial) and $\widehat{f} = \widehat{f}^*$, we can prove (4) and thereby, the dual problem (D) is equivalent to the in-sample dual problem ($\widehat{\text{D}}$) in the following sense: $v_D = v_{\widehat{D}}$, and every optimal policy $\widehat{f}^* \in \widehat{\mathcal{F}}$ of ($\widehat{\text{D}}$) can be extended to an optimal policy of (D) via the Shapley extension (S). Since the primal problems (P) and ($\widehat{\text{P}}$) are equivalent to their dual problems (D) and ($\widehat{\text{D}}$), respectively, the theorem is proved. We refer to Appendix A for a complete proof.

3.2. Further Intuition Behind the Shapley Policy

Here we provide some intuition to explain the structural property of the Shapley policy that makes it robust optimal. We would like to emphasize that the discussion below argues from the perspective of the primal problem, whose main purpose is to offer some insights of the structural properties of the problem, possibly at the sacrifice of mathematical rigor. A complete proof will be established from the dual perspective in Appendix A.

3.2.1. Extension based on Slope Minimization As discussed in Section 3.1, the solution to the problem (P) is related to the solution to the problem ($\widehat{\text{P}}$). On the one hand, since we restrict the uncertainty set in ($\widehat{\text{P}}$), the worst-case cost of a policy f in problem (P) is always greater or equal than the worst-case cost of its restriction policy $\widehat{f} = f|_{\widehat{X}}$ in problem ($\widehat{\text{P}}$). On the other hand, if for any in-sample policy $\widehat{f} : \widehat{\mathcal{X}} \to \mathcal{Z}$ we can find an extended policy $f : \mathcal{X} \to \mathcal{Z}$ such that a worst case distribution is guaranteed to be supported on $\widehat{\mathcal{X}}$, then f has the same worst-case expected cost in (P) as $\widehat{f}$ in ($\widehat{\text{P}}$), which makes f an optimal extension, thus the two problems become equivalent. Below we show that the extension defined by the slope minimization problem (L) indeed leads to a worst-case distribution supported on $\widehat{\mathcal{X}}$ and thus is optimal.

Pick any $x \in \mathcal{X} \setminus \widehat{\mathcal{X}}$ and $z \in \mathcal{Z}$. Denote $y_k = \widehat{f}(\widehat{x}_k)$, $y = f(x)$, $d_k = \|\widehat{x}_k - x\|$, and the slope of a secant line connecting x and $\widehat{x}_k$ by $L_k := \frac{y_k - y}{d_k}$, and define $L := \max_{k\in[K]} |L_k|$. We claim that L can be simultaneously achieved by some k^- and k^+, with $L_{k^-} = -L$ and $L_{k^+} = L$. Indeed, if either $-L$ or L is not attained, we can always perturb y in one direction or the other to balance between the two extreme slopes. To prove that the worst-case distribution should not transport any probability mass to x, suppose on the contrary that the probability mass is transported from $(\widehat{x}_j, \widehat{z}_j)$ to (x, z) for some $j \in [K]$. If $y_j \le y$, then there exists some $\delta \in [0,1]$ such that $y = \delta y_j + (1-\delta) y_{k^+}$, because $y_{k^+} = y + L d_{k^+} \ge y$. Now we can propose another transport plan, which instead moves δ fraction of the mass on $(\widehat{x}_j, \widehat{z}_j)$ to $(\widehat{x}_j, z)$ and $1-\delta$ fraction of the mass on $(\widehat{x}_j, \widehat{z}_j)$ to $(\widehat{x}_{k^+}, z)$. Then

- The new transport plan incurs a higher cost due to the convexity of $\Psi(\cdot, z)$: $\delta\Psi(y_j, z) + (1-\delta)\Psi(y_{k^+}, z) \ge \Psi(y, z)$.
- The new plan has a smaller transport cost: distance in z-direction remains the same, while in x-direction the distance is shorter: $\delta\|\widehat{x}_j - \widehat{x}_j\| + (1-\delta)\|\widehat{x}_j - \widehat{x}_{k^+}\| \le 0 + (1-\delta)(d_j + d_{k^+}) \le d_j = \|\widehat{x}_j - x\|$. Here in the first inequality we have used the triangle inequality $\|\widehat{x}_j - \widehat{x}_{k^+}\| \le \|\widehat{x}_j - x\| + \|x - \widehat{x}_{k^+}\|$, and the second inequality is because

$$y = \delta y_j + (1-\delta) y_{k^+} = \delta(y + d_j L_j) + (1-\delta)(y + d_{k^+} L) \ge y - \delta d_j L + (1-\delta) d_{k_+} L \Rightarrow (1-\delta) d_{k_+} \le \delta d_j.$$

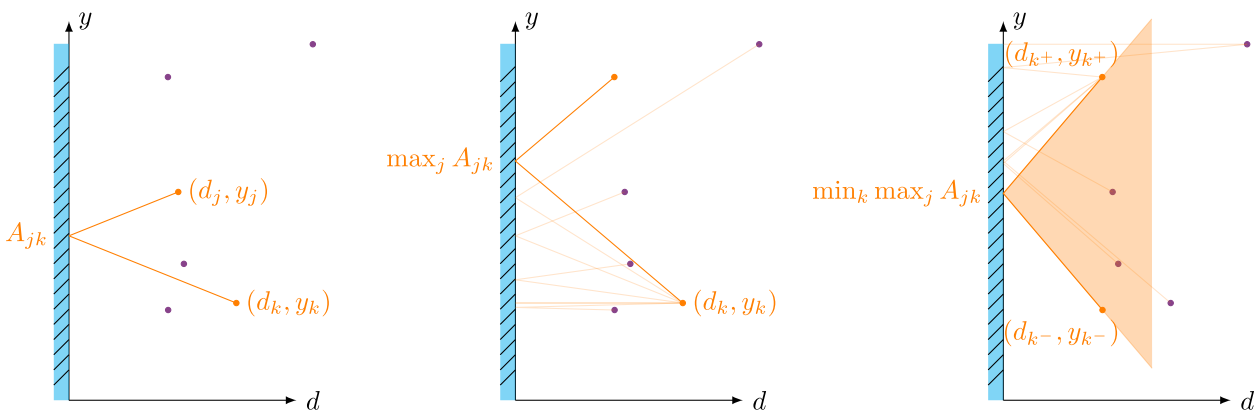
Hence, moving probability mass to $\widehat{\mathcal{X}}$ always lead to a worse distribution than moving to $\mathcal{X}$, thus we prove the claim and the policy defined by (L) is optimal.

3.2.2. Shapley Policy as the Solution to Slope Minimization Next, we show geometrically in Figure 3 that the Shapley extension (S) is the closed-form solution to (L). To visualize $A_{jk}(x)$, we plot the d–y plane on which a point (d_k, y_k) means $\widehat{x}_k$ is of distance d_k away from x and is assigned a policy value y_k. Imagine a mirror at $d = 0$ facing right. It is not hard to see that $A_{jk}(x)$ is the reflection point of the point j in the mirror from the point k's viewpoint (Figure 3(a)). Thereby $\max_j A_{jk}(x)$ corresponds to the highest reflection points among all points in the mirror of point k'th horizon, as shown in Figure 3(b). Minimizing over k gives the Shapley saddle point $f(x) = \min_k \max_j A_{jk}(x)$ (Figure 3(c)). Geometrically, since $L_{k_+} = -L_{k_-}$, the shadow region in 3(c) is a symmetric cone with the smallest opening that covers all the points $\{(d_k, y_k)\}_{k\in[K]}$. Thus, it is apparent that the minimax theorem holds by symmetry, and the vertex of such smallest opening is precisely determined by the Lipschitz-minimization problem (L).

Furthermore, by introducing an auxiliary variable L denoting the inner maximum of (L), (L) can be solved by the following linear program

$$\begin{aligned} \min_{L\geq 0, y\in\mathbb{R}} \quad & L \\ \text{subject to} \quad & y - Ld_k \leq y_k \leq y + Ld_k, \qquad \text{for all } k \in [K] \end{aligned} \tag{6}$$

with y corresponding to the vertex and L corresponding to the slope of the symmetric cone. Note that the optimal L never exceeds the Lipschitz norm of the in-sample policy. As we shall see in the following subsection, penalizing the Lipschitz norm will be another equivalent formulation for the optimal policy. The above linear program also provides a numerical scheme to locate the saddle point. Naïve computation of the saddle point of a $K \times K$ matrix has time complexity $O(K^2)$, but using linear program allows much faster computation for large K empirically, as shown in Section 5.



(a) $A_{jk}(x)$ (b) $\max_j A_{jk}(x)$ (c) $\min_k \max_j A_{jk}(x)$

Figure 3 Visualization of the Shapley saddle point

Before closing this subsection, we remark that the analysis above mainly relies on (a) the convexity of the newsvendor cost; (b) the triangle inequality of the transport cost $\|\cdot\|$; and (c) one-dimensional decision for which the interpolation and extrapolation are well-defined. Consequently, our results Theorem 1 remain to hold for any cost function that is convex in the one-dimensional decision variable, but only applies to 1-Wasserstein distance for which the triangle inequality of the transport cost function applies. Extensions to other cases appear to be nontrivial, if possible at all, and are left for future investigation.

4. Discussions

In this section, we provide additional properties of our distributionally robust formulation (P) or its dual (D). Unlike Section 3, results in this section relies on the specific form of the newsvendor cost beyond convexity. In Section 4.1, we show that the problems (D) ($\widehat{\text{D}}$) can be equivalently interpreted as the Lipschitz regularization on the policy. Based on this observation, we develop a finite-dimensional linear program to compute the optimal in-sample robust policy in Section 4.2, and derive the generalization bound of the Shapley policy in Section 4.3. Finally, in Section 4.4, we discuss the structure of the optimal robust policy.

4.1. Interpretation as Lipschitz Regularization

In this subsection, we establish an equivalence between our robust formulation and Lipschitz regularization, as already hinted in Section 3.2.

Consider the following Lipschitz regularization problem defined as

$$v_R := \min_{f \in \mathcal{F}} \left\{ (b \vee h)(1 \vee \|f\|_{\mathrm{Lip}})\rho + \mathbb{E}_{(\widehat{X},\widehat{Z})\sim\widehat{\mathbb{P}}}[\Psi(f(\widehat{X}),\widehat{Z})] \right\}, \tag{R}$$

where we denote by $a_1 \vee a_2$ the maximum between a_1 and a_2, and by $\|\cdot\|_{\mathrm{Lip}}$ the Lipschitz norm of a function (which is infinite for non-Lipschitz functions). Problem (R) balances between the variation of a policy f (reflected by the term $1 \vee \|f\|_{\mathrm{Lip}}$) and its expected in-sample cost. If we set $\rho = 0$, (R) would degenerate to the non-robust empirical risk minimization. In this case, the optimal policy f^* is defined only on the in-sample data, which equals to the critical fractile of the empirical conditional distribution; and can take any value outside $\widehat{\mathcal{X}}$. If we prohibit perturbing any data in z-direction, then the lower cut off 1 of the Lipschitz norm $\|f\|_{\mathrm{Lip}}$ in the first term of (R) vanishes. In this case, when $\rho \to \infty$, the Lipschitz penalty term forces the optimal policy f^* to be a constant function, thereby (R) reduces to the classical newsvendor problem without feature information

$$\min_{y\in\mathbb{R}} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}}}[\Psi(y,\widehat{Z})],$$

and the optimal policy equals the unconditional critical fractile of $\widehat{\mathbb{P}}_{\widehat{Z}}$. Let us also define the in-sample Lipschitz regularization problem

$$v_{\hat{R}} := \min_{\widehat{f} \in \widehat{\mathcal{F}}} \left\{ (b \vee h)(1 \vee \|\widehat{f}\|_{\mathrm{Lip}})\rho + \mathbb{E}_{(\widehat{X},\widehat{Z})\sim\widehat{\mathbb{P}}}[\Psi(\widehat{f}(\widehat{X}),\widehat{Z})] \right\}. \tag{$\widehat{\mathrm{R}}$}$$

The following result enables us to further reformulate the problem as Lipschitz regularization. The proof is provided in Appendix B.

Proposition 1. *Problems* ($\widehat{\mathrm{D}}$) *and* ($\widehat{\mathrm{R}}$) *are equivalent. Moreover, if* $\widehat{f}^*$ *is an optimal solution to* ($\widehat{\mathrm{R}}$)*, then its Shapley extension defined by* (S) *is an optimal solution to* (R).

In the literature, it is known that 1-Wasserstein distributionally robust optimization is upper bounded by Lipschitz regularization (Mohajerin Esfahani and Kuhn 2018, Shafieezadeh-Abadeh et al. 2019, Gao et al. 2022), and the two problems are equivalent under certain assumptions. That said, Proposition 1 is arguably surprising because our considered problem does not satisfy the assumptions imposed in the references above that ensure the equivalence. The key observation in the proof (Lemma EC.3 in Appendix B) is that there exists a robust optimal in-sample policy $\widehat{f}^*$ with sufficiently small Lipschitz norm, which gives a direct restriction on the range of the dual multiplier λ and transforms ($\widehat{\mathrm{D}}$) to ($\widehat{\mathrm{R}}$).

Thus far, combining all results in Sections 3.1 and 4.1, we have shown that problems ($\widehat{\mathrm{P}}$), ($\widehat{\mathrm{D}}$) and ($\widehat{\mathrm{R}}$) share an in-sample optimal robust policy $\widehat{f}^* \in \widehat{\mathcal{F}}$, which can be extended to an optimal robust policy $f^* \in \mathcal{F}$ for problems (P), (D) and (R).

4.2. Linear Programming Reformulation for In-sample Robust Problem

Based on the Lipschitz regularization reformulation, we derive a linear programming reformulation for the in-sample problem ($\widehat{\mathrm{P}}$). From Proposition 1, by introducing auxiliary variables, we directly conclude the following.

Proposition 2. *Problem* ($\widehat{\mathsf{R}}$) *is equivalent to*

$$\begin{aligned} \min_{y\in\mathbb{R}^K,\, L\geq 1,\, \psi_k\in\mathbb{R}^{n_k},\, 1\leq k\leq K} \quad & (b\vee h)\rho L + \frac{1}{n}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\psi_{ki} \\ \text{s.t.} \quad & |y_j - y_k| \leq L\|\widehat{x}_j - \widehat{x}_k\|,\ \forall 1\leq j,k\leq K, \\ & h(y_k - \widehat{z}_{ki}) \leq \psi_{ki},\ b(\widehat{z}_{ki} - y_k) \leq \psi_{ki},\ \forall 1\leq i\leq n_k, 1\leq k\leq K. \end{aligned}$$

This linear program has $N+K+1$ variables with K^2+2N+1 constraints. The variable L stands for $1\vee\|\widehat{f}\|_{\mathrm{Lip}}$; and ψ_{ki} represents the cost of ordering quantity y_k when the demand is $\widehat{z}_{ki}$ and the feature value is $\widehat{x}_k$. In practice, we may impose different norm scaling parameter $\beta>0$ in $\|(x,z)\| = \|x\| + \frac{1}{\beta}|z|$ that balances between the uncertainty in the feature information and in the demand. In this new setting, $v_{\hat{R}}$ becomes $v_{\hat{R}} = (b\vee h)(\beta\vee\|\widehat{f}\|_{\mathrm{Lip}})\rho + \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_{\widehat{f}}]$, and in the linear programming $L\geq 1$ becomes $L\geq\beta$. In Appendix C, we show in Figure EC.2 how the performance of the algorithm is affected by the choice of parameters ρ and β.

4.3. Generalization Error Bound

Another implication of the Lipschitz regularization reformulation is on the statistical property of the Shapley policy.

For simplicity, in this subsection we consider the norm parameter $\beta=0$, so the product norm in the joint space $\mathcal{X}\times\mathcal{Z}$ is defined as $\|(x,z)\| = \|x\| + \infty\mathbf{1}_{\{z\neq 0\}}$. From the proof of Proposition 1 and following the discussion of norm parameter after Proposition 2, we can see that

$$\|\Psi_{f^*}\|_{\mathrm{Lip}} = (b\vee h)\|f^*\|_{\mathrm{Lip}} = \lambda^*,$$

where λ^* is the optimal dual variable in ($\widehat{\mathsf{D}}$). Hence, using the Lipschitz composition property of the Radamacher complexity, the expected generalization gap of Ψ_{f^*} is dominated by λ^* times the expected Radamacher complexity of a 1-Lipschitz ball in $\mathcal{X}$, $\mathfrak{R}_n(\mathrm{Lip}_1(\mathcal{X})) = O(n^{-\frac{1}{d}})$, where d is the dimension of $\mathcal{X}$ (Luxburg and Bousquet 2004, Theorems 15 and 18). We have the following result.

Proposition 3. *Assume the demand is upper bounded by $\bar{D}>0$. Then the expected generalization gap of the optimal robust policy f^* is upper bounded by*

$$\mathbb{E}_{\otimes}\left[\mathbb{E}_{\mathbb{P}_{\mathrm{true}}}[\Psi_{f^*}] - \mathbb{E}_{\mathbb{P}_n}[\Psi_{f^*}]\right] \leq \frac{2(b\vee h)\bar{D}}{\rho}\mathfrak{R}_n(\mathrm{Lip}_1(\mathcal{X})),$$

where $\mathbb{E}_{\otimes}$ denotes expectation over the random sampling distribution $\mathbb{P}_n$.

Unlike the parametric results in Gao (2022), An and Gao (2021), this bound is dimension-dependent, which is reasonable since it essentially considers a non-parametric statistical setting. We note that existing performance guarantees Mohajerin Esfahani and Kuhn (2018) on Wasserstein DRO is not directly applicable to our original problem, because it constrains the loss functions to a certain class, such as those with sublinear growth. However, in our case the loss function Ψ_f is a priori unconstrained as there is no assumption on $f\in\mathcal{F}$. It is our result on the Lipschitz regularization that enables us to derive an upper bound on the generalization bound.

4.4. Structure of the Optimal Robust Policy

In this subsection, we discuss the structure of the optimal robust policy, which provides further interpretation of the in-sample optimal robust policy. The proofs are provided in Appendix D.

Define the *empirical conditional critical fractiles* for $\widehat{x} \in \widehat{\mathcal{X}}$ as

$$\overline{q}(\widehat{x}) := \max\left\{z \in \mathcal{Z} : \widehat{\mathbb{P}}\{Z < z | X = \widehat{x}\} \le \frac{b}{b+h}\right\}, \tag{7}$$

$$\underline{q}(\widehat{x}) := \min\left\{z \in \mathcal{Z} : \widehat{\mathbb{P}}\{Z \le z | X = \widehat{x}\} \ge \frac{b}{b+h}\right\}. \tag{8}$$

It is easy to see that $\underline{q}(\widehat{x}) \le \overline{q}(\widehat{x})$. We also define the subsets of historical observations of features

$$\widehat{\mathcal{X}}_< := \left\{\widehat{x} \in \widehat{X} : \ \overline{q}(\widehat{x}) < \widehat{f}^*(\widehat{x})\right\}, \quad \widehat{\mathcal{X}}_> := \left\{\widehat{x} \in \widehat{X} : \ \underline{q}(\widehat{x}) > \widehat{f}^*(\widehat{x})\right\}.$$

The following result gives a finer description of the in-sample optimal robust policy $\widehat{f}^*$.

Proposition 4.

(I) *If $\underline{q}(\widehat{x}_j) - \overline{q}(\widehat{x}_k) \le \|\widehat{x}_j - \widehat{x}_k\|$ for all $1 \le j, k \le K$, then $\widehat{f}^*$ is 1-Lipschitz and an empirical conditional critical fractile, i.e. $\underline{q} \le \widehat{f}^* \le \overline{q}$. In this case, $\widehat{f}^*$ is optimal to ($\widehat{\mathrm{R}}$) for any $\rho \ge 0$.*

(II) *Otherwise, $\|\widehat{f}^*\|_{\mathrm{Lip}} = L \ge 1$. For every $\widehat{x}_k \in \widehat{\mathcal{X}}_>$, there exists $\widehat{x}_j \in \widehat{\mathcal{X}} \setminus \widehat{\mathcal{X}}_>$, such that $\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j) = L\|\widehat{x}_k - \widehat{x}_j\|$. Similarly, for every $\widehat{x}_k \in \widehat{\mathcal{X}}_<$ there exists $\widehat{x}_j \in \widehat{\mathcal{X}} \setminus \widehat{\mathcal{X}}_<$ such that $\widehat{f}^*(\widehat{x}_j) - \widehat{f}^*(\widehat{x}_k) = L\|\widehat{x}_k - \widehat{x}_j\|$.*

Proposition 4 separates two cases. First, if the empirical conditional critical fractile is already 1-Lipschitz, it must be an optimal policy since it minimizes both terms in ($\widehat{\mathrm{R}}$). Otherwise, if the variation of the empirical conditional critical fractile is too large, i.e., a large value of $|\widehat{f}^*(\widehat{x}_j) - \widehat{f}^*(\widehat{x}_k)|/\|\widehat{x}_j - \widehat{x}_k\|$ for some $j \ne k$, then to reduce the variation of the policy, we would order more than the empirical critical fractile $\overline{q}(\widehat{x}_k)$ at the cost of holding more, resulting in a set $\widehat{\mathcal{X}}_<$; or would order less than the empirical critical fractile $\underline{q}(\widehat{x}_k)$ at the cost of back ordering more, resulting in a set $\widehat{\mathcal{X}}_>$.

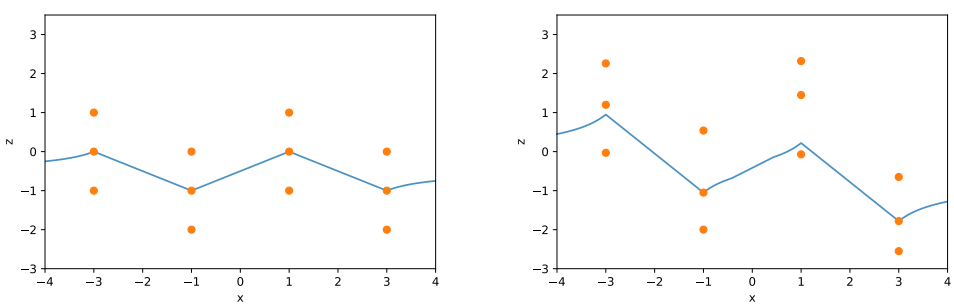


Figure 4 Optimal policy $\widehat{f}^*$ (blue curves) given the empirical distribution of feature and demand (orange dots). If the empirical conditional critical fractile q is 1-Lipschitz, then $\widehat{f}^* = q$ (left). Otherwise, $\widehat{f}^*$ regularizes q by reducing its Lipschitz norm (right).

The discussion above is illustrated in Figure 4. We have $b = 3$, $h = 2$, $\rho = 10$, $\mathcal{X} = \mathbb{R}$, $\widehat{\mathcal{X}} = \{-3, -1, 1, 3\}$, and $\widehat{\mathbb{P}}_n$ is supported on $n = 12$ points (as indicated by the orange dots in x–z plot), where each $\widehat{x} \in \widehat{\mathcal{X}}$ are associated to three demand realizations with the middle level corresponding to the empirical conditional critical fractile $q = \overline{q} = \underline{q}$. In the left example, q is 1-Lipschitz, hence the optimal policy $\widehat{f}^*$ (represented by the blue curve) passes through all empirical conditional critical fractiles; on the right, $\widehat{f}^*$ regularizes q by ordering less than q on $\widehat{x} = -3, 1$ so as to reduce the variation of the policy.

Given the structure of the optimal policy, in Proposition EC.1 and Figure EC.3 in Appendix D, we investigate the worst-case distribution $\mathbb{P}^*$, which sheds light on the (non-)conservativeness of our formulation.

5. Numerical Experiments

In this section, using synthetic data and real data in Sections 5.1 and 5.2 respectively, we compare the empirical performance of our proposed Shapley policy against several benchmarks, including: empirical risk minimization using affine policy with ℓ^1 and ℓ^2 regularization (ERM2 (ℓ^1/ℓ^2)) and kernel-weights optimization (KO) in Ban and Rudin (2019); conditional stochastic optimization using random forests (RandForest) and k-nearest neighbors (kNN) in Bertsimas and Kallus (2020); and stochastic optimization forests with different splitting criteria (StochOptForest (apx-soln/apx-risk)) in Kallus and Mao (2022).

The hyperparmeters for each methods are shown in Table EC.1 in Appendix E. Particularly, in our Shapley, the hyperparameters are the radius ρ and the norm scaling parameter β introduced after Proposition 2. For models with hyperparameters, we use 5-fold cross-validation to tune their values: first, we randomly shuffle the n training samples and partition them into five groups; next, for each group, we use the remaining four groups of samples for calibrating the model and use this group for validation; finally, we pick the hyperparameters with the least average validation cost averaged over the five groups. Once the hyperparameters are tuned, we use all n samples together to train the models again. All experiments are performed in Ubuntu 18.04 using Python 3.6.9 with a convex optimization solver MOSEK 9.2.35 and CVXPY 1.1.7 interface, on a Dell Precision 5820 Tower Workstation with Intel® Xeon® W-2125 CPU (32 cores) and 32GB RAM (DDR4 2666MHz).

5.1. Synthetic Data

Consider the following nonlinear model with heteroscedastic noises, based on the experiment setup in Zhu et al. (2012) that studies high-dimensional quantile regression

$$Z = \sin\left(2(X^\top \beta_0)\right) + 2\exp\left(-16(X^\top \beta_0)^2\right) + \exp(X^\top \beta_0)\varepsilon,$$

where the feature vector $X \in \mathbb{R}^d$ is a multivariate Gaussian random variable with mean zero and covariance matrix $(\sigma_{ij})_{d\times d}$ with $\sigma_{ij} = (\frac{199}{200})^{|i-j|}$, the coefficient parameter $\beta_0 = c(200, -200, 199, -199, \ldots, 1, -1, 0, 0, \ldots, 0)^\top$ is a d-dimensional vector with $d = 5000$, where $c = 25/\sqrt{\sum_{k=1}^{200} 2k^2}$ is chosen such that $\|\beta\|_2 = 25$ and $\beta^\top X$ is normally distributed with mean zero and roughly standard deviation 1; and ε is standard Gaussian. To study the impact of the extremity of quantile, we vary $h \in \{0.2, 0.5, 1\}$ while fixing $b = 1$. To study the impact of the sample size n, we vary the number of data points $n \in \{10, 100, 1000\}$. The testing data size is fixed as 10000. We use Euclidean norm for all distance-based approaches, namely, Shapley, KO, kNN.

We run 20 repeated experiments and draw the boxplots in Figure 5, which are generated by the relative differences of out-of-sample costs between our method and each of the benchmarks using the same training and testing data set; a positive number indicates that Shapley leads to an $x\%$ reduction on contrast. We have the following observations.

(I) In general, Shapley has superior performance under the regimes with small sample size ($n = 10, 100$) and extreme quantiles ($h = 0.2, 0.5$). This demonstrates the advantages of Shapley resulting from the distributionally robust formulation (P) that usually works well in small-sample regimes. As the sample size grows larger, the absolute differences among all methods are getting smaller, but our Shapley policy still maintains a competitive median performance.

(II) For benchmark methods, kNN is outperformed by Shapley policy in most situations (8 out of 9 boxplots in terms of the median performance). The affine policy ERM2 (ℓ^1) and KO have a clear advantage over our Shapley when $n = 1000, h = 1$. But for small sample size ($n = 10$) they are outperformed by our Shapley. Forest-based methods (RandForest, StochOptForest (apx-soln)) have a relatively good performance when the sample size is small. We note, however, that the performance plotted in Figure 5 may not show the full strength of the forest-based methods since

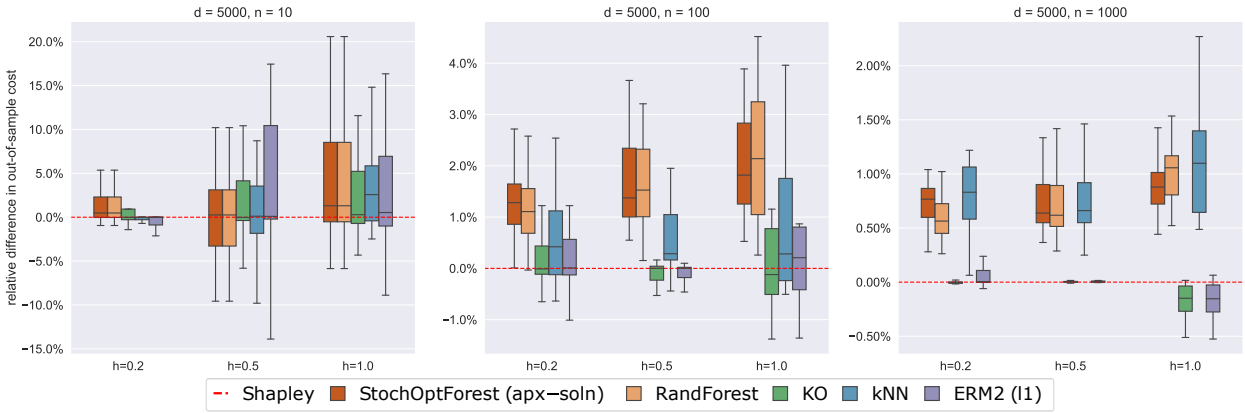


Figure 5 Boxplots of the relative differences in the out-of-sample performance between `Shapley` and other benchmarks

we run a reduced version of the algorithm due to time limit; see (II) below. Moreover, even under this case, the running time of `StochOptForest (apx-risk)` still exceeds the time threshold so we are unable to show its performances under this experiment, but will do so later using a real data set.

We report the average computational time of these methods in Table 1. The training time is defined as the time spent on training the model using n samples with fixed hyperparameter values and constructing the policy function, and the testing time is defined as the time spent on finding the ordering quantity and evaluating the average cost using the full testing set. The computational time is averaged over all 20 experiments and all choices of parameter h (as it appears that the value of h does not have an evident impact on the running time). We have the following observations.

Model	Average Training Time			Average Testing Time		
	$n=10$	$n=100$	$n=1000$	$n=10$	$n=100$	$n=1000$
`Shapley`	0.0089	0.2591	20.9073	0.8290	2.5210	28.8658
k`NN`	0.0003	0.0010	0.0092	98.4019	134.6318	496.1717
`KO`	0.0003	0.0010	0.0101	100.7559	136.6974	500.7450
`ERM2` (ℓ^1)	1.6626	3.0285	23.7068	8.6050	8.6662	8.5541
`RandForest`*	0.0877	21.1517	479.5995	19.8414	25.0150	28.6125
`StochOptForest (apx-soln)`*	0.0880	24.1451	536.7108	19.7084	24.9899	28.4322

Table 1 Average computational time (in seconds) per problem instance
*The algorithm runs with a reduced version (see (II) below)

(I) For the average training time, `KO` and k`NN` takes the least amount of time since they simply involve data sorting. `ERM2` (ℓ^1) involves solving linear programming (ℓ^1), and the training time increases mildly in the training sample size. `Shapley` also involves solving the linear programming reformulation of the in-sample robust problem (Proposition 2) with $O(n)$ variables and $O(n^2)$ constraints when the number of distinct historical features $K \approx n$ (which holds in our experiments since almost every sample has a distinct feature); empirically the training time increases linearly in sample size. Forest-based methods (`RandForest`, `StochOptForest (apx-soln)`) take the longest time to train, and often longer in magnitudes.

(II) For the average testing time, affine policy `ERM2` (ℓ^1) takes the least time for testing when the testing size is large ($n=1000$) and is independent of the training sample size by construction. The average testing time of `KO` and k`NN` roughly grows linearly in the training sample size n (subtracting some overhead time) as their theoretical bounds are $O(n)$. We find empirically that the testing

time of Shapley grows sub-linearly in the sample size using our equivalent formulation (6), which is faster than computing the matrix saddle point naïvely. In this experiment forest-based methods (RandForest, StochOptForest (apx-soln)) are run with a reduced version, which reduces the number of the trees by randomly selecting one-tenth of the features for the candidate splits, as we set a time threshold of one hour for each instance. Comparisons with the full version of the algorithms will be conducted in Section 5.2.

5.2. Real Data

We test the performance using a real-world dataset considered in Oroojlooyjadid et al. (2020) modified from Pentaho (2008), which consists of historical demands of baskets from a retailer in two years. Each demand is associated with three categorical variables that are observable before the ordering decision is made: department id (1 to 24), month of year (1 to 12), and day of week (1 to 7). In the affine policy model (ERM2 (ℓ^1/ℓ^2)) and forest-based model (RandForest, StochOptForest (apx-soln/apx-risk)), these feature variables are converted into 43 dummy variables, so the dimension is $d = 40$ (43 dummy variables minus 3 reference categories). For distance-based methods (Shapley, kNN, KO), the metric on $\mathcal{X}$ is defined as the following. Let $x_k = (i_k, m_k, d_k) \in \mathcal{X}$, $k = 1, 2$, where $i_k = 1, \ldots, 24$, $m_k = 1, \ldots, 12$, $d_k = 1, \ldots, 7$ represent department id, month and weekday respectively, and we define the distance on $\mathcal{X}$ as

$$\|x_1 - x_2\| = \left(\|i_1 - i_2\|_{\text{cat}}^2 + \|m_1 - m_2\|_{\text{mod}_{12}}^2 + \|d_1 - d_2\|_{\text{mod}_7}^2 \right)^{\frac{1}{2}},$$

where we choose the discrete metric $\|\cdot\|_{\text{cat}}$ by $\|i_1 - i_2\|_{\text{cat}} = 1$ if $i_1 \neq i_2$ and 0 otherwise, since there is no apparent relationship between different departments; $\|\cdot\|_{\text{mod}_{12}}$ and $\|\cdot\|_{\text{mod}_7}$ are metrics on the ring of integers modulo 12 and 7, where $\|n\|_{\text{mod}_q} = \frac{1}{q} \min\left\{|r| : r \equiv n \bmod q\right\}$, which measures the similarity among months/days based on their closeness.

The raw data set contains 9877 data for training and 3293 data for testing. To generate the training data sets for our purpose, in each of the repeated experiments, we randomly draw n distinct samples from the raw training data set to train the model. The out-of-sample cost is defined as the average newsvendor cost over the full set of the testing data. To investigate in different sample size regimes, we consider $n \in \{20, 40, 100\}$, representing the cases of $n < d$, $n = d$, and $n > d$. We vary the unit holding cost $h \in \{0.2, 0.5, 1\}$, while fixing the unit backorder cost $b = 1$.

The average out-of-sample costs and their 95% confidence intervals are shown in Table 2. We observe that our Shapley policy has a superior performance comparing with all other methods in terms of the average out-of-sample cost; the advantage is most apparent when the h/b-ratio is extreme, namely, $h/b = 0.2$.

Model		Shapley	KO	kNN	StochOptForest (apx-risk)	StochOptForest (apx-soln)	RandForest	ERM2 (ℓ^1)	ERM2 (ℓ^2)
$h = 0.2$	$n = 20$	**24.85 ± 1.12**	26.54 ± 0.90	28.71 ± 1.39	27.30 ± 0.66	27.31 ± 0.66	27.31 ± 0.66	30.60 ± 1.68	28.95 ± 1.51
	$n = 40$	**23.38 ± 0.98**	26.30 ± 0.53	27.75 ± 0.94	29.40 ± 1.76	29.29 ± 1.67	29.05 ± 1.74	27.68 ± 1.31	27.72 ± 1.05
	$n = 100$	**20.47 ± 0.55**	25.79 ± 0.26	22.38 ± 0.70	27.96 ± 0.84	27.90 ± 0.86	28.02 ± 0.85	23.40 ± 1.04	26.68 ± 0.31
$h = 0.5$	$n = 20$	**37.70 ± 1.22**	38.74 ± 0.64	39.97 ± 0.89	39.31 ± 0.50	39.29 ± 0.49	39.28 ± 0.49	42.33 ± 2.18	40.68 ± 1.49
	$n = 40$	**34.93 ± 1.25**	38.80 ± 0.71	38.95 ± 1.04	39.64 ± 1.03	39.65 ± 1.03	39.83 ± 1.04	40.54 ± 1.81	38.97 ± 0.26
	$n = 100$	**30.41 ± 0.43**	37.53 ± 0.35	31.34 ± 0.76	39.29 ± 0.42	39.28 ± 0.42	39.33 ± 0.46	31.51 ± 1.45	38.66 ± 0.25
$h = 1$	$n = 20$	**44.14 ± 0.80**	45.82 ± 0.76	47.03 ± 0.54	46.83 ± 0.53	46.81 ± 0.51	46.83 ± 0.53	49.57 ± 3.40	47.50 ± 1.35
	$n = 40$	**43.99 ± 0.75**	45.83 ± 0.36	45.75 ± 0.63	47.11 ± 0.88	46.97 ± 0.78	47.12 ± 0.91	46.27 ± 1.29	46.52 ± 0.45
	$n = 100$	40.28 ± 0.73	45.34 ± 0.21	**39.17 ± 0.87**	46.60 ± 0.33	46.57 ± 0.33	46.58 ± 0.32	40.65 ± 1.31	46.33 ± 0.48

Table 2 Average out-of-sample costs with 95% confidence intervals

To better compare the instance-wise performance, we also draw 3 boxplots in Figure 6 under different choices of h and n, which again shows the differences of out-of-sample costs between our Shapley to other benchmarks using same training/testing data set. We observe the following.

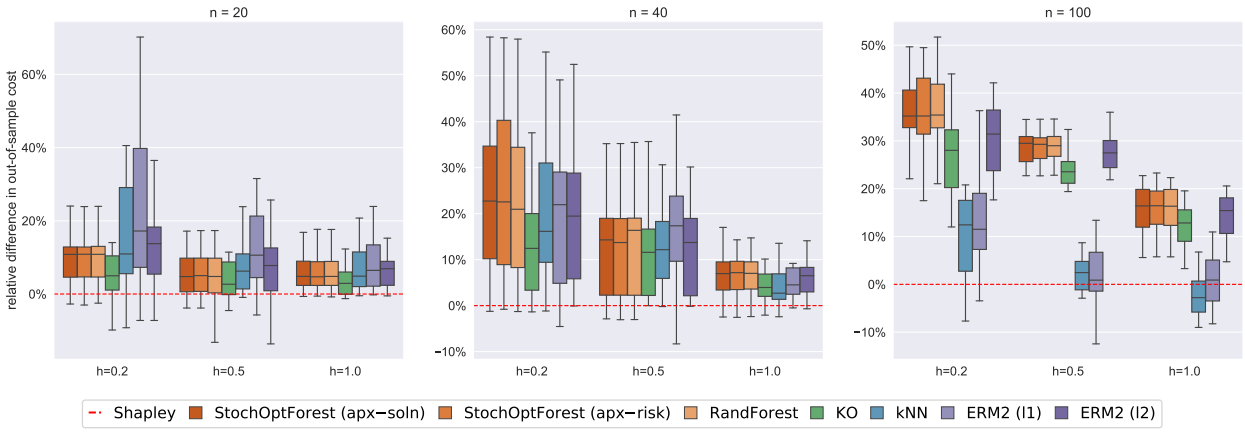


Figure 6 Boxplots of the differences in the out-of-sample performance between `Shapley` and other benchmarks

(I) Affine policies with regularization do not perform very well, and are outperformed by `Shapley` in terms of the median out-of-sample cost. This demonstrates the advantage of using a richer policy class beyond affine.

(II) Compared with non-parametric distance-based approaches `KO` and k`NN`, in terms of the median out-of-sample costs, `Shapley` outperforms both of them in 8 out of 9 cases. Especially when the h/b-ratio is 0.2, the difference is most obvious, and this holds for all sample sizes n. This demonstrates the advantages of `Shapley` for extreme critical fractiles.

(III) Forest-based methods (`RandForest`, `StochOptForest (apx-soln/apx-risk)`) are outperformed by `Shapley` in terms of the median of out-of-sample cost in all 9 cases. Note that unlike the synthetic experiments, here we use the full version of these methods.

We also report the average computational time in this experiment. The results shows in Table 3. The observations are mostly consistent with observations in the synthetic experiment. In particular, the average testing time of `KO`, k`NN` and `Shapley` are comparable, but the full version (without reducing the number of splits) of forest-based methods take a longer time in testing.

Model	Average Training Time			Average Testing Time		
	$n=20$	$n=40$	$n=100$	$n=20$	$n=40$	$n=100$
`KO`	0.0002	0.0001	0.0002	5.2130	5.2173	5.2485
k`NN`	0.0002	0.0001	0.0002	5.2287	5.2465	5.3025
`ERM2` (ℓ^1)	0.0123	0.0128	0.0145	0.4261	0.4270	0.4339
`ERM2` (ℓ^2)	0.0169	0.0181	0.0214	0.4262	0.4268	0.4288
`Shapley`	0.0456	0.1006	0.3155	5.3749	5.4800	5.8135
`RandForest`	0.3526	1.7674	6.5025	65.9638	68.5409	73.9359
`StochOptForest (apx-soln)`	0.3533	1.7964	6.6897	65.9666	68.6409	73.8397
`StochOptForest (apx-risk)`	0.3520	1.7462	6.4172	65.9379	68.6739	73.7745

Table 3 Average computational time (in seconds) per problem instance

6. Concluding Remarks

In this paper, we have developed an efficient approach to finding a robust optimal end-to-end policy for the feature-based newsvendor, which balances between the variation of the ordering quantity with respect to features and the expected cost. Our proposed Shapley extension provides a novel

family of policies for adjustable robust optimization with provably zero optimality gap, and can be easily extended to contextual decision-making problems with any convex cost beyond the newsvendor cost. For the future work, it would be interesting to extend it to other adjustable robust optimization problems.

References

An Y, Gao R (2021) Generalization bounds for (Wasserstein) robust optimization. *Advances in Neural Information Processing Systems* 34.

Ban GY, Gallien J, Mersereau AJ (2019) Dynamic procurement of new products with covariate information: The residual tree method. *Manufacturing & Service Operations Management* 21(4):798–815.

Ban GY, Rudin C (2019) The big data newsvendor: Practical insights from machine learning. *Operations Research* 67(1):90–108.

Bayraksan G, Love DK (2015) Data-driven stochastic programming using phi-divergences. *The Operations Research Revolution*, 1–19 (INFORMS).

Ben-Tal A, Den Hertog D, De Waegenaere A, Melenberg B, Rennen G (2013) Robust solutions of optimization problems affected by uncertain probabilities. *Management Science* 59(2):341–357.

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) Robust optimization. *Robust optimization* (Princeton university press).

Bertsimas D, Georghiou A (2015) Design of near optimal decision rules in multistage adaptive mixed-integer optimization. *Operations Research* 63(3):610–627.

Bertsimas D, Goyal V (2012) On the power and limitations of affine policies in two-stage adaptive optimization. *Mathematical programming* 134(2):491–531.

Bertsimas D, Iancu DA, Parrilo PA (2010) Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research* 35(2):363–394.

Bertsimas D, Iancu DA, Parrilo PA (2011) A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Transactions on Automatic Control* 56(12):2809–2824.

Bertsimas D, Kallus N (2020) From predictive to prescriptive analytics. *Management Science* 66(3):1025–1044.

Bertsimas D, Koduri N (2022) Data-driven optimization: A reproducing kernel Hilbert space approach. *Operations Research* 70(1):454–471.

Bertsimas D, McCord C, Sturt B (2022) Dynamic optimization with side information. *European Journal of Operational Research (Forthcoming)* .

Blanchet J, Murthy K (2019) Quantifying distributional model risk via optimal transport. *Mathematics of Operations Research* 44(2):565–600.

Carrizosa E, Olivares-Nadal AV, Ramírez-Cobo P (2016) Robust newsvendor problem with autoregressive demand. *Computers & Operations Research* 68:123–133.

Chen X, Sim M, Sun P, Zhang J (2008) A linear decision-based approximation approach to stochastic programming. *Operations Research* 56(2):344–357.

Chen Z, Xie W (2021) Regret in the newsvendor model with demand and yield randomness. *Production and Operations Management* 30(11):4176–4197.

Das B, Dhara A, Natarajan K (2021) On the heavy-tail behavior of the distributionally robust newsvendor. *Operations Research* 69(4):1077–1099.

El Balghiti O, Elmachtoub AN, Grigas P, Tewari A (2019) Generalization bounds in the predict-then-optimize framework. *Advances in Neural Information Processing Systems* 32:14412–14421.

El Housni O, Goyal V (2021) On the optimality of affine policies for budgeted uncertainty sets. *Mathematics of Operations Research* 46(2):674–711.

Elmachtoub AN, Grigas P (2022) Smart "predict, then optimize". *Management Science* 68(1):9–26.

Esteban-Pérez A, Morales JM (2021) Distributionally robust stochastic programs with side information based on trimmings. *Mathematical Programming* 1–37.

Fu M, Li X, Zhang L (2021) Data-driven feature-based newsvendor: A distributionally robust approach. *Available at SSRN 3885663* .

Gallego G, Moon I (1993) The distribution free newsboy problem: review and extensions. *Journal of the Operational Research Society* 44(8):825–834.

Gallego G, Ryan JK, Simchi-Levi D (2001) Minimax analysis for finite-horizon inventory models. *IIE Transactions* 33(10):861–874.

Gao R (2022) Finite-sample guarantees for Wasserstein distributionally robust optimization: Breaking the curse of dimensionality. *Operations Research (Forthcoming)* .

Gao R, Chen X, Kleywegt AJ (2022) Wasserstein distributionally robust optimization and variation regularization. *Operations Research (Forthcoming)* .

Gao R, Kleywegt A (2022) Distributionally robust stochastic optimization with Wasserstein distance. *Mathematics of Operations Research (Forthcoming)* .

Georghiou A, Tsoukalas A, Wiesemann W (2021) On the optimality of affine decision rules in robust and distributionally robust optimization. *Optimization Online* .

Han Q, Du D, Zuluaga LF (2014) A risk-and ambiguity-averse extension of the max-min newsvendor order formula. *Operations Research* 62(3):535–542.

Hanasusanto GA, Kuhn D, Wallace SW, Zymler S (2015a) Distributionally robust multi-item newsvendor problems with multimodal demand distributions. *Mathematical Programming* 152(1):1–32.

Hanasusanto GA, Kuhn D, Wiesemann W (2015b) K-adaptability in two-stage robust binary programming. *Operations Research* 63(4):877–891.

Hanasusanto GA, Kuhn D, Wiesemann W (2016) K-adaptability in two-stage distributionally robust binary programming. *Operations Research Letters* 44(1):6–11.

Hannah L, Powell W, Blei D (2010) Nonparametric density estimation for stochastic optimization with an observable state variable. *Advances in Neural Information Processing Systems* 23:820–828.

Hu Y, Kallus N, Mao X (2022) Fast rates for contextual linear optimization. *Management Science (Forthcoming)* .

Iancu DA, Sharma M, Sviridenko M (2013) Supermodularity and affine policies in dynamic robust optimization. *Operations Research* 61(4):941–956.

Kallus N, Mao X (2022) Stochastic optimization forests. *Management Science (Forthcoming)* .

Kannan R, Bayraksan G, Luedtke JR (2020a) Data-driven sample average approximation with covariate information. *Optimization Online* preprint.

Kannan R, Bayraksan G, Luedtke JR (2020b) Residuals-based distributionally robust optimization with covariate information. *arXiv preprint* arXiv:2012.01088.

Kuhn D, Esfahani PM, Nguyen VA, Shafieezadeh-Abadeh S (2019) Wasserstein distributionally robust optimization: Theory and applications in machine learning. *Operations Research & Management Science in the Age of Analytics*, 130–166 (INFORMS).

Lee S, Homem-de Mello T, Kleywegt AJ (2012) Newsvendor-type models with decision-dependent uncertainty. *Mathematical Methods of Operations Research* 76(2):189–221.

Lee S, Kim H, Moon I (2020) A data-driven distributionally robust newsvendor model with a Wasserstein ambiguity set. *Journal of the Operational Research Society* 1–19.

Liyanage LH, Shanthikumar JG (2005) A practical inventory control policy using operational statistics. *Operations Research Letters* 33(4):341–348.

Loke GG, Tang Q, Xiao Y (2020) Decision-driven regularization: Harmonizing the predictive and prescriptive. *Available at SSRN* 3623006.

Luxburg Uv, Bousquet O (2004) Distance-based classification with Lipschitz functions. *Journal of Machine Learning Research* 5:669–695, ISSN 1532-4435.

Meng Q, Shi Y, Qi Y, Ma C, Yuan R, Wu D, Shen ZM (2021) A practical end-to-end inventory management model with deep learning. *Management Science* (Forthcoming).

Mohajerin Esfahani P, Kuhn D (2018) Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming* 171(1):115–166.

Natarajan K, Sim M, Uichanco J (2018) Asymmetry and ambiguity in newsvendor models. *Management Science* 64(7):3146–3167.

Oroojlooyjadid A, Snyder LV, Takáč M (2020) Applying deep learning to the newsvendor problem. *IISE Transactions* 52(4):444–463.

Pentaho (2008) Foodmart's database tables. http://pentaho.dlpage.phi-integration.com/mondrian/mysql-foodmart-database.

Perakis G, Roels G (2008) Regret in the newsvendor model with partial information. *Operations research* 56(1):188–203.

Postek K, Hertog Dd (2016) Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS Journal on Computing* 28(3):553–574.

Rahimian H, Bayraksan G, Homem-de Mello T (2019a) Controlling risk and demand ambiguity in newsvendor models. *European Journal of Operational Research* 279(3):854–868.

Rahimian H, Bayraksan G, Homem-de Mello T (2019b) Identifying effective scenarios in distributionally robust stochastic programs with total variation distance. *Mathematical Programming* 173(1):393–430.

Scarf H (1958) A min-max solution of an inventory problem. *Studies in the mathematical theory of inventory and production* .

See CT, Sim M (2010) Robust approximation to multiperiod inventory management. *Operations research* 58(3):583–594.

Shafieezadeh-Abadeh S, Kuhn D, Esfahani PM (2019) Regularization via mass transportation. *Journal of Machine Learning Research* 20(103):1–68.

Shalev-Shwartz S, Ben-David S (2014) *Understanding machine learning: From theory to algorithms* (Cambridge University press).

Shapiro A, Dentcheva D, Ruszczyński A (2014) *Lectures on stochastic programming: modeling and theory* (SIAM).

Shapley L (1964) Some topics in two-person games. *Advances in Game Theory* 52:1–29.

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM national conference*, 517–524.

Srivastava PR, Wang Y, Hanasusanto GA, Ho CP (2021) On data-driven prescriptive analytics with side information: A regularized Nadaraya–Watson approach. *arXiv preprint arXiv:2110.04855* .

Sturt B (2021) A nonparametric algorithm for optimal stopping based on robust optimization. *arXiv preprint arXiv:2103.03300* .

Subramanyam A, Gounaris CE, Wiesemann W (2019) K-adaptability in two-stage mixed-integer robust optimization. *Mathematical Programming Computation* 1–32.

Toktay LB, Wein LM (2001) Analysis of a forecasting-production-inventory system with stationary demand. *Management Science* 47(9):1268–1281.

Wang Z, Glynn PW, Ye Y (2016) Likelihood robust optimization for data-driven problems. *Computational Management Science* 13(2):241–261.

Xin L, Goldberg DA (2021) Time (in) consistency of multistage distributionally robust inventory models with moment constraints. *European Journal of Operational Research* 289(3):1127–1141.

Xin L, Goldberg DA (2022) Distributionally robust inventory control when demand is a martingale. *Mathematics of Operations Research (Forthcoming)* .

Yanıkoğlu İ, Gorissen BL, den Hertog D (2019) A survey of adjustable robust optimization. *European Journal of Operational Research* 277(3):799–813.

Zhu K, Thonemann UW (2004) An adaptive forecasting algorithm and inventory policy for products with short life cycles. *Naval Research Logistics (NRL)* 51(5):633–653.

Zhu L, Huang M, Li R (2012) Semiparametric quantile regression with high-dimensional covariates. *Statistica Sinica* 22(4):1379.

Zhu T, Xie J, Sim M (2022) Joint estimation and robustness optimization. *Management Science* 68(3):1659–1677.

Appendices to *"Optimal Robust Policy for Feature-based newsvendor"*

Appendix A: Proofs for Section 3.1

The following Lemma EC.1 is a direct consequence of the strong duality result in Wasserstein distributionally robust optimization (e.g., Gao and Kleywegt (2022), Mohajerin Esfahani and Kuhn (2018), Blanchet and Murthy (2019)). To ease the notation, in the sequel we denote $\Psi_f(x,z) := \Psi(f(x),z)$ and $\Psi_{\widehat{f}}(\widehat{x},z) = \Psi(\widehat{f}(\widehat{x}),z)$.

Lemma EC.1. *For each $f \in \mathcal{F}$, the inner primal problem*

$$v_P^f := \sup_{\mathbb{P}\in\mathcal{P}_1(\mathcal{X}\times\mathcal{Z})} \left\{ \mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi_f(X,Z)] : \mathcal{W}(\mathbb{P},\widehat{\mathbb{P}}) \le \rho \right\}.$$

is equal to the following inner dual problem,

$$v_D^f := \inf_{\lambda\ge 0} \left\{ \lambda\rho + \mathbb{E}_{(\widehat{X},\widehat{Z})\sim\widehat{\mathbb{P}}} \left[\sup_{(x,z)\in\mathcal{X}\times\mathcal{Z}} \left\{ \Psi_f(x,z) - \lambda(\|x-\widehat{X}\| + |z-\widehat{Z}|) \right\} \right] \right\}.$$

Note that $v_P^f = v_D^f$ can be infinite if $\limsup_{(x,z)\to\infty} \Psi(f(x),z) = \infty$, but in this case f cannot be a minimizer of (P).

Proof of Lemma 1. Denote $y_k = \widehat{f}(\widehat{x}_k)$, and we define

$$\begin{aligned}
&A_{kk}(x) := y_k, \qquad k = 1,\dots,K,\\
&A_{jk}(x) := \frac{\|x-\widehat{x}_k\|}{\|x-\widehat{x}_j\| + \|x-\widehat{x}_k\|} y_j + \frac{\|x-\widehat{x}_j\|}{\|x-\widehat{x}_j\| + \|x-\widehat{x}_k\|} y_k, \qquad j \ne k,\\
&A^+(x) := \min_{1\le k\le K} \max_{1\le j\le K} A_{jk}(x), \qquad A^-(x) := \max_{1\le k\le K} \min_{1\le j\le K} A_{jk}(x).
\end{aligned}$$

In Figure EC.1, we plot the graph of the function A_{12} when $K=2$ (left) and A_{12}, A_{23}, A_{13} when $K=3$ (right), in the case $\mathcal{X} = \mathbb{R}^2$. Same as the setting of Figure 2, when $K=2$ the Shapley policy is A_{12}, and when $K=3$ the Shapley policy is the middle one among A_{12}, A_{13} and A_{23}, which is rendered with a mesh in this figure.

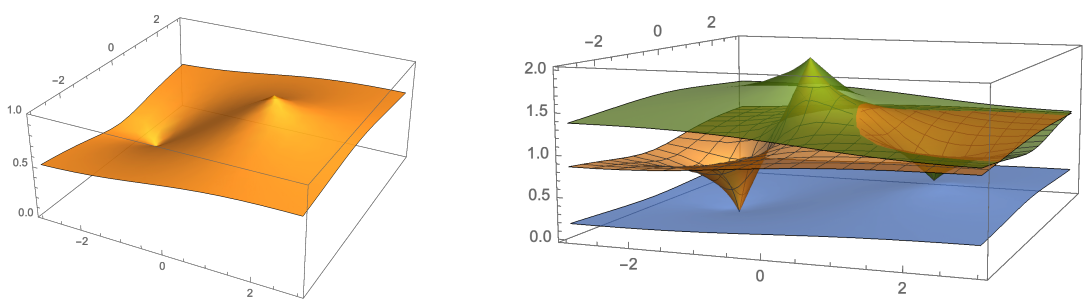


Figure EC.1 Graph of the Shapley policy $y = f(x)$ when $K = 2, 3$, $x \in \mathbb{R}^2$

We claim A^+ and A^- both satisfy the four properties. First, we show they are indeed extensions. Fix $\ell \in 1,\dots,K$, then $A_{j\ell}(\widehat{x}_\ell) = A_{\ell j}(\widehat{x}_\ell) = y_\ell$ for every j. This implies

$$\begin{aligned}
A^+(\widehat{x}_\ell) &= \min_{1\le k\le K} \max_{1\le j\le K} A_{jk}(\widehat{x}_\ell) \le \max_{1\le j\le K} A_{j\ell}(\widehat{x}_\ell) = y_\ell,\\
A^-(\widehat{x}_\ell) &= \max_{1\le k\le K} \min_{1\le j\le K} A_{jk}(\widehat{x}_\ell) \ge \min_{1\le j\le K} A_{j\ell}(\widehat{x}_\ell) = y_\ell.
\end{aligned}$$

However, $A^+ \ge A^-$, so in fact $A^+(\widehat{x}_\ell) = A^-(\widehat{x}_\ell) = y_\ell$, that is, A^+ and A^- interpolate given data.

Next we show the boundedness and the Lipschitzness. It suffices to show them for each A_{jk}, because both bounds are compatible with min max operations. Because $A_{jk}(x)$ is just an interpolation between y_j and y_k,

clearly we have $\min\{y_j, y_k\} \le A_{jk}(x) \le \max\{y_j, y_k\}$. As for the Lipschitz bound of A_{jk}, when $j = k$, $A_{kk} \equiv y_k$ are constant functions, so they always satisfy the Lipschitz bound. When $j \ne k$, fix $x, x' \in \mathcal{X}$, and we denote

$$d_{xj} = \|\widehat{x}_j - x\|, \quad d_{x'j} = \|\widehat{x}_j - x'\|, \quad d_{xk} = \|\widehat{x}_k - x\|, \quad d_{x'k} = \|\widehat{x}_k - x'\|, \quad d_{xx'} = \|x - x'\|, \quad d_{jk} = \|\widehat{x}_j - \widehat{x}_k\|.$$

Then

$$\begin{aligned}
A_{jk}(x) - A_{jk}(x') &= \frac{d_{xk}}{d_{xj} + d_{xk}} y_j + \frac{d_{xj}}{d_{xj} + d_{xk}} y_k - \frac{d_{x'k}}{d_{x'j} + d_{x'k}} y_j - \frac{d_{x'j}}{d_{x'j} + d_{x'k}} y_k \\
&= (y_j - y_k) \left(\frac{d_{xk}}{d_{xj} + d_{xk}} - \frac{d_{x'k}}{d_{x'j} + d_{x'k}} \right) \\
&= (y_j - y_k) \left(\frac{d_{xk} d_{x'j} + d_{xk} d_{x'k} - d_{xj} d_{x'k} - d_{xk} d_{x'k}}{(d_{xj} + d_{xk})(d_{x'j} + d_{x'k})} \right) \\
&= (y_j - y_k) \left(\frac{d_{xk} d_{x'j} - d_{xk} d_{xj} + d_{xj} d_{xk} - d_{xj} d_{x'k}}{(d_{xj} + d_{xk})(d_{x'j} + d_{x'k})} \right) \\
&= (y_j - y_k) \left(\frac{d_{xk}(d_{x'j} - d_{xj}) + d_{xj}(d_{xk} - d_{x'k})}{(d_{xj} + d_{xk})(d_{x'j} + d_{x'k})} \right).
\end{aligned}$$

By triangular inequality,

$$\begin{aligned}
|A_{jk}(x) - A_{jk}(x')| &\le |y_j - y_k| \left(\frac{d_{xk} d_{xx'} + d_{xj} d_{xx'}}{(d_{xj} + d_{xk})(d_{x'j} + d_{x'k})} \right) \\
&\le |y_j - y_k| \left(\frac{d_{xx'}}{d_{x'j} + d_{x'k}} \right) \\
&\le |y_j - y_k| \left(\frac{d_{xx'}}{d_{jk}} \right) \\
&= \frac{|y_j - y_k|}{\|\widehat{x}_j - \widehat{x}_k\|} \|x - x'\|.
\end{aligned}$$

Thus if we denote $L := \max_{j \ne k} \dfrac{|y_j - y_k|}{\|\widehat{x}_j - \widehat{x}_k\|}$ to be the discrete Lipschitz constant of the given data, then all the A_{jk} are L-Lipschitz in x, so there min and max are also L-Lipschitz in x.

It remains to prove (5), which is to show that $y = A^+(x), A^-(x)$ satisfy the following condition for every k:

$$\Phi(y) - \|x - \widehat{x}_k\| \le \max_{j=1,\dots,K} \left\{ \Phi(y_j) - \|\widehat{x}_j - \widehat{x}_k\| \right\} =: M_k. \tag{Mk}$$

We first claim that $y = A_{jk}(x)$ satisfy the bound (Mk). By convexity of Φ,

$$\begin{aligned}
\Phi(A_{jk}(x)) - \|x - \widehat{x}_k\| &= \Phi\left(\frac{\|x - \widehat{x}_k\|}{\|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|} y_j + \frac{\|x - \widehat{x}_j\|}{\|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|} y_k \right) - \|x - \widehat{x}_k\| \\
&\le \frac{\|x - \widehat{x}_k\|}{\|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|} \Phi(y_j) + \frac{\|x - \widehat{x}_j\|}{\|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|} \Phi(y_k) - \|x - \widehat{x}_k\|.
\end{aligned}$$

Using definition of M_k in (Mk),

$$\Phi(y_j) \le M_k + \|\widehat{x}_j - \widehat{x}_k\|, \qquad \Phi(y_k) \le M_k + \|\widehat{x}_k - \widehat{x}_k\| = M_k,$$

Plug in into the above inequality,

$$\begin{aligned}
\Phi(A_{jk}(x)) - \|x - \widehat{x}_k\| &\le M_k + \frac{\|x - \widehat{x}_k\|}{\|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|} \|\widehat{x}_j - \widehat{x}_k\| - \|x - \widehat{x}_k\| \\
&= M_k + \|x - \widehat{x}_k\| \left(\frac{\|\widehat{x}_j - \widehat{x}_k\|}{\|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|} - 1 \right) \le M_k.
\end{aligned}$$

In the last step we used the triangle inequality $\|\widehat{x}_j - \widehat{x}_k\| \le \|x - \widehat{x}_j\| + \|x - \widehat{x}_k\|$. Fixing a k in $1, \dots, K$, then from the definition we can find j_1, j_2 such that $A_{j_1 k}(x) \le A^-(x) \le A^+(x) \le A_{j_2 k}(x)$. Because in a closed interval

$[A_{j_1k}(x), A_{j_2k}(x)]$ convex function Φ can only attain maximum at the endpoints due to the maximum principle, we conclude that

$$\Phi(A^-(x)), \Phi(A^+(x)) \le \max\{\Phi(A_{j_1k}(x)), \Phi(A_{j_2k}(x))\}.$$

As a result, since both $y = A_{j_1k}(x)$ and $y = A_{j_2k}(x)$ satisfy (Mk), we conclude that $y = A^+(x), A^-(x)$ also satisfy (Mk), which implies (5).

The existence of the saddle point is proved in Lemma EC.2 below. It is purely algebraic and it makes use of the Shapley's theorem. □

Lemma EC.2. *With the same notation as Lemma 1,* $A^+ = A^-$.

Proof. We fix an x and then omit the x in $A_{jk}(x)$ to ease the notation. Denote the symmetric matrix $A = (A_{jk})_{jk}$, and we want to show that for this matrix, $\min_k \max_j A_{jk} = \max_j \min_k A_{jk}$, i.e., a saddle point exists. By Theorem 2.1 in Shapley (1964), to show the existence of a saddle point for A, it is sufficient to show that any 2×2 submatrix of A has a saddle point.

For a general 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we claim that if it has no saddle point then it has the *diagonal dominant property*: the two elements on one diagonal is strictly greater than the two elements on the other diagonal. To show this claim, we note that the maximin being different from the minimax means

$$(a \wedge c) \vee (b \wedge d) \ne (a \vee b) \wedge (c \vee d).$$

Without loss of generality, assume a is the smallest entry. Then the above inequality is simplified to

$$b \wedge d \ne b \wedge (c \vee d).$$

If $d \ge c$, then both sides equals to $b \wedge d$. If $d \ge b$, then both sides equals to b. Hence the above inequality can only hold if $d < c$ and $d < b$, which implies a, d are both strictly smaller than b, c.

Applying to our case, we need to show that for any i, j, k, l, the matrix

$$B = \begin{pmatrix} A_{ik} & A_{jk} \\ A_{i\ell} & A_{j\ell} \end{pmatrix}$$

doesn't have the diagonal dominant property. Recall that A_{jk} is an interpolation of y_j and y_k, so it is important to compare the values of y_i, y_j, y_k, y_ℓ. Without loss of generality, assume $y_i \ge y_j$, $y_k \ge y_\ell$, and assume $y_i \ge y_k$ using the transpose symmetry. There are three possibilities:

(I) $y_i \ge y_k \ge y_\ell > y_j$.

(II) $y_i \ge y_k \ge y_j \ge y_\ell$.

(III) $y_i \ge y_j > y_k \ge y_\ell$.

In case (I), $A_{ik} \ge y_k \ge A_{jk}$, $A_{j\ell} \le y_\ell \le A_{il}$, so neither diagonal can dominate the other. In case (II), $A_{ik} \ge y_k \ge A_{jk} \ge y_j \ge A_{j\ell}$, again neither diagonal can dominate the other. The third case needs a further discussion.

Since A^+, A^- are both extension of $\widehat{f}$, they always agree on $\widehat{\mathcal{X}}$, so we may assume $x \ne \widehat{x}_i, \widehat{x}_j, \widehat{x}_k, \widehat{x}_\ell$ thus

$$d_i = \|x - \widehat{x}_i\|, \qquad d_j = \|x - \widehat{x}_j\|, \qquad d_k = \|x - \widehat{x}_k\|, \qquad d_\ell = \|x - \widehat{x}_\ell\|$$

are all positive. Recall that $A_{jk} = \frac{d_k}{d_j+d_k} y_j + \frac{d_j}{d_j+d_k} y_k$. We prove by contradiction and assume B has the diagonal dominant property. If the main diagonal is strictly greater than the off diagonal, then

$$A_{ik}, A_{j\ell} > A_{jk}, A_{i\ell}.$$

For instance, from $A_{ik} > A_{jk}$ we have

$$\begin{aligned}
\frac{d_k}{d_i+d_k} y_i + \frac{d_i}{d_i+d_k} y_k &> \frac{d_k}{d_j+d_k} y_j + \frac{d_j}{d_j+d_k} y_k \\
\frac{d_k}{d_i+d_k}(y_i - y_k) + y_k &> \frac{d_k}{d_j+d_k}(y_j - y_k) + y_k \\
\frac{d_k}{d_i+d_k}(y_i - y_k) &> \frac{d_k}{d_j+d_k}(y_j - y_k) \\
(d_j+d_k)(y_i - y_k) &> (d_i+d_k)(y_j - y_k).
\end{aligned} \tag{EC.1}$$

Similarly, from $A_{ik} > A_{i\ell}, A_{j\ell} > A_{jk}, A_{j\ell} > A_{i\ell}$ we conclude

$$\begin{aligned}(d_i + d_k)(y_i - y_\ell) &> (d_i + d_\ell)(y_i - y_k),\\ (d_j + d_\ell)(y_j - y_k) &> (d_j + d_k)(y_j - y_\ell),\\ (d_i + d_\ell)(y_j - y_\ell) &> (d_j + d_\ell)(y_i - y_\ell).\end{aligned}$$

Note that in case (III), every term in (EC.1) and the above three inequalities is positive. So if we multiply four inequalities, we would reach a contradiction.

If the main diagonal is strictly smaller than the off diagonal, then all the inequalities above flip sign, and we would still reach a contradiction. In conclusion, B never has the diagonal dominant property. In other words, B admits a saddle point. By Theorem 2.1 in Shapley (1964), A also admits a saddle point, therefore $A^+ = A^-$. □

Proof of Theorem 1. To show the direction $v_D \geq v_{\hat{D}}$, note that $v_{\hat{D}}$ can be written with $f \in \mathcal{F}$ instead of $\widehat{f} \in \widehat{\mathcal{F}}$:

$$v_{\hat{D}} = \inf_{f \in \mathcal{F}} \inf_{\lambda \geq 0} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\sup_{z \in \mathcal{Z}} \max_{x \in \widehat{\mathcal{X}}} \left\{ \Psi(f(x), z) - \lambda \|x - \widehat{X}\| - \lambda |z - \widehat{Z}| \right\} \right] \right\}$$

since the target function depends only on the value of the restriction $\widehat{f} = f|_{\widehat{\mathcal{X}}}$. From this, $v_D \geq v_{\hat{D}}$ is straightforward because we are restricting the set over which the supremum of x is taken.

To show $v_D \leq v_{\hat{D}}$, we let f be the Shapley extension of a given $\widehat{f} \in \widehat{\mathcal{F}}$ as defined in Lemma 1, and split into two cases: $\lambda > 0$ and $\lambda = 0$. If $\lambda > 0$, by Lemma 1 we know that f satisfies (5). In particular, If we choose $\Phi(y) = \frac{1}{\lambda}\Psi(y, z) - |z - \widehat{Z}|$, then

$$\sup_{x \in \mathcal{X}} \left\{ \Psi(f(x), z) - \lambda \|x - \widehat{X}\| \right\} \leq \max_{x \in \widehat{\mathcal{X}}} \left\{ \Psi(\widehat{f}(x), z) - \lambda \|x - \widehat{X}\| \right\}, \qquad \text{for all } \hat{X} \in \widehat{\mathcal{X}}, z \in \mathcal{Z}.$$

Consequently,

$$\begin{aligned}&\lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\sup_{z \in \mathcal{Z}} \sup_{x \in \mathcal{X}} \left\{ \Psi(f(x), z) - \lambda \|x - \widehat{X}\| - \lambda |z - \widehat{Z}| \right\} \right]\\ &\quad \leq \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\sup_{z \in \mathcal{Z}} \max_{x \in \widehat{\mathcal{X}}} \left\{ \Psi(\widehat{f}(x), z) - \lambda \|x - \widehat{X}\| - \lambda |z - \widehat{Z}| \right\} \right].\end{aligned}$$

If $\lambda = 0$, by Lemma 1 we know that the range of f is $[\min \widehat{f}, \max \widehat{f}]$. Since $\Psi(\cdot, z)$ is convex, the supremum of $\Psi(f(\cdot), z)$ is attained at the extreme points, so

$$\mathbb{E}_{\widehat{\mathbb{P}}} \left[\sup_{z \in \mathcal{Z}} \sup_{x \in \mathcal{X}} \{ \Psi(f(x), z) \} \right] \leq \mathbb{E}_{\widehat{\mathbb{P}}} \left[\sup_{z \in \mathcal{Z}} \max_{x \in \widehat{\mathcal{X}}} \left\{ \Psi(\widehat{f}(x), z) \right\} \right].$$

Therefore, taking infimum in $\lambda \geq 0$ gives

$$\begin{aligned}&\inf_{\lambda \geq 0} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\sup_{z \in \mathcal{Z}} \sup_{x \in \mathcal{X}} \left\{ \Psi(f(x), z) - \lambda \|x - \widehat{X}\| - \lambda |z - \widehat{Z}| \right\} \right] \right\}\\ &\quad \leq \inf_{\lambda \geq 0} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\sup_{z \in \mathcal{Z}} \max_{x \in \widehat{\mathcal{X}}} \left\{ \Psi(\widehat{f}(x), z) - \lambda \|x - \widehat{X}\| - \lambda |z - \widehat{Z}| \right\} \right] \right\}.\end{aligned}$$

The left dominates v_D, so taking the inf over $\widehat{f} \in \widehat{\mathcal{F}}$ on the right gives $v_D \leq v_{\hat{D}}$, which completes the proof of $v_D = v_{\hat{D}}$. Note that the above also proves that if $\widehat{f}^*$ is a minimizer of $v_{\hat{D}}$, then the Shapley extension of $\widehat{f}^*$ is a minimizer of v_D, also a minimizer of v_P by Lemma EC.1. □

Appendix B: Proofs for Section 4.1

Lemma EC.3. *For each fixed $\lambda \geq b \vee h$ and $\widehat{f} \in \widehat{\mathcal{F}}$, we can always find a $\frac{\lambda}{b \vee h}$-Lipschitz policy $\widehat{g} \in \widehat{\mathcal{F}}$ satisfying $\underline{z}_k \leq \widehat{g}(\widehat{x}_k) \leq \overline{z}_k$, where*

$$\underline{z}_k := \min_{(x,z) \in \operatorname{supp} \widehat{\mathbb{P}}} \{ z + \|x - \widehat{x}_k\| \}, \qquad \overline{z}_k := \max_{(x,z) \in \operatorname{supp} \widehat{\mathbb{P}}} \{ z - \|x - \widehat{x}_k\| \},$$

such that for all $(\widehat{x}, \widehat{z}) \in \operatorname{supp} \widehat{\mathbb{P}}$,

$$\max_{x \in \widehat{\mathcal{X}}} \{ \Psi(\widehat{g}(x), \widehat{z}) - \lambda \|x - \widehat{x}\| \} \leq \max_{x \in \widehat{\mathcal{X}}} \left\{ \Psi(\widehat{f}(x), \widehat{z}) - \lambda \|x - \widehat{x}\| \right\}.$$

Proof of Lemma EC.3. Denote $y_k = \widehat{f}(\widehat{x}_k)$. First, we show that if $\widehat{f}$ is "not Lipschitz enough" at some point, in the sense that its local Lipschitz constant at a point is too large, then we can reduce it by modifying $\widehat{f}$. Suppose there exists j_0, k_0 such that

$$y_{j_0} \geq y_{k_0} + L\|\widehat{x}_{k_0} - \widehat{x}_{j_0}\| := \tilde{y}_{j_0}$$

for some constant $L > 0$ to be specify later. We claim that replacing decision y_{j_0} by $\tilde{y}_{j_0}$ will not deteriorate the objective value, in the sense that for any $(\widehat{x}, \widehat{z})$ in the support of $\widehat{\mathbb{P}}$,

$$h(\tilde{y}_{j_0} - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_{j_0})_+ - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| \leq \max_{j=1,\ldots,K} \left\{h(y_j - \widehat{z})_+ + b(\widehat{z} - y_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\|\right\} =: \text{RHS}.$$

We consider two cases. If $\tilde{y}_{j_0} \geq \widehat{z}$, then

$$\begin{aligned} h(\tilde{y}_{j_0} - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_{j_0})_+ - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| &= h(\tilde{y}_{j_0} - \widehat{z}) - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| \\ &\leq h(y_{j_0} - \widehat{z}) - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| \leq \text{RHS}. \end{aligned}$$

If $\tilde{y}_{j_0} < \widehat{z}$, then

$$\begin{aligned} h(\tilde{y}_{j_0} - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_{j_0})_+ - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| &= b(\widehat{z} - \tilde{y}_{j_0}) - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| \\ &= b(\widehat{z} - y_{k_0} - L\|\widehat{x}_{k_0} - \widehat{x}_{j_0}\|) - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| \\ &= b(\widehat{z} - y_{k_0}) - bL\|\widehat{x}_{k_0} - \widehat{x}_{j_0}\| - \lambda\|\widehat{x}_{j_0} - \widehat{x}\|. \end{aligned}$$

If we choose any $L \geq \frac{\lambda}{b}$, then

$$\begin{aligned} h(\tilde{y}_{j_0} - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_{j_0})_+ - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| &\leq b(\widehat{z} - y_{k_0}) - \lambda\|\widehat{x}_{k_0} - \widehat{x}_{j_0}\| - \lambda\|\widehat{x}_{j_0} - \widehat{x}\| \\ &\leq b(\widehat{z} - y_{k_0}) - \lambda\|\widehat{x}_{k_0} - \widehat{x}\| \leq \text{RHS}. \end{aligned}$$

This completes the proof of the claim.

Now we make use of the above claim recursively.

`Step 0.` Denote $y_k^{(1)} = y_k$ for every $k \in [K]$.

`Step 1.` Pick $k_1 \in \arg\min_{k\in[K]}\{y_k^{(1)}\}$, and define $y_j^{(2)} = y_j^{(1)} \wedge (y_{k_1}^{(1)} + L\|\widehat{x}_j - \widehat{x}_{k_1}\|)$ for every $j \in [K]$.

`Step 2.` Pick $k_2 \in \arg\min_{k\in[K]\setminus\{k_1\}}\{y_k^{(2)}\}$, and define $y_j^{(3)} = y_j^{(2)} \wedge (y_{k_2}^{(2)} + L\|\widehat{x}_j - \widehat{x}_{k_2}\|)$ for every $j \in [K]$.

`Step 3.` Pick $k_3 \in \arg\min_{k\in[K]\setminus\{k_1,k_2\}}\{y_k^{(3)}\}$, and define $y_j^{(4)} = y_j^{(3)} \wedge (y_{k_3}^{(3)} + L\|\widehat{x}_j - \widehat{x}_{k_3}\|)$ for every $j \in [K]$ ...

The above process terminates after `Step K-1`. We have a sequence of policies $\widehat{f}^{(m)}$ defined by $\widehat{f}^{(m)}(\widehat{x}_k) = y_k^{(m)}$. According to the previous claim, each step does not deteriorate the objective value: for any $1 \leq m \leq K-1$,

$$\max_{x\in\widehat{\mathcal{X}}} \left\{\Psi(\widehat{f}^{(m+1)}(x), \widehat{z}) - \lambda\|x - \widehat{x}\|\right\} \leq \max_{x\in\widehat{\mathcal{X}}} \left\{\Psi(\widehat{f}^{(m)}(x), \widehat{z}) - \lambda\|x - \widehat{x}\|\right\}.$$

It is easy to conclude that our selection has the following properties:

(I) It is decreasing: $\widehat{f}^{(m+1)} \leq \widehat{f}^{(m)}$.

(II) The sequence $y_{k_m}^{(m)}$ is increasing in m, that is,

$$y_{k_1}^{(1)} \leq y_{k_2}^{(2)} \leq y_{k_3}^{(3)} \leq \cdots \leq y_{k_K}^{(K)}.$$

This is because $y_{k_{m+1}}^{(m)} \geq y_{k_m}^{(m)}$ since k_m is the argmin in step k, and by definition we have

$$y_{k_{m+1}}^{(m+1)} = y_{k_{m+1}}^{(m)} \wedge (y_{k_m}^{(m)} + L\|\widehat{x}_{k_m} - \widehat{x}_{k_{m+1}}\|) \geq y_{k_m}^{(m)}.$$

(III) The above increasing order implies the value at $\widehat{x}_{k_m}$ stops decreasing after step m:

$$y_{k_1}^{(1)} = y_{k_1}^{(2)} = \cdots = y_{k_1}^{(K)}, \qquad y_{k_2}^{(2)} = y_{k_2}^{(3)} = \cdots = y_{k_2}^{(K)}, \qquad y_{k_3}^{(3)} = y_{k_3}^{(4)} = \cdots = y_{k_3}^{(K)}, \qquad \ldots$$

again following the definition. Therefore $\widehat{f}^{(K)}(\widehat{x}_{k_m}) = y_{k_m}^{(m)}$.

Combine the above three properties, we have for any $m < n$,

$$y_{k_m}^{(m)} \le y_{k_n}^{(n)} \le y_{k_n}^{(m+1)} \le y_{k_m}^{(m)} + L\|\widehat{x}_{k_n} - \widehat{x}_{k_m}\|.$$

Now we define $\tilde{f} = \widehat{f}^{(K)}$, then it is L-Lipschitz. A similar argument works for $L \ge \frac{\lambda}{h}$, so we can pick $L = \frac{\lambda}{b} \wedge \frac{\lambda}{h} = \frac{\lambda}{b \vee h}$, and by the above construction $\tilde{f}$ is $\frac{\lambda}{b\vee h}$-Lipschitz and satisfy

$$\max_{j=1,\dots,K} \left\{h(\tilde{f}(\widehat{x}_j) - \widehat{z})_+ + b(\widehat{z} - \tilde{f}(\widehat{x}_j))_+ - \lambda\|\widehat{x}_j - \widehat{x}\|\right\} \le \max_{j=1,\dots,K} \left\{h(y_j - \widehat{z})_+ + b(\widehat{z} - y_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\|\right\}$$

for all $(\widehat{x}, \widehat{z}) \in \operatorname{supp} \widehat{\mathbb{P}}$, that is,

$$\max_{x\in\widehat{\mathcal{X}}} \left\{\Psi(\tilde{f}(x), \widehat{z}) - \lambda\|x - \widehat{x}\|\right\} \le \max_{x\in\widehat{\mathcal{X}}} \left\{\Psi(\widehat{f}(x), \widehat{z}) - \lambda\|x - \widehat{x}\|\right\}.$$

Now we deal with upper and lower bound. By the first part of this proof we can assume without loss of generality that $\widehat{f}$ is already $\frac{\lambda}{b\vee h}$-Lipschitz to begin with. Define $\tilde{y}_j = \overline{z}_j \wedge y_j$ for every j, we claim that for any $(\widehat{x}, \widehat{z})$ in the support of $\widehat{\mathbb{P}}$,

$$\max_{j=1,\dots,K} \left\{h(\tilde{y}_j - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\|\right\} \le \max_{j=1,\dots,K} \left\{h(y_j - \widehat{z})_+ + b(\widehat{z} - y_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\|\right\}.$$

Indeed, if $\tilde{y}_j = y_j$, then we are not changing anything, directly we have

$$h(\tilde{y}_j - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\| = h(y_j - \widehat{z})_+ + b(\widehat{z} - y_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\| \le \text{RHS}.$$

When $\tilde{y}_j = \overline{z}_j \le y_j$, we split to two cases. On the one hand, if $\widehat{z} \le \tilde{y}_j$, then

$$h(\tilde{y}_j - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\| \le h(\tilde{y}_j - \widehat{z}) - \lambda\|\widehat{x}_j - \widehat{x}\| \le h(y_j - \widehat{z}) - \lambda\|\widehat{x}_j - \widehat{x}\| \le \text{RHS}.$$

On the other hand, if $\widehat{z} \ge \tilde{y}_j$, using $\tilde{y}_j = \overline{z}_j \ge \widehat{z} - \|\widehat{x} - \widehat{x}_j\|$ by the definition of $\overline{z}$, so

$$h(\tilde{y}_j - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\| = b(\widehat{z} - \tilde{y}_j) - \lambda\|\widehat{x}_j - \widehat{x}\| \le b\|\widehat{x} - \widehat{x}_j\| - \lambda\|\widehat{x}_j - \widehat{x}\| \le 0 \le \text{RHS}.$$

Here we used that $\lambda \ge b$, and the right hand side is always nonnegative because it is nonnegative when $\widehat{x}_j = \widehat{x}$. In conclusion, for every scenario we have

$$h(\tilde{y}_j - \widehat{z})_+ + b(\widehat{z} - \tilde{y}_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\| \le \max_{j=1,\dots,K} \left\{h(y_j - \widehat{z})_+ + b(\widehat{z} - y_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\|\right\}.$$

Take maximum on the left over j completes the proof of the claim.

Similarly we can let $\tilde{\tilde{y}}_j = \underline{z}_j \vee \tilde{y}_j$, and $\tilde{\tilde{y}}_j$ will satisfy

$$\max_{j=1,\dots,K} \left\{h(\tilde{\tilde{y}}_j - \widehat{z})_+ + b(\widehat{z} - \tilde{\tilde{y}}_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\|\right\} \le \max_{j=1,\dots,K} \left\{h(y_j - \widehat{z})_+ + b(\widehat{z} - y_j)_+ - \lambda\|\widehat{x}_j - \widehat{x}\|\right\}.$$

Now we define $\widehat{g}(\widehat{x}_j) = \underline{z}_j \vee (\overline{z}_j \wedge \tilde{f}(\widehat{x}_j))$ for every j, with $\tilde{f}$ being the $\frac{\lambda}{b\vee h}$-Lipschitz function define in the first part of the proof. Then $\widehat{g}$ satisfies

$$\max_{x\in\widehat{\mathcal{X}}} \left\{\Psi(\widehat{g}(x), \widehat{z}) - \lambda\|x - \widehat{x}\|\right\} \le \max_{x\in\widehat{\mathcal{X}}} \left\{\Psi(\tilde{f}(x), \widehat{z}) - \lambda\|x - \widehat{x}\|\right\} \le \max_{x\in\widehat{\mathcal{X}}} \left\{\Psi(\widehat{f}(x), \widehat{z}) - \lambda\|x - \widehat{x}\|\right\}.$$

Moreover, note that $\widehat{x}_k \mapsto \overline{z}_k$ and $\widehat{x}_k \mapsto \underline{z}_k$ are 1-Lipshitz since they are the max and min of a family of 1-Lipschitz function of $\widehat{x}_k$, and $1 \le \frac{\lambda}{b\vee h}$, so $\widehat{g}$ is $\frac{\lambda}{b\vee h}$-Lipschitz, and by definition $\underline{z}_k \le \widehat{g}(\widehat{x}_k) \le \overline{z}_k$. □

Proof of Proposition 1. We start by proving that

$$v_{\hat{D}} = \min_{\widehat{f}\in\widehat{\mathcal{F}},\, \lambda\ge b\vee h} \left\{\lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}}\left[\max_{x\in\widehat{\mathcal{X}}} \left\{\Psi_{\widehat{f}}(x, \widehat{Z}) - \lambda\|x - \widehat{X}\|\right\}\right]\right\}. \tag{EC.2}$$

To see this, consider maximizing over z first in the inner maximization of the dual problem

$$v_{\hat{D}} = \inf_{\widehat{f}\in\widehat{\mathcal{F}}} \inf_{\lambda\ge 0} \left\{\lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}}\left[\max_{x\in\widehat{\mathcal{X}}} \left\{\sup_{z\in\mathcal{Z}} \left\{\Psi(\widehat{f}(x), z) - \lambda|z - \widehat{Z}|\right\} - \lambda\|x - \widehat{X}\|\right\}\right]\right\}.$$

If $\lambda < b \vee h = b$, then the sup of

$$\Psi(\widehat{f}(x), z) - \lambda|z - \widehat{Z}| = h(\widehat{f}(x) - z)_+ + b(z - \widehat{f}(x))_+ - \lambda|z - \widehat{Z}|$$

will be infinity as $z \to \infty$. Therefore, in order to find the minimum over λ we can disregard this case and constrain $\lambda \geq b \vee h$. In this case, sup over z is $\Psi(\widehat{f}(x), \widehat{Z})$ attained at $z = \widehat{Z}$, which proves (EC.2).

For each $\lambda \geq b \vee h$, denote

$$\bar{\widehat{\mathcal{F}}} := \left\{ \widehat{g} \in \widehat{\mathcal{F}} : \underline{z}_k \leq \widehat{g}(\widehat{x}_k) \leq \bar{z}_k, \forall k \right\}, \qquad \widehat{\mathcal{F}}_\lambda := \left\{ \widehat{g} \in \widehat{\mathcal{F}} : \|\widehat{g}\|_{\mathrm{Lip}} \leq \frac{\lambda}{b \vee h} \right\}.$$

Then

$$v_{\hat{D}} = \inf_{\lambda \geq b \vee h} \inf_{\widehat{f} \in \bar{\widehat{\mathcal{F}}} \cap \widehat{\mathcal{F}}_\lambda} \left\{ \lambda \rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\max_{x \in \mathcal{X}} \left\{ \Psi(\widehat{f}(x), \widehat{Z}) - \lambda \|x - \widehat{X}\| \right\} \right] \right\},$$

where we replace $\widehat{\mathcal{F}}$ by $\bar{\widehat{\mathcal{F}}} \cap \widehat{\mathcal{F}}_\lambda$ in (EC.2) using Lemma EC.3. For each $\widehat{f} \in \widehat{\mathcal{F}}_\lambda$, because $\|\Psi(\widehat{f}(\cdot), \widehat{z})\|_{\mathrm{Lip}} \leq \|\widehat{f}\|_{\mathrm{Lip}}(b \vee h) \leq \lambda$, the max over x is attained at $x = \widehat{X}$. Therefore,

$$v_{\hat{D}} = \inf_{\lambda \geq b \vee h} \inf_{\widehat{f} \in \bar{\widehat{\mathcal{F}}} \cap \widehat{\mathcal{F}}_\lambda} \left\{ \lambda \rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\Psi(\widehat{f}(\widehat{X}), \widehat{Z}) \right] \right\}.$$

Now we switch the inf over λ and inf over $\widehat{f}$,

$$\begin{aligned} v_{\hat{D}} &= \inf_{\lambda \geq b \vee h} \inf_{\substack{\widehat{f} \in \bar{\widehat{\mathcal{F}}} \\ \|\widehat{f}\|_{\mathrm{Lip}} \leq \frac{\lambda}{b \vee h}}} \left\{ \lambda \rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\Psi(\widehat{f}(\widehat{X}), \widehat{Z}) \right] \right\} \\ &= \inf_{\widehat{f} \in \bar{\widehat{\mathcal{F}}}} \inf_{\substack{\lambda \geq b \vee h \\ \lambda \geq (b \vee h) \|\widehat{f}\|_{\mathrm{Lip}}}} \left\{ \lambda \rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\Psi(\widehat{f}(\widehat{X}), \widehat{Z}) \right] \right\} \\ &= \inf_{\widehat{f} \in \bar{\widehat{\mathcal{F}}}} \left\{ (b \vee h)(1 \vee \|\widehat{f}\|_{\mathrm{Lip}}) \rho + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\Psi(\widehat{f}(\widehat{X}), \widehat{Z}) \right] \right\}. \end{aligned}$$

Using the change of variables $y_k = \widehat{f}(\widehat{x}_k)$, $k = 1, \ldots, K$, the above is equivalent to

$$v_{\hat{D}} = \inf_{y_k \in [\underline{z}_k, \bar{z}_k], 1 \leq k \leq K} \left\{ (b \vee h) \left(1 \vee \max_{i \neq j} \frac{|y_i - y_j|}{\|\widehat{x}_i - \widehat{x}_j\|} \right) \rho + \frac{1}{n} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \Psi(y_k, \widehat{z}_{ki}) \right\}.$$

This is an infimum over K variables which take values in closed intervals, a compact set, so the infimum is attained on a convex subset. Repeat the above argument with $\widehat{\mathcal{F}}$ in place of $\bar{\widehat{\mathcal{F}}}$, we conclude that $v_{\hat{D}} = v_{\hat{R}}$ and ($\widehat{\mathrm{D}}$) is equivalent to ($\widehat{\mathrm{R}}$). □

Appendix C: Additional Results for Section 4.2 and Proofs for Section 4.3

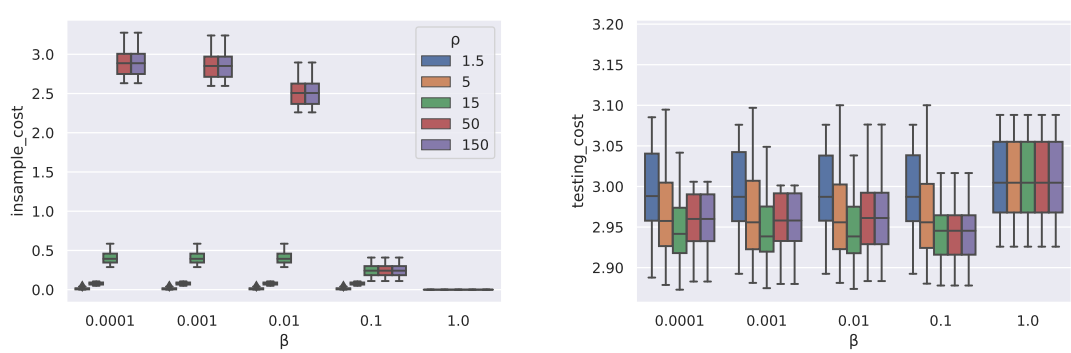


Figure EC.2 Impact of hyperparameters ρ and β on the in-sample and out-of-sample performance

In Figure EC.2, we illustrate the effect of different choices of hyperparameters on the model performance. We plot the in-sample costs and the out-of-sample costs in the case when $d = 5000$, $n = 100$, and $h = b = 1$, with various hyperparameters, under the same setup as in Section 5.1. The in-sample cost increases in the radius

ρ and decreases in the distance weight β. As β becomes large ($\beta = 1$), in-sample cost reduces to zero and the policy becomes empirical risk minimization. This can be seen from the regularization point of view ($\widehat{\mathrm{R}}$): when we are indifferent in Lipschitz norms that are below a very high threshold—higher than the Lipschitz norm of the critical conditional quantile—the optimal policy is simply the unconditional quantile function. The out-of-sample performance seems less sensitive in β as long as it is sufficiently small (less than 0.1), but it has a clear preference in suitable tuning value of ρ. For large or small ρ, the policies are either too conservative or too restrictive, which lead to undesirable out-of-sample cost.

Proof for Proposition 3. First we give an upper bound for λ^* in terms of ρ, using the duality of the primal and the dual problem:

$$\lambda^*\rho \le v_D = v_P \le \max_{\mathcal{X}\times\mathcal{Z}} \Psi(f^*(x), z).$$

If the demand is bounded by $\bar{D}$, then $\mathcal{Z} \subset [0, \bar{D}]$ and subsequently the decision is also bounded by $f^*(x) \in [0, \bar{D}]$ for all x, by the boundedness of the Shapley extension. Therefore $|f^*(x) - z| \le \bar{D}$, thus $\Psi_{f^*}(x, z) \le (b \vee h)\bar{D}$, so

$$\|\Psi_{f^*}\|_{\mathrm{Lip}} = \lambda^* \le \frac{(b \vee h)\bar{D}}{\rho}.$$

Then the result follows from (Shalev-Shwartz and Ben-David 2014, Theorem 26.3 and Lemma 26.9). □

Appendix D: Proofs for Section 4.4

We recall and define $L = \|\widehat{f}^*\|_{\mathrm{Lip}}$ and

$$\begin{aligned} &\widehat{\mathcal{X}}_< := \{\widehat{x} \in \widehat{X} : \ \overline{q}(\widehat{x}) < \widehat{f}^*(\widehat{x})\}, \quad \widehat{\mathcal{X}}_> := \{\widehat{x} \in \widehat{X} : \ \underline{q}(\widehat{x}) > \widehat{f}^*(\widehat{x})\}, \\ &\widehat{\mathcal{X}}_= := \{\widehat{x} \in \widehat{X} : \ \underline{q}(\widehat{x}) \le \widehat{f}^*(\widehat{x}) \le \overline{q}(\widehat{x})\}, \quad \widehat{\mathcal{X}}_\ge := \widehat{\mathcal{X}}_> \cup \widehat{\mathcal{X}}_=, \ \ \widehat{\mathcal{X}}_\le := \widehat{\mathcal{X}}_< \cup \widehat{\mathcal{X}}_=. \end{aligned} \tag{EC.3}$$

Proof of Proposition 4. For the first case we prove by constructing an optimal policy, and for the second case we prove by contradiction.

If condition (I) is satisfied, then an optimal policy can be constructed by the following algorithm.

`Step 1.` Define $y_k^* \leftarrow \overline{q}(\widehat{x}_k)$, $\forall k \in [K]$. By (I), we know that y_k^* satisfies

$$\underline{q}(\widehat{x}_j) - y_k^* \le \|\widehat{x}_j - \widehat{x}_k\|, \qquad \forall j, k \in [K]. \tag{EC.4}$$

Note that this also implies $\underline{q}(\widehat{x}_j) \le y_j^*$ for all j by setting $k = j$.

`Step 2.` Choose $k_1 \in \arg\min_k y_k^*$. For any $k \ne k_1$, denote $y_k^{(1)} = y_{k_1}^* + \|\widehat{x}_k - \widehat{x}_{k_1}\|$. Then for all j, k,

$$\underline{q}(\widehat{x}_j) - y_k^{(1)} = \underline{q}(\widehat{x}_j) - y_{k_1}^* - \|\widehat{x}_k - \widehat{x}_{k_1}\| \le \|\widehat{x}_j - \widehat{x}_{k_1}\| - \|\widehat{x}_k - \widehat{x}_{k_1}\| \le \|\widehat{x}_j - \widehat{x}_k\|.$$

This means that if we reassign values to $y_k^* \leftarrow y_k^* \wedge y_k^{(1)}$, $\forall k \ne k_1$, then (EC.4) would still hold. Note that since $y_k^{(1)} \ge y_{k_1}^*$, after reassignment $y_{k_1}^*$ is still the smallest among all y_k^*.

`Step 3.` Choose $k_2 \in \arg\min_{k \ne k_1} y_k^*$. For any $k \notin \{k_1, k_2\}$, denote

$$y_k^{(2)} = y_{k_2}^* + \|\widehat{x}_k - \widehat{x}_{k_2}\|,$$

and reassign $y_k^* \leftarrow y_k^* \wedge y_k^{(2)}$. Same as `Step 2`, (EC.4) still holds, and $y_{k_1}^* \le y_{k_2}^* \le y_k^*$ for all $k \notin \{k_1, k_2\}$.

`Step 4.` Repeat `Step 3`. Eventually we would have $y_{k_1}^* \le y_{k_2}^* \le \cdots \le y_{k_K}^*$, with (EC.4) still holds. Now for every $i < j$, we have

$$0 \le y_{k_j}^* - y_{k_i}^* \le y_{k_j}^{(i)} - y_{k_i}^* = \|\widehat{x}_{k_j} - \widehat{x}_{k_i}\|.$$

Therefore, define $\widehat{f}(\widehat{x}_k) := y_k^*$, then $\widehat{f}$ is a 1-Lipschitz function. In the above process y_k^* is decreasing its value, so $y_k^* \le \overline{q}(\widehat{x}_k)$ which is its initial value, and (EC.4) ensures $y_k^* \ge \underline{q}(\widehat{x}_k)$ by setting $j = k$.

Since $\underline{q} \le \widehat{f} \le \overline{q}$, $\widehat{f}$ is a conditional $\frac{b}{b+h}$-quantile, with $\|\widehat{f}\|_{\mathrm{Lip}} \le 1$. Then it must be a minimizer of ($\widehat{\mathrm{R}}$) for any $\rho \ge 0$, because it minimizes both terms. Any other optimal policy $\widehat{f}^*$ must also be a 1-Lipschitz quantile function to reach this minimum value. This completes the proof for the first part.

To see the second part, if condition (I) is not satisfied, then we claim that the optimizer $\widehat{f}^*$ must have Lipschitz constant $\|\widehat{f}^*\|_{\mathrm{Lip}} = L \ge 1$. Indeed, if $L < 1$, we can always adjust the value of $\widehat{f}^*$ to reduce costs in the second

term of $v_{\hat{R}}$ without paying more cost in the first term. So, the only possibility that $\|\widehat{f}^*\|_{\mathrm{Lip}} < 1$ is that the second term is already optimized, that is $\underline{q} \le \widehat{f}^* \le \overline{q}$. However, this would imply (I) holds, which is a contradiction.

Now we partition $\widehat{\mathcal{X}}$ according to (EC.3). First we fix $\widehat{x}_k \in \widehat{\mathcal{X}}_>$. Indeed, there must be $\widehat{x}_{j_1} \in \widehat{\mathcal{X}} \setminus \{\widehat{x}_k\}$ such that $\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_{j_1}) = L\|\widehat{x}_k - \widehat{x}_{j_1}\|$, otherwise we can increase the value of $\widehat{f}^*(\widehat{x}_k)$ and optimize the second term in ($\widehat{\mathrm{R}}$) without jeopardizing the first term. If $\widehat{x}_{j_1} \in \widehat{\mathcal{X}}_\le$, then the claim is proved. For the same reason, if $\widehat{x}_{j_1} \in \widehat{\mathcal{X}}_>$, then we can find $\widehat{x}_{j_2} \in \widehat{\mathcal{X}} \setminus \{\widehat{x}_k, \widehat{x}_{j_1}\}$ such that $\widehat{f}^*(\widehat{x}_{j_1}) - \widehat{f}^*(\widehat{x}_{j_2}) = L\|\widehat{x}_{j_1} - \widehat{x}_{j_2}\|$, thus $\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_{j_2}) = L\|\widehat{x}_k - \widehat{x}_{j_1}\| + L\|\widehat{x}_{j_1} - \widehat{x}_{j_2}\| \ge L\|\widehat{x}_k - \widehat{x}_{j_2}\|$, and here inequality sign must be equality because $\widehat{f}^*$ is L-Lipschitz. Note that this also shows that $(\widehat{x}_k, \widehat{f}^*(\widehat{x}_k)), (\widehat{x}_{j_1}, \widehat{f}^*(\widehat{x}_{j_1}))$ and $(\widehat{x}_{j_2}, \widehat{f}^*(\widehat{x}_{j_2}))$ are on the same straight line if $\|\cdot\| = \|\cdot\|_2$. If $\widehat{x}_{j_2} \in \widehat{\mathcal{X}}_\le$ then we finish the proof of the claim, otherwise $\widehat{x}_{j_3}$ can be found. Note that $\widehat{f}^*(\widehat{x}_k) > \widehat{f}^*(\widehat{x}_{j_1}) > \widehat{f}^*(\widehat{x}_{j_2}) > \dots$ is strictly decreasing, thus $\widehat{x}_k, \widehat{x}_{j_1}, \widehat{x}_{j_2} \dots$ are distinct. After finitely many steps, we must have $\widehat{x}_j \in \widehat{\mathcal{X}}_\le$ and $\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j) = L\|\widehat{x}_k - \widehat{x}_j\|$ before we run out of points. □

Now we study the worst-case distribution under a optimal robust policy.

Proposition EC.1 **(Worst-case Distribution)**. *Let $\widehat{f}^*$ be an in-sample optimal robust policy. Then there exists a worst-case distribution $\mathbb{P}^*$ of ($\widehat{\mathrm{P}}$) such that the following holds.*

(I) *If $\|\widehat{f}^*\|_{\mathrm{Lip}} \le 1$, then $\mathbb{P}^*$ perturbs $\widehat{\mathbb{P}}$ by moving $(\widehat{x}, \widehat{z})$ with $\widehat{z} \ge \widehat{f}^*(\widehat{x})$ toward $(\widehat{x}, z')$ for some $z' > \widehat{z}$.*

(II) *If $\|\widehat{f}^*\|_{\mathrm{Lip}} > 1$, then $\mathbb{P}^*$ perturbs $\widehat{\mathbb{P}}$ by moving each $(\widehat{x}, \widehat{z})$ with $\widehat{x} \in \widehat{\mathcal{X}}_>$ and $\widehat{z} \ge \widehat{f}^*(\widehat{x})$ toward $(x', \widehat{z})$ for some $x' \in \widehat{\mathcal{X}} \setminus \widehat{\mathcal{X}}_>$ and $\widehat{f}^*(\widehat{x}) - \widehat{f}^*(x') = \|\widehat{f}^*\|_{\mathrm{Lip}}\|\widehat{x} - x'\|$.*

In Figure EC.3, we have the identical setting as in Figure 4. Above the graph of f^* is a blue shadow region representing $\{(x, z) : z \ge f^*(x)\}$, and $\mathbb{P}^*$ moves the probability mass in this region when backorder costs more than holding. In the left figure $\|\widehat{f}^*\|_{\mathrm{Lip}} < 1$, so it is more cost-efficient to move along the direction of z. In the right figure $\|\widehat{f}^*\|_{\mathrm{Lip}} > 1$, and it is more cost-efficient to move along the direction of x; since the worst-case distribution is for the in-sample problem, it perturbs x from one empirical value to another.

Proof of Proposition EC.1. To ease the notation we use f to represent $\widehat{f}^*$ in this proof.

(I) If $\|f\|_{\mathrm{Lip}} \le 1$, then $\|\Psi_f\|_{\mathrm{Lip}} = b \vee h$. In this case we choose to transport Z instead of X. We define a transport map $T : \widehat{\mathcal{X}} \times \mathcal{Z} \to \widehat{\mathcal{X}} \times \mathcal{Z}$ by

$$T(x, z) := \begin{cases} (x, z + t) & z \ge f(x) \\ (x, z) & z < f(x) \end{cases}$$

for some t to be determined. Let $\mathbb{P} = T_\# \widehat{\mathbb{P}}$ be the push-forward of $\widehat{\mathbb{P}}$ via $T_\# \widehat{\mathbb{P}}$ defined by $\mathbb{P}[A] = \widehat{\mathbb{P}}[T^{-1}(A)]$ for every measurable set $A \subset \widehat{\mathcal{X}} \times \mathcal{Z}$, then

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}[\Psi_f] - \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f] &= \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f \circ T(X, Z)] - \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f(X, Z)] \\
&= \mathbb{E}_{\widehat{\mathbb{P}}}[\mathbf{1}_{\{Z \ge f(X)\}}(\Psi(f(X), Z + t) - \Psi(f(X), Z))] \\
&= \mathbb{E}_{\widehat{\mathbb{P}}}[\mathbf{1}_{\{Z \ge f(X)\}}(b(Z + t - f(X)) - b(Z - f(X)))] \\
&= bt\widehat{\mathbb{P}}[Z \ge f(X)] \\
&= b\mathbb{E}_{\widehat{\mathbb{P}}}[\|T(X, Z) - (X, Z)\|] \\
&\ge b\mathcal{W}_1(\mathbb{P}, \widehat{\mathbb{P}}).
\end{aligned}$$

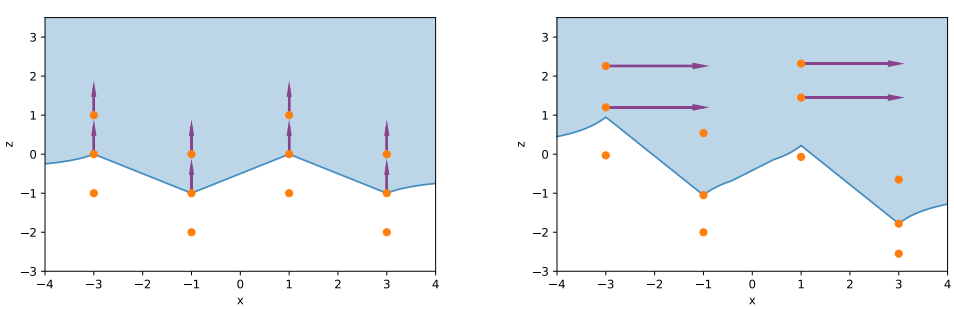


Figure EC.3 Transport map of worst-case distribution (purple arrows). When the optimal policy is 1-Lipschitz, worst-case distribution moves in z (left). Otherwise, worst-case distribution moves in x (right).

By choosing $t = \rho/\widehat{\mathbb{P}}[Z \geq f(X)]$, we have $\mathcal{W}(\mathbb{P}, \widehat{\mathbb{P}}) = \rho$, so $\mathbb{P}$ is feasible, and $\mathbb{E}_{\mathbb{P}}[\Psi_f] = \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f] + b\rho$ is a worst case distribution. Note that the denominator $\widehat{\mathbb{P}}[Z \geq f(X)]$ is never zero, otherwise $\widehat{\mathcal{X}} = \widehat{\mathcal{X}}_<$ and $\widehat{\mathcal{X}}_\geq = \emptyset$ which contradicts with Proposition 4.

(II) If $\|f\|_{\text{Lip}} = L > 1$, then $\|\Psi_f\|_{\text{Lip}} = (b \vee h)L$. In this case we choose to transport X instead of Z. In order to find a worst case distribution, we are interested in how far this $\widehat{x}_j \in \widehat{\mathcal{X}}_\leq$ can be from $\widehat{x}_k \in \widehat{\mathcal{X}}_>$ in Proposition 4 (II). For every k we define

$$\tau(k) \in \arg\min_{j \neq k} \left\{ \widehat{f}^*(\widehat{x}_j) : \widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j) = L\|\widehat{x}_k - \widehat{x}_j\| \right\}, \qquad \Delta(\widehat{x}_k) := \|\widehat{x}_k - \widehat{x}_{\tau(k)}\| = \frac{1}{L}(\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_{\tau(k)})). \tag{EC.5}$$

Intuitively, whenever $\widehat{x}_k \in \widehat{\mathcal{X}}_>$, $\tau(k)$ specifies a moving direction from $\widehat{x}_k$ to $\widehat{x}_{\tau(k)}$, and $\Delta(\widehat{x}_k)$ denotes the moving distance.

We define a transport map $T : \widehat{\mathcal{X}} \times \mathcal{Z} \to \widehat{\mathcal{X}} \times \mathcal{Z}$ by

$$T(\widehat{x}_k, z) := \begin{cases} (\widehat{x}_{\tau(k)}, z) & x \in \mathcal{X}_>, z \geq f(x) \\ (\widehat{x}_k, z) & x \in \mathcal{X}_\leq \text{ or } z < f(x) \end{cases}.$$

This implies that for every k,

$$\begin{aligned} \Psi_f \circ T(\widehat{x}_k, z) - \Psi_f(\widehat{x}_k, z) &= \mathbf{1}_{\{\widehat{x}_k \in \mathcal{X}_>, z \geq \widehat{f}^*(\widehat{x}_k)\}}(b(z - \widehat{f}^*(\widehat{x}_{\tau(k)})) - b(z - \widehat{f}^*(\widehat{x}_k))) \\ &= b\mathbf{1}_{\{\widehat{x}_k \in \mathcal{X}_>, z \geq \widehat{f}^*(\widehat{x}_k)\}} L\|\widehat{x}_k - \widehat{x}_{\tau_k}\| \\ &= bL\mathbf{1}_{\{\widehat{x}_k \in \mathcal{X}_>, z \geq \widehat{f}^*(\widehat{x}_k)\}} \Delta(\widehat{x}_k) \\ &= bL\|T(\widehat{x}_k, z) - (\widehat{x}_k, z)\| \end{aligned}$$

Let $\tilde{\mathbb{P}} = T_\# \widehat{\mathbb{P}}$, then

$$\mathbb{E}_{\tilde{\mathbb{P}}}[\Psi_f] - \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f] = \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f \circ T(X, Z) - \Psi_f(X, Z)] = bL\mathbb{E}_{\widehat{\mathbb{P}}}[\|T(\widehat{x}_k, z) - (\widehat{x}_k, z)\|] \geq bL\mathcal{W}_1(\tilde{\mathbb{P}}, \widehat{\mathbb{P}}).$$

We will show that

$$\mathbb{E}_{\tilde{\mathbb{P}}}[\Psi_f] - \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f] \geq bL\rho. \tag{EC.6}$$

If this is true, we can construct a feasible $\mathbb{P}$ by a convex combination of the form $\mathbb{P} = \alpha\tilde{\mathbb{P}} + (1-\alpha)\widehat{\mathbb{P}}$, such that

$$\mathbb{E}_{\mathbb{P}}[\Psi_f] - \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f] = bL\rho \geq bL\mathcal{W}(\mathbb{P}, \widehat{\mathbb{P}}),$$

so $\mathbb{P}$ is a feasible worst case distribution. It is not hard to see that $\mathbb{P}$ is exactly ρ away from $\widehat{\mathbb{P}}$.

Recall Δ is defined in (EC.5). To prove (EC.6), we construct another solution $f^\varepsilon(\widehat{x}_k) := \widehat{f}^*(\widehat{x}_k) - \varepsilon\Delta(\widehat{x}_k)$ which means "ordering less" than $\widehat{f}^*$. We claim when ε is small, $\|f^\varepsilon\|_{\text{Lip}} = L - \varepsilon$. To see why this is true, we can consider only the pairs of points $\widehat{x}_k$ and $\widehat{x}_j$ with $\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j) = L\|\widehat{x}_k - \widehat{x}_j\|$ since ε can be chosen sufficiently small. In this situation,

$$f^\varepsilon(\widehat{x}_k) - f^\varepsilon(\widehat{x}_j) = \widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j) - \varepsilon(\Delta(\widehat{x}_k) - \Delta(\widehat{x}_j)) = \widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j) - \frac{\varepsilon}{L}\left((\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_{\tau(k)})) - (\widehat{f}^*(\widehat{x}_j) - \widehat{f}^*(\widehat{x}_{\tau(j)}))\right).$$

It can be seen that $\widehat{f}^*(\widehat{x}_{\tau(k)}) \leq \widehat{f}^*(\widehat{x}_{\tau(j)})$ by the minimality of $\tau(k)$ (see the last paragraph in the proof of Proposition 4), hence

$$f^\varepsilon(\widehat{x}_k) - f^\varepsilon(\widehat{x}_j) \leq \widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j) - \frac{\varepsilon}{L}(\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j)) = \left(1 - \frac{\varepsilon}{L}\right)(\widehat{f}^*(\widehat{x}_k) - \widehat{f}^*(\widehat{x}_j)) = (L - \varepsilon)\|\widehat{x}_k - \widehat{x}_j\|.$$

Therefore f^ε is $(L-\varepsilon)$-Lipschitz.

Since f minimizes $v_{\hat{R}}$, we have

$$\begin{aligned} (b \vee h)L\rho + \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f] &\leq (b \vee h)(L - \varepsilon)\rho + \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_{f^\varepsilon}] \\ (b \vee h)\varepsilon\rho &\leq \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_{f^\varepsilon} - \Psi_f]. \end{aligned}$$

Here, "ordering less" means we need to pay more backorder cost and less holding cost, so

$$\Psi(f^\varepsilon(\widehat{x}_k), z) - \Psi(f(\widehat{x}_k), z) = \begin{cases} b\varepsilon\Delta(\widehat{x}_k), & \widehat{f}^*(\widehat{x}_k) \leq z, \\ -h\varepsilon\Delta(\widehat{x}_k), & \widehat{f}^*(\widehat{x}_k) > z. \end{cases}$$

Here we choose ε small such that $\widehat{f}^*(\widehat{x}_k) > z$ implies $\widehat{f}^*(\widehat{x}_k) > z + \varepsilon\Delta(\widehat{x}_k)$ for every $(\widehat{x}_k, z) \in \operatorname{supp}\widehat{\mathbb{P}}$. Now take the conditional expectation,

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}[\Psi(f^{\varepsilon}(X), Z) - \Psi(f(X), Z)|X = \widehat{x}_k] = \varepsilon\Delta(\widehat{x}_k)\left(b\widehat{\mathbb{P}}[Z \geq \widehat{f}^*(\widehat{x}_k)|X = \widehat{x}_k] - h\widehat{\mathbb{P}}[Z < \widehat{f}^*(\widehat{x}_k)|X = \widehat{x}_k]\right).$$

If $\widehat{x}_k \in \mathcal{X}_{\leq}$, $\widehat{f}^*(\widehat{x}_k)$ is no less than the conditional quantile, so the above will be nonpositive. Thus

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}[\Psi(f^{\varepsilon}(X), Z) - \Psi(f(X), Z)|X = \widehat{x}_k] \leq \begin{cases} b\varepsilon\Delta(\widehat{x}_k)\widehat{\mathbb{P}}[Z \geq \widehat{f}^*(\widehat{x}_k)|X = \widehat{x}_k], & \widehat{x}_k \in \mathcal{X}_{>}, \\ 0, & \widehat{x}_k \in \mathcal{X}_{\leq}. \end{cases}$$

Finally, take expectation in X, we have

$$(b \vee h)\varepsilon\rho \leq \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_{f^{\varepsilon}} - \Psi_f] \leq b\varepsilon\mathbb{E}_{\widehat{\mathbb{P}}}[\Delta(X)\mathbf{1}_{\{X \in \mathcal{X}_{>}, Z \geq f(X)\}}].$$

In particular, $\mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_f \circ T - \Psi_f] = bL\mathbb{E}_{\widehat{\mathbb{P}}}[\Delta(X)\mathbf{1}_{\{X \in \mathcal{X}_{>}, Z \geq f(X)\}}] \geq bL\rho$, which proves (EC.6). □

Remark EC.1. Similarly, when $b < h$, the transport map should move $X \in \widehat{\mathcal{X}}_{<}$ in $\{Z \leq \widehat{f}^*(X)\}$ when $L > 1$, and should decrease Z in $\{Z \leq \widehat{f}^*(X)\}$ when $L \leq 1$. Considering the demand must be nonnegative, in the case $L \leq 1$ we should use the transport map

$$T(x, z) := \begin{cases} (x, (z - t)_+) & z \geq \widehat{f}^*(x) \\ (x, z) & z \leq \widehat{f}^*(x) \end{cases}.$$

for some $t \geq 0$ if $\rho \leq \mathbb{E}_{\widehat{\mathbb{P}}}[Z\mathbf{1}\{Z < \widehat{f}^*(X)\}]$. From this proposition we can see that, when ρ is sufficiently small, the worst case distribution $\mathbb{P}$ is still supported in $\widehat{\mathcal{X}} \times \mathcal{Z}$.

Essentially the only place where $b \geq h$ really matters is the proof of (EC.2), where we send $z \to \infty$. When $b < h$, one would send $z \to -\infty$ instead, which would be absurd because $z \geq 0$. However, if we start with $\mathcal{Z} = \mathbb{R}$ instead of $\mathbb{R}_+$ in ($\widehat{\mathsf{P}}$):

$$v_{\hat{P}} = \min_{\widehat{f}:\widehat{\mathcal{X}} \to \mathbb{R}} \sup_{\mathbb{P} \in \mathcal{P}_1(\widehat{\mathcal{X}} \times \mathbb{R})} \left\{\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi_{\widehat{f}}(X, Z)] : \mathcal{W}(\mathbb{P}, \widehat{\mathbb{P}}) \leq \rho\right\},$$

we would have the same worst case distribution supported over $\{Z \geq 0\}$, so the value of $v_{\hat{P}}$ is going to be the same. Since the proof of $v_P = v_D = v_{\hat{D}} = v_{\hat{P}}$ in Theorem 1 doesn't rely on whether $\mathcal{Z} = \mathbb{R}$ or $\mathbb{R}_+$ because of the property (iii) in the Lemma 1, $\mathcal{Z}$ in (P), (D), ($\widehat{\mathsf{P}}$), ($\widehat{\mathsf{D}}$) can all be replaced by $\mathbb{R}$, and thus we can send $z \to -\infty$ to fix the proof for (EC.2).

Appendix E: Additional Numerical Results for Section 5.1

In this section, we provide additional numerical results for robustness check.

Methods	Hyperparameters
`Shapley`	radius and norm scaling parameter
`KO`	Bandwidth
k`NN`	No. of neighbours
`ERM2` (ℓ^1/ℓ^2)	Regularization parameter
`RandForest`	
`StochOptForest (apx-soln)`	None
`StochOptForest (apx-risk)`	

Table EC.1 Hyperparameters used in the methods

In Figure EC.4, we consider various noise distributions while fixing $n = 100$, and $d = 5000$. The model setup is the same as in Section 5.1 and the results are presented in Figures EC.4. We consider three different noise distributions for ε as considered by Zhu et al. (2012): (i) the Laplace distribution; (ii) the student-t distribution with 3 degrees of freedom; and (iii) Gaussian mixture $0.8\mathcal{N}(0, 1) + 0.2\mathcal{N}(0, 9)$. It can be seen that the observations that we made consistently hold for all three noise distributions.

In Figure EC.5, we consider a finer and larger grid for h while fixing $h + b = 1$ to study the impact on the extremity of fractiles. Need to notice that the performance is not symmetric since the distribution of the non-linear model is asymmetric. We found that the performance of `Shapley` tends to be better than other non-parametric

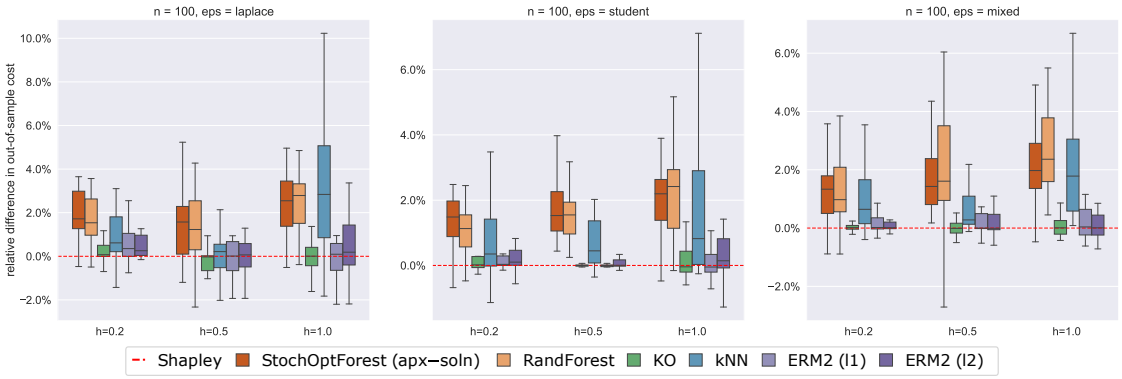


Figure EC.4 Boxplots of the out-of-sample performance for models with different noises

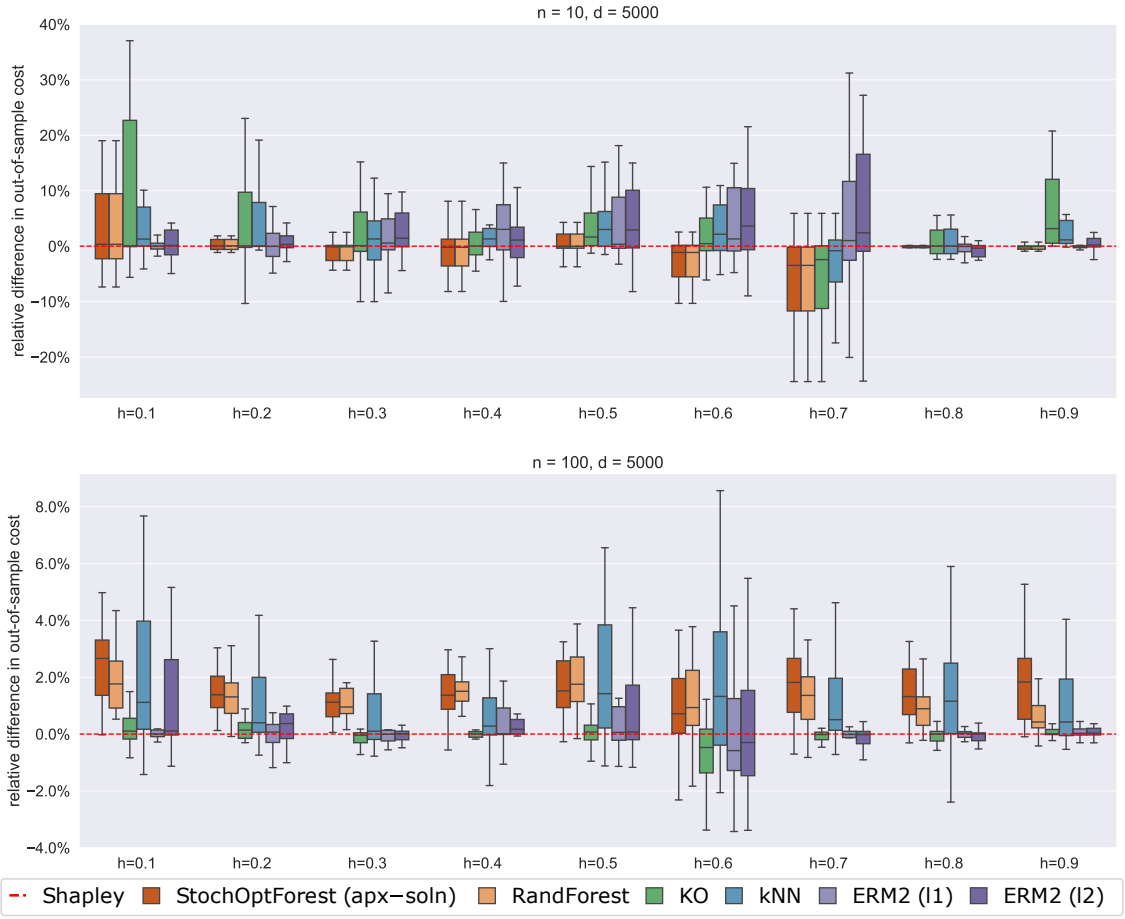


Figure EC.5 Boxplots of the out-of-sample performance for models with different h/b-ratios

benchmarks (KO, kNN and forest-based methods) when the fractile becomes extreme, which demonstrate the advantages of a robust approach.

In Figure EC.6, to study the impact of sparsity, we consider a sparse model as in Zhu et al. (2012) where $\beta_0 = (2, -2, -1, 1, 0, \ldots, 0)^\top / \sqrt{10}$ is a d-dimensional vector, $d = 5000$, and the covariate $X \in \mathbb{R}^d$ is a multivariate Gaussian random variable with mean zero and covariance matrix $(\sigma_{ij})_{d\times d}$ with $\sigma_{ij} = 0.5^{|i-j|}$, and ε is the standard Gaussian distribution. We consider $n = 10, 100$ only because 1000 seems too large for this sparse model. Under this case, our Shapley policy still maintains competitive performance, although its advantage is no longer dominant.

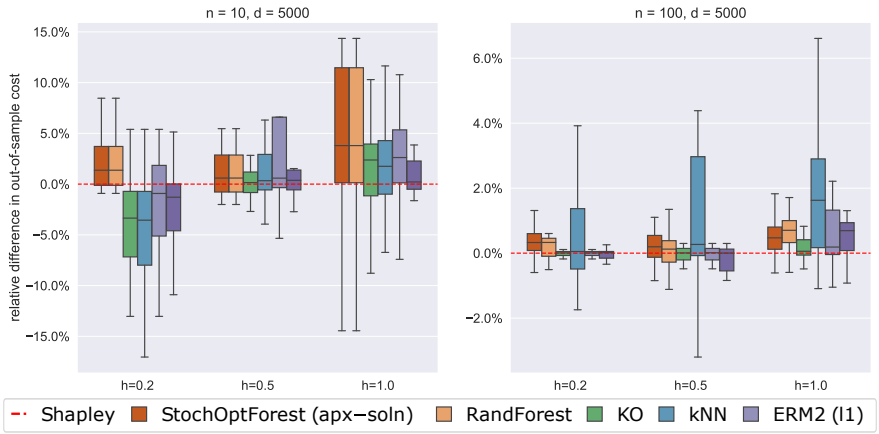


Figure EC.6 Differences in the out-of-sample performance between `Shapley` and other benchmarks in sparse case