Optimal Robust Policy for Feature-Based Newsvendor

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

Jincheng Yang, Luhao Zhang  
Department of Mathematics, The University of Texas at Austin, jincheng, luhaozhang@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. Unlike existing works that restrict the policies to some parametric class which may suffer from sub-optimality (such as affine class) or lack of interpretability (such as neural networks), we aim to optimize over all measurable functions of features. In this case, the classical empirical risk minimization yields a policy that are not well-defined on unseen features. To avoid such degeneracy, we consider a 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 features 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 real data and demonstrate its superior empirical performance.

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

History: First submitted on March 18, 2021.

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 economic—are available prior to the decision-making and reveals partial information on the product demand. The feature information reduces the uncertainty and facilitates the decision-maker to make customized ordering decisions under each individual feature. For example, in E-commerce, ordering decisions are made for a large number of customers with different demographics, and personalized decisions help to achieve good performance over the entire population. It is crucial to involve the feature information in decision-making, as 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). In this work, we are interested in the policy optimization problem that the decision-maker faces, which seeks a policy (a.k.a. decision rule) that outputs an ordering decision for every feature so as to minimize the expected cost over a distribution of features. We refer to 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 be equal to the conditional critical fractile of the demand distribution under each feature. Unfortunately, many real-world problems comprise a potentially large set of features that historical data cannot exhaust, thus the true underlying conditional distribution of the demand under a new feature 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 would lead to a pathological policy that is defined only on historical features but can take arbitrary values on unseen features; see more detailed discussion in Section 2. This necessitates the need for other approaches such as regularization or robust formulations to make the decisions generalize to unseen features.

Computationally, the policy optimization problem involves a challenging functional optimization over the space of policies, which is often infinite-dimensional. To mitigate the computational issue, one typical strategy is to confine the search space to a parametric policy class. For example, Ban and Rudin (2019, ERM) studied affine policies, which can be efficiently solved using convex optimization. The
an affine class is often restrictive, though, because the true optimal policy may not be linear in covariates at all, and numerical experiments in Ban and Rudin (2019) showed that affine policies are outperformed by their proposed kernel optimization method in the same paper. One possible remedy is to lift the policy class onto a higher-dimensional space by considering nonlinear transformations of covariates (basis functions), thereby one can enlarge the policy search space to an arbitrarily complex class with coefficients affinely dependent on the basis functions. 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 exhibit nice empirical performance, but are often hard to interpret and data-demanding. As such, the following question remains open: can we find an optimal policy from a non-restrictive class while maintaining computational efficiency and interpretability?

To answer this question, we consider a policy optimization framework that optimizes over all policies that are measurable functions of the feature, without any parametrization that could possibly be restrictive. This distinguishes our model from all existing works. To resolve the above-mentioned pathological issue in the empirical risk minimization, we consider a Wasserstein distributionally robust formulation (Kuhn et al. 2019) that hedges against demand and feature data uncertainty. This leads to an adjustable robust optimization (Yanikoglu et al. 2019), whose optimal solutions are generally unknown except for a few notable cases. Perhaps surprisingly, we identify a new class of policies that are proven to be optimal and can be computed efficiently. More specifically,

(I) We show that to solve the proposed infinite-dimensional robust policy optimization, it suffices to first solve a finite-dimensional in-sample robust policy optimization in which the policy is learned only on observed in-sample features, and then extend the in-sample optimal robust policy to unseen features such that it is optimal to the original robust policy optimization (Theorem 1). The extended policy, termed as Shapley policy, prescribes the decision on unseen features by computing the pure equilibrium of a matrix whose elements are pairwise weighted averages of the in-sample policy values on the historical features. This provides a new class of optimal policies for adjustable optimization that may be of independent interest.

(II) We further show that the in-sample robust policy optimization can be solved by linear programming (Theorem 2). In addition, the optimal robust policy can be nicely 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. We compare the out-of-sample cost of the Shapley policy with various benchmarks using real data and demonstrate its superior empirical performance.

![Figure 1](image)

**Figure 1** Illustration for conditional median regression, where training data are generated from a two-dimensional continuous distribution of $(x, z)$. Our proposed Shapley policy has a smaller variation (Lipschitz norm) than the linear interpolation of the conditional median of the empirical distribution.
Related Literature

On feature-based newsvendor. We consider a policy optimization formulation aiming at minimizing the expected newsvendor cost over a population. Once solved, it outputs an ordering decision for every new feature without explicitly estimating any intermediate value such as the conditional expected newsvendor cost or solving/tuning any additional optimization problems. As such, this approach is in contrast with two other approaches 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 and then optimizes for the ordering quantity (e.g., Toktay and Wein (2001), Zhu and Thonemann (2004)). 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 hard 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 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 feature and demand data. They demonstrated the nice generalization capability of this approach both theoretically and empirically. Nonetheless, as commented in Kallus and Mao (2020), 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 (2020) proposed a learning scheme to construct random forests and weights directly targeted for the decision optimization objective and illustrate 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, Ho and Hanasusanto 2019), trees and forests (Ban et al. 2019, Bertsimas and Kallus 2020, Kallus and Mao 2020), Dirichlet process (Hannah et al. 2010), local regression and classification (Bertsimas and Kallus 2020), smart prediction-then-optimization (Elmachtoub and Grigas 2021), robustness optimization and regularization (Zhu et al. 2021, Loke et al. 2020, Esteban-Pérez and Morales 2020), empirical residuals (Kannan et al. 2020a,b), etc. All these works aims at solving a conditional problem under every single feature separately, whereas our policy optimization marginalizes these conditional problems and provides an end-to-end framework.

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 2015, 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, Esfahani and Kuhn 2018, Gao and Kleywegt 2016, Lee et al. 2020, Chen and Xie 2020), etc. Except for See and Sim (2010), most of these works do not consider featureual information. In our analysis, we use the duality results for minimax Wasserstein distributionally robust optimization (Esfahani and Kuhn 2018, Blanchet and Murthy 2019, Gao and Kleywegt 2016) to obtain an equivalent reformulation of the inner 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 outer 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, Housni and Goyal 2018, Bertsimas et al. 2019, Georgiou et al. 2021), binary decision rules (Bertsimas and Georgiou 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, with the exception of Bertsimas et al. (2019), which considers dynamic decision-making with side information using affine decision rules. Different from 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, Georgiou et al. 2021). Our result complements the literature by providing a new class of provably optimal policies for adjustable robust optimization.

2. **Robust 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 $\Psi$ 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 on 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 $X = x$:

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

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

$$\mathbb{E} \left[ \inf_{y \in \mathcal{Z}} \mathbb{E}[\Psi(y, Z) | X] \right] = \inf_{f: \mathcal{X} \rightarrow \mathcal{Z}} \mathbb{E}[\Psi(f(X), Z)],$$

(1)

here: on the left-hand side, the outer expectation is taken over the marginal distribution of $X$ and it measures the optimal true risk, i.e., the expected cost when the distribution is the underlying true distribution and when $y$ is chosen as the true optimal ordering quantity for every feature; while on the right-hand side, the expectation is taken over the joint distribution of feature $X$ and demand $Z$, and 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 the optimal policy on the right-hand side takes a value $f^*(x)$ for any $x$ in the support of $X$, where $f^*(x)$ is 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 out-of-sample performance measure considered by Ban and Rudin (2019) among other literature on decision-making with feature information.
In practice, the true underlying distribution is often unknown. Instead, the decision maker often has historical data at disposal. Suppose they formulate an empirical distribution of the form

$$\hat{\mathbb{P}} = \frac{1}{n} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \delta(\hat{x}_k, z_k),$$

where $\hat{x}_k$’s are distinct features, $k = 1, \ldots, K$ and $n = \sum_{k=1}^{K} n_k$. A conventional wisdom is to consider the empirical risk minimization by replacing the true distribution with the empirical distribution $\hat{\mathbb{P}}$

$$\inf_{f: X \rightarrow Z} \mathbb{E}_{(X,Z) \sim \hat{\mathbb{P}}} [\Psi(f(X), Z)].$$

Unfortunately, this would yield a degenerate solution that is only defined on the set of observed features $\hat{\mathcal{X}} := \{\hat{x}_k : k \in [K]\}$, but can take arbitrary values elsewhere. Here $[K]$ denotes the set $\{1,2,\ldots,K\}$. In addition, the optimal policy takes a value $\hat{f}(\hat{x})$ for any $\hat{x} \in \mathcal{X}$, where $\hat{f}(\hat{x})$ the the critical fractile of the empirical conditional distribution; particularly, if the historical samples are generated from some continuous underlying distribution, then with probability one we have $n_k = 1$ for all $k$ and $f(\hat{x}_k) = \hat{z}_k$, which could be far away from optimal.

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

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

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

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

where $\mathbb{F}$ is the set of all measurable functions on $\mathcal{X}$. Throughout the paper, we assume $b > 0$ and $0 \leq h \leq b$. We remark that when $b < h$, all results in the paper still hold for sufficiently small $\rho$ (see Remark 1 in Appendix C).

3. Policy Optimization

Problem (P) is an infinite-dimensional optimization whose main difficulty is that, we need not only to assign an ordering quantity for every observed feature in $\mathcal{X}$ but also to each unseen feature in $\mathcal{X}\setminus\hat{\mathcal{X}}$. In this section, we aim to derive an equivalent tractable reformulation for (P).

By restricting the support of the involved distributions on in-sample data $\hat{\mathcal{X}}$ only, we define the in-sample robust primal problem as

$$v_p := \min_{f \in \mathcal{F}} \sup_{\mathbb{P} \in \mathbb{P}_1(\mathcal{X} \times \mathcal{Z})} \left\{ \mathbb{E}_{(X,Z) \sim \mathbb{P}} [\Psi(\hat{f}(X), Z)] : \mathcal{W}(\mathbb{P}, \hat{\mathbb{P}}) \leq \rho \right\},$$

where in-sample here means instead of considering the set of all features, we restrict our attention to historical features only, which is a finite subset. By minimizing over $\mathcal{F} = \{\hat{f} : \mathcal{X} \rightarrow \mathcal{Z}\} = \mathbb{R}_+^K$, the set of
in-sample policies defined on $\mathcal{X}$, the in-sample problem is a finite-dimensional problem that turns out to be tractable.

Our key result is to show that to solve the primal problem (P), it suffices to first solve the in-sample problem ($\bar{P}$) to obtain an in-sample optimal robust policy $\hat{f}^*$ that maps each in-sample feature in $\mathcal{X}$ to an ordering quantity, and then extend $\hat{f}^*$ to the entire space $\mathcal{X}$, based on Shapley extension defined in the Lemma 1 below. In Section 3.1, we prove that such extension renders an optimal policy to the primal problem (P). Next, in Section 3.2, we derive a finite-dimensional linear programming reformulation for the in-sample problem ($\bar{P}$).

### 3.1. Shapley Policy and its Optimality

In this subsection, we show that the infinite-dimensional functional optimization (P) can be reduced to a finite-dimensional problem.

To begin with, applying the duality for Wasserstein distributionally robust optimization (Lemma 3 in Appendix A) on the inner maximization of (P) and ($\bar{P}$) yields their strong dual problems

$$
\begin{align}
\nu_D := \inf_{f \in \mathcal{F}} \min_{\lambda \geq 0} \left\{ \lambda \rho + \mathbb{E}_{(X,Z)} \left[ \sup_{z \in \mathcal{Z}, x \in \mathcal{X}} \left\{ \Psi(f(x), z) - \lambda \|x - \hat{X}\| - \lambda |z - \hat{Z}| \right\} \right] \right\}, \\
\nu_D := \min_{f \in \mathcal{F}, \lambda \geq 0} \left\{ \lambda \rho + \mathbb{E}_{(X,Z)} \left[ \sup_{z \in \mathcal{Z}, x \in \mathcal{X}} \left\{ \Psi(f(x), z) - \lambda \|x - \hat{X}\| - \lambda |z - \hat{Z}| \right\} \right] \right\}.
\end{align}
$$

(D)

(D)

Observe that ($\hat{D}$) remains unchanged if one replaces the minimization over $\hat{f} \in \hat{\mathcal{F}}$ with minimization over $f \in \mathcal{F}$ because the objective value does not depend on the policy value outside $\mathcal{X}$. 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 $\nu_D \geq \nu_D$ because the supremum in (D) is taken over a larger set. To show the other direction, it suffices to show that the minimizer $\hat{f}^*$ of ($\hat{D}$) admits an extension $f^*$ such that for every $(X, Z)$, $\hat{f}^*$ is taken in the support of $\hat{P}$,

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

(3)

To this end, we establish the following key lemma.

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

$$
\begin{align}
f(x) := \min_{1 \leq k \leq K} \max_{1 \leq j \leq K} A_{jk}(x) = \max_{1 \leq j \leq K} \min_{1 \leq k \leq K} A_{jk}(x), \quad \forall x \in \mathcal{X},
\end{align}
$$

where $A_{jk}(x) := \frac{\|x - \hat{x}_k\|}{\|x - \hat{x}_j\| + \|x - \hat{x}_k\|} \hat{f}(\hat{x}_j) + \frac{\|x - \hat{x}_j\|}{\|x - \hat{x}_j\| + \|x - \hat{x}_k\|} \hat{f}(\hat{x}_k),
$$

(S)

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

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

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

$$
\sup_{x \in \mathcal{X}} \left\{ \Phi(f(x)) - \|x - \hat{x}_k\| \right\} \leq \max_{j=1,...,K} \left\{ \Phi(\hat{f}(\hat{x}_j)) - \|\hat{x}_j - \hat{x}_k\| \right\}.
$$

(4)

(iii) [Boundedness] $\min_{k \in [K]} \hat{f}(\hat{x}_k) \leq f \leq \max_{k \in [K]} \hat{f}(\hat{x}_k)$.

(iv) [Lipschitzness] The Lipschitz norm of $f$, denoted as $\|f\|_{1,\text{Lip}}$, is upper bounded by

$$
\max_{j \neq k} \frac{\|f(\hat{x}_j) - \hat{f}(\hat{x}_k)\|}{\|\hat{x}_j - \hat{x}_k\|}.
$$
For any \( x \) on the line segment connecting \( \hat{x}_j \) and \( \hat{x}_k \), \( A_{jk}(x) \) is simply a linear interpolation, elsewhere \( A_{jk}(x) \) is a distance-dependent weighted average, which is equivalent to 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 4 in Appendix A), thus we call it the Shapley policy. The first property states that \( f \) and \( \hat{f} \) coincide on in-sample data, thus \( f \) is indeed an extension. The second property implies (3) 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 top 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 \( \hat{x}_1 \) and \( \hat{x}_2 \), the interpolation is linear, and is curved elsewhere. As \( |x| \to \infty \), the policy converges to a “non-informative” ordering quantity \((f^∗(\hat{x}_1) + f^∗(\hat{x}_2))/2\) which, intuitively, means that the observed features provides little guidance on a faraway new feature \( x \) and thus the policy simply takes the average of the two in-sample policy values. When \( K = 3 \), for each pair of three observed features \( \hat{x}_1, \hat{x}_2, \hat{x}_3 \), 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 \( \hat{x}_k \), \( f^∗(x) \) is close to \( \hat{f}^∗(\hat{x}_k) \). When \( x \) is away from all observed features, \( f^∗(x) \) converges to \( \frac{1}{3} \left( \min_{k \in [K]} f^∗(\hat{x}_k) + \max_{k \in [K]} f^∗(\hat{x}_k) \right) \). The case of two-dimensional feature space \( \mathcal{X} = \mathbb{R}^2 \) is similar, as shown in Figure 6 in the Appendix A.

![Figure 2](image)

By applying Lemma 1 with \( \Phi(y) = \frac{1}{\lambda} \sup_{z \in \mathbb{Z}} \{ \Psi(y, z) − \lambda|z| \} \) (the degenerate case \( \lambda = 0 \) is trivial) and \( \hat{f} = \hat{f}^∗ \), we can prove (3) and thereby the following result is immediate.

**Theorem 1 (Dimension Reduction).** The dual problem \((D)\) is equivalent to the in-sample dual problem \((\hat{D})\) in the following sense: \( v_D = v_{\hat{D}} \), and every optimal policy \( f^∗ \in \hat{F} \) of \((\hat{D})\) can be extended to an optimal policy of \((D)\) via the Shapley extension \((\hat{S})\).

Theorem 1 shows that to solve \((D)\) (or \((P)\)), it suffices to solve the in-sample problem \((\hat{D})\) (or \((\hat{P})\)) and then extend the optimal in-sample policy to the entire space using \((\hat{S})\).

**3.1.1. Intuition Behind the Shapley Policy** We here provide some intuition to explain what is so special about the Shapley policy to make it robust optimal, which also sheds light on the proof of Theorem 1.

As discussed at the beginning of this section, the solution to the problem \((P)\) is related to the solution to the problem \((\hat{P})\). On the one hand, since we restrict the uncertainty set in \((\hat{P})\), 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.
\( \hat{f} = f|_{\hat{X}} \) in problem (\( \hat{P} \)). On the other hand, if for any in-sample policy \( \hat{f} : \hat{X} \to D \) we can find an extended policy \( f : X \to D \) such that a worst case distribution is guaranteed to be supported on \( \hat{X} \), then \( f \) has the same worst-case cost in (\( P \)) as \( \hat{f} \) in (\( \hat{P} \)), which makes \( f \) the optimal extension, thus the two problems become equivalent. Consider the extension defined by minimizing the absolute slope

\[
\hat{f}(x) = \arg \min_{y \in \mathbb{R}} \left\{ \max_{k \in [K]} \left| \frac{\hat{f}(\hat{x}_k) - y}{||\hat{x}_k - x||} \right| \right\}.
\]

Below we show that it leads to a worst-case distribution supported on \( \hat{X} \) and thus is optimal.

Fix \( x \in X \setminus \hat{X} \). Denote \( y_k = \hat{f}(\hat{x}_k) \), \( y = f(x) \), \( d_k = ||\hat{x}_k - x|| \), and the slope of a secant line connecting \( x \) and \( \hat{x}_k \) 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 \( \hat{x}_j, \hat{z}_j \) to \( (x, z) \) for some \( j \in [K] \). If \( y_j \leq y \), then there exists some \( \delta \in [0, 1] \) such that \( y = \delta y_j + (1 - \delta)y_{k^+} \), because \( y_{k^+} = y + d_{k^+} \geq y \). Now we can propose another transport plan, which instead moves \( \delta \) fraction of the mass on \( (\hat{x}_j, \hat{z}_j) \) to \( (x, z) \) and \( 1 - \delta \) fraction of the mass on \( (\hat{x}_j, \hat{z}_j) \) to \( (\hat{x}_{k^+}, z) \). Then

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

\[
y = \delta y_j + (1 - \delta)y_{k^+} = \delta(y + d_j L_j) + (1 - \delta)(y + d_{k^+} L) \geq y - \delta d_j L + (1 - \delta)d_{k^+} L \Rightarrow (1 - \delta)d_{k^+} \leq \delta d_j.
\]

Hence, moving probability mass to \( \hat{X} \) always lead to a worse distribution than moving to \( X \), thus we prove the claim and the policy defined by (5) is optimal.

Next, we show geometrically in Figure 3 that (5) is indeed the Shapley extension (\( S \)). To visualize \( A_{jk}(x) \), we plot the \( d-y \) plane on which a point \( (d_k, y_k) \) means \( \hat{x}_k \) 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 from point \( k \)'th view, 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 \( \{(dk, y_k)\}_{k \in [K]} \). 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 (5). Furthermore, by denoting the inner maximum of (5) as \( L \), (5) can be solved by the following linear program

\[
\min_{L \geq 0, y \in \mathbb{R}} L \\
\text{subject to} \quad y - Ld_k \leq y_k \leq y + Ld_k, \quad \text{for all } k \in [K]
\]

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. Naive 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.

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 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.

### 3.2. Linear Programming Reformulation

The discussion in the previous subsection suggests a relationship between our robust formulation and Lipschitz regularization. In this subsection, we formalize this result, which further renders a linear programming reformulation for the in-sample problem (\( \hat{P} \)).

Consider the following Lipschitz regularization problem defined as

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

where we denote by \( a_1 \vee a_2 \) the maximum between \( a_1 \) and \( a_2 \), and by \( \|\cdot\|_{\mathcal{Lip}} \) 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 \( 1 \vee \|f\|_{\mathcal{Lip}} \)) 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 \( \mathcal{X} \). If we prohibit perturbing any data in \( z \)-direction, then the lower cut off \( 1 \) of the Lipschitz norm \( \|f\|_{\mathcal{Lip}} \) 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

\[
\min_{y \in \mathbb{R}} \mathbb{E}_{\hat{P}\hat{Z}} \left[ \Psi(y, \hat{Z}) \right],
\]

and the optimal policy equals the unconditional critical fractile of \( \hat{P}\hat{Z} \).

Let us also define the in-sample Lipschitz regularization problem

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

The following Lemma 2 enables us to further reformulate the problem as Lipschitz regularization. The proof is provided in Appendix B.
Lemma 2. Problems \( \hat{\mathcal{D}} \) and \( \hat{\mathcal{R}} \) are equivalent. Moreover, if \( \hat{f}^* \) is an optimal solution to \( \hat{\mathcal{R}} \), then the Shapley policy defined by \( \mathcal{S} \) is an optimal solution to \( \hat{\mathcal{R}} \).

In the literature, it is known that 1-Wasserstein distributionally robust optimization is upper bounded by Lipschitz regularization (Esfahani and Kuhn 2018, Shafieezadeh-Abadeh et al. 2019, Gao et al. 2017), and the two problems are equivalent under certain assumptions. That said, Lemma 2 is actually a bit 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 5 in Appendix B) is that there exists a robust optimal in-sample policy \( f^* \) with sufficiently small Lipschitz norm, which gives a direct restriction on the range of the dual multiplier \( \lambda \) and transforms \( \hat{\mathcal{D}} \) to \( \hat{\mathcal{R}} \).

From Lemma 2, by introducing auxiliary variables, we directly conclude the following.

Theorem 2 (Linear Programming Reformulation). Problem \( \hat{\mathcal{R}} \) is equivalent to

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

This linear program has \( N + K + 1 \) variables with \( K^2 + 2N + 1 \) constraints. The variable \( L \) stands for \( 1 \vee \|f\|_{\text{Lip}} \); and \( \psi_{ki} \) represents the cost of ordering quantity \( y_k \) when the demand is \( \hat{z}_{ki} \) and the feature is \( \hat{x}_k \). In practice, we may impose different norm scaling parameter \( \beta > 0 \) in \( \|\|x\|| = \beta \|x\| + |z| \) that balances between the uncertainty in the feature information and in the demand, in which case \( |y_j - y_k| \leq L \|\hat{x}_j - \hat{x}_k\| \) becomes \( |\hat{y}_k - \hat{z}_{ki}| \leq L \beta \|\hat{x}_j - \hat{x}_k\| \).

Thus far, combining all results in Sections 3.1 and 3.2, we have shown that problems \( \hat{\mathcal{P}} \), \( \hat{\mathcal{D}} \) and \( \hat{\mathcal{R}} \) share an in-sample optimal robust policy \( f^* \in \mathcal{F} \), which can be extended to an optimal robust policy \( \hat{f}^* \in \hat{\mathcal{F}} \) for problems \( \mathcal{P} \), \( \mathcal{D} \) and \( \mathcal{R} \).

3.3. Optimal Robust Policy and Worst-case Scenarios

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 C.

Define the empirical conditional critical fractiles for \( \hat{x} \in \hat{X} \) as

\[
\bar{q}(\hat{x}) := \max \left\{ z \in \mathcal{Z} : \hat{P}[Z < z|X = \hat{x}] \leq \frac{b}{b + h} \right\}, \quad (7)
\]

\[
\underline{q}(\hat{x}) := \min \left\{ z \in \mathcal{Z} : \hat{P}[Z \leq z|X = \hat{x}] \geq \frac{b}{b + h} \right\}. \quad (8)
\]

It is easy to see that \( \underline{q}(\hat{x}) \leq \bar{q}(\hat{x}) \). We also define the subsets of historical features

\[
\hat{X}_\leq := \left\{ \hat{x} \in \hat{X} : \bar{q}(\hat{x}) < \hat{f}^*(\hat{x}) \right\}, \quad \hat{X}_\geq := \left\{ \hat{x} \in \hat{X} : \underline{q}(\hat{x}) > \hat{f}^*(\hat{x}) \right\}.
\]

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

Proposition 1 (Structure of the Optimal Policy).

(I) If \( q(\hat{x}_j) - \bar{q}(\hat{x}_k) \leq \|\hat{x}_j - \hat{x}_k\| \) for all \( 1 \leq j, k \leq K \), then \( \hat{f}^* \) is 1-Lipschitz and an empirical conditional critical fractile, i.e. \( q \leq \hat{f}^* \leq \bar{q} \). In this case, \( \hat{f}^* \) is optimal to \( \hat{\mathcal{R}} \) for any \( \rho \geq 0 \).
(II) Otherwise, \( \| f^* \|_{\text{Lip}} = L \geq 1 \). For every \( \tilde{x}_k \in \tilde{X}_\geq \), there exists \( \tilde{x}_j \in \tilde{X} \setminus \tilde{X}_> \), such that \( y_k - y_j = L \| \tilde{x}_k - \tilde{x}_j \| \). Similarly, for every \( \tilde{x}_k \in \tilde{X}_< \) there exists \( \tilde{x}_j \in \tilde{X} \setminus \tilde{X}_< \) such that \( y_j - y_k = L \| \tilde{x}_k - \tilde{x}_j \| \).

Proposition 1 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 \( \check{R} \). Otherwise, if the variation of the empirical conditional critical fractile is too large, i.e., a large value of \( |y_j - y_k|/\|\tilde{x}_j - \tilde{x}_k\| \) for some \( j \neq k \), then to reduce the variation of the policy, we would order more than the empirical critical fractile \( \bar{q}(\tilde{x}_k) \) at the cost of holding more, resulting in a set \( \tilde{X}_< \); or would order less than the empirical critical fractile \( q(\bar{x}_k) \) at the cost of back ordering more, resulting in a set \( \tilde{X}_\geq \).

![Figure 4](image)

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

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

Given the structure of the optimal policy, in Proposition 2 and Figure 7 in Appendix C, we investigate the worst-case distribution \( \mathbb{P}^* \), which sheds light on the (non-)conservativeness of our formulation.

4. Numerical Experiments

In this section, we compare our Shapley policy against several benchmarks, including: empirical risk minimization of affine policy with \( \ell^1 \) and \( \ell^2 \) regularization (ERM2 (\( \ell^1/\ell^2 \))) and kernel-weights optimization (K0) 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 criterion (StochOptForest (apx-soln/apx-risk)) in Kallus and Mao (2020).

We test the out-of-sample 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 (\( \ell^1/\ell^2 \))) 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 subtracting 3 reference categories). For distance-based methods (Shapley, kNN, K0), 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

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

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; \(\| \cdot \|_{\text{mod}_{i_2}}\) and \(\| \cdot \|_{\text{mod}_7}\) are metrics on the ring of integers modulo 12 and 7, where \(\|n\|_{\text{mod}_q} = \frac{1}{q} \min \{|r| : r \equiv n \mod q\}\), 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. If the model has hyperparameters, we use 5-fold cross-validation: first, we randomly shuffle the \(n\) samples and partition them into 5 groups; then for each group, we use the remaining 4 groups of samples for training and use this group for validation; finally, we pick the hyperparameters with the least average validation cost among the 5 groups. Once the hyperparameters are tuned, we use all \(n\) samples together to train the model again. The hyperparameters of these benchmarks are: the number of neighbors in \(k\)NN, the bandwidth in \(K_0\), the regularization coefficient in \(\text{ERM2} (\ell^1/\ell^2)\); and in our Shapley, the radius \(\rho\) and the norm scaling parameter \(\beta\) introduced after Theorem 2. The out-of-sample cost is defined as the average newsvendor cost of 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 \(b = 1\). We run 20 repeated experiments 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).

The average out-of-sample costs and their 95\% confidence intervals are shown in Table 1. 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, \(h/b = 0.2\).

<table>
<thead>
<tr>
<th>Model</th>
<th>Shapley</th>
<th>(K_0)</th>
<th>(k)NN</th>
<th>(\text{StochasticForest} (\text{apx-risk}))</th>
<th>(\text{StochasticForest} (\text{apx-so/l.Varn}))</th>
<th>(\text{RandForest})</th>
<th>(\text{ERM2} (\ell^2))</th>
<th>(\text{ERM2} (\ell^1))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(h = 0.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 20)</td>
<td>26.62±1.41</td>
<td>27.77±1.56</td>
<td>28.68±1.14</td>
<td>28.48±1.58</td>
<td>29.48±1.58</td>
<td>28.48±1.58</td>
<td>24.91±2.98</td>
<td>37.40±1.42</td>
</tr>
<tr>
<td>(n = 40)</td>
<td>23.96±1.25</td>
<td>25.31±0.95</td>
<td>27.23±0.87</td>
<td>27.72±0.88</td>
<td>27.73±0.88</td>
<td>27.73±0.88</td>
<td>30.05±1.52</td>
<td>33.91±1.99</td>
</tr>
<tr>
<td>(n = 100)</td>
<td>20.44±0.52</td>
<td>24.12±0.53</td>
<td>21.18±0.55</td>
<td>26.68±0.24</td>
<td>26.68±0.24</td>
<td>26.68±0.24</td>
<td>22.69±0.93</td>
<td>22.32±0.96</td>
</tr>
<tr>
<td>(h = 0.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 20)</td>
<td>37.30±1.07</td>
<td>37.08±1.58</td>
<td>40.09±1.02</td>
<td>39.85±1.06</td>
<td>39.85±1.06</td>
<td>39.85±1.06</td>
<td>51.78±3.16</td>
<td>43.48±2.02</td>
</tr>
<tr>
<td>(n = 40)</td>
<td>33.71±0.84</td>
<td>35.35±0.96</td>
<td>37.60±0.62</td>
<td>39.01±0.40</td>
<td>39.01±0.40</td>
<td>39.01±0.40</td>
<td>43.56±2.58</td>
<td>40.50±1.14</td>
</tr>
<tr>
<td>(n = 100)</td>
<td>30.43±0.52</td>
<td>32.95±0.57</td>
<td>30.58±0.46</td>
<td>38.70±0.23</td>
<td>38.68±0.23</td>
<td>38.68±0.23</td>
<td>30.72±0.79</td>
<td>30.57±0.72</td>
</tr>
<tr>
<td>(h = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 20)</td>
<td>45.21±1.11</td>
<td>45.43±1.31</td>
<td>48.21±1.49</td>
<td>46.85±0.60</td>
<td>46.85±0.60</td>
<td>46.85±0.60</td>
<td>58.59±4.29</td>
<td>50.20±2.48</td>
</tr>
<tr>
<td>(n = 40)</td>
<td>42.03±0.53</td>
<td>44.71±0.46</td>
<td>44.85±0.64</td>
<td>46.55±0.35</td>
<td>46.55±0.35</td>
<td>46.55±0.35</td>
<td>53.95±5.57</td>
<td>46.27±2.26</td>
</tr>
<tr>
<td>(n = 100)</td>
<td>38.31±0.31</td>
<td>42.52±0.60</td>
<td>38.60±0.63</td>
<td>46.26±0.26</td>
<td>46.26±0.26</td>
<td>46.26±0.26</td>
<td>38.88±0.96</td>
<td>39.69±0.83</td>
</tr>
</tbody>
</table>

Table 1 Average out-of-sample costs with 95\% confidence intervals

To better compare the instance-wise performance, we also draw 9 boxplots in Figure 5 under different choices of \(h\) and \(n\), which are generated by the differences of out-of-sample costs between our method and each of the benchmarks using the same training/testing data set; a positive number indicates that the benchmark is outperformed by Shapley. We have the following observations.

(I) Affine policies with regularization (\(\text{ERM2} (\ell^1/\ell^2)\)) performs relatively well only when \(n = 100\), and even in this case, it is outperformed by Shapley in terms of the median of out-of-sample cost. This demonstrates the advantage of using a larger policy class beyond affine.
Figure 5  
Boxplots of the differences in the out-of-sample performance between Shapley and other benchmarks

(II)  
kNN and KO are two distance-based non-parametric approaches, so is our Shapley policy. In terms of the median of 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 of the 9 cases, and the difference becomes more apparent when the sample size increases.

<table>
<thead>
<tr>
<th>Model</th>
<th>Average Training Time</th>
<th>Average Testing Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 20$</td>
<td>$n = 40$</td>
</tr>
<tr>
<td>KO</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
<tr>
<td>kNN</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
<tr>
<td>ERM2 ($\ell^1$)</td>
<td>0.0112</td>
<td>0.0153</td>
</tr>
<tr>
<td>ERM2 ($\ell^2$)</td>
<td>0.0363</td>
<td>0.0523</td>
</tr>
<tr>
<td>Shapley</td>
<td>0.0592</td>
<td>0.1272</td>
</tr>
<tr>
<td>RandForest</td>
<td>0.3618</td>
<td>1.7841</td>
</tr>
<tr>
<td>StochOptForest (apx-soln)</td>
<td>0.3762</td>
<td>1.8495</td>
</tr>
<tr>
<td>StochOptForest (apx-risk)</td>
<td>0.3640</td>
<td>1.8585</td>
</tr>
</tbody>
</table>

Table 2  
Average computational time (in seconds) per problem instance

We also report the average computational time of these methods in Table 2, in which 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 total 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 3 choices of holding cost (as it appears that the holding cost does not have an evident effect on the running time). We have the following observations.

(I)  
For the average training time, KO and kNN takes the least amount of time since they simply involve data sorting. ERM2 ($\ell^1/\ell^2$) involves solving linear programming ($\ell^1$) or convex quadratic programming ($\ell^2$), and the training time increases mildly in the training sample size. Shapley also involves solving linear programming 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) takes the longest time to train, and often longer in magnitudes.

(II) For the average testing time, affine policies (ERM2 ($\ell^1/\ell^2$)) takes the least time for testing and is independent of the training size by construction. The average testing time of K0 and kNN seems not affected much by the training sample size $n$ although their theoretical bounds are $O(n)$; this may be due to the parallel computing mechanism of the pandas.DataFrame library in Python. We find empirically that the testing time of Shapley is size-independent using our equivalent formulation (6), which is faster than computing the matrix saddle point naïvely. Forest-based methods (RandForest, StochOptForest) take the longest time that increases in the sample size.

5. Concluding Remarks

In this paper, we have developed an efficient approach to finding a robust optimal end-to-end policy for the feature-based newssoft, 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 newssoft cost. For the future work, it would be interesting to investigate its statistical properties and extend it to other adjustable robust optimization problems.

Appendix A: Proofs for Section 3.1

The following Lemma 3 is a direct consequence of the strong duality result in Wasserstein distributionally robust optimization (e.g., Gao and Kleywegt (2016), 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 $\Psi_f(x, z) = \Psi(f(x), z)$.

**Lemma 3.** For each $f \in \mathcal{F}$, the inner primal problem

$$v_f^p := \sup_{P \in P_1(X \times Z)} \left\{ \mathbb{E}_{(X,Z) \sim P} [\Psi_f(X, Z)] : \mathbb{W}(P, \tilde{P}) \leq \rho \right\},$$

is equal to the following inner dual problem,

$$v_f^d := \inf_{\lambda \geq 0} \left\{ \lambda \rho + \mathbb{E}_{(X,Z) \sim \tilde{P}} \left[ \sup_{(x,z) \in X \times Z} \{ \Psi_f(x, z) - \lambda(\|x - \tilde{X}\| + |z - \tilde{Z}|) \} \right] \right\}.$$

Note that $v_f^d = v_f^p$ 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 $y_k = f(\hat{x}_k)$, and we define

$$A_{kk}(x) := y_k, \quad k = 1, \ldots, K,$$

$$A_{jk}(x) := \frac{\|x - \hat{x}_j\|}{\|x - \hat{x}_j\| + \|x - \hat{x}_k\|} y_j + \frac{\|x - \hat{x}_k\|}{\|x - \hat{x}_j\| + \|x - \hat{x}_k\|} y_k, \quad j \neq k,$$

$$A^+(x) := \max_{1 \leq k \leq K} \max_{1 \leq j \leq K} A_{jk}(x), \quad A^-(x) := \max_{1 \leq k \leq K} \min_{1 \leq j \leq K} A_{jk}(x).$$

In Figure 6, 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 $X = \mathbb{R}^2$. 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.

We claim $A^+$ and $A^-$ both satisfy the four properties. First, we show they are indeed extensions. Fix $l = 1, \ldots, K$, then $A_{jj}(\hat{x}_l) = A_{jj}(\hat{x}_l) = y_l$ for every $j$. This implies

$$A^+(\hat{x}_j) = \min_{1 \leq k \leq K} \max_{1 \leq j \leq K} A_{jk}(\hat{x}_j) \leq \max_{1 \leq j \leq K} A_{jj}(\hat{x}_j) = y_l,$$

$$A^-(\hat{x}_j) = \max_{1 \leq k \leq K} \min_{1 \leq j \leq K} A_{jk}(\hat{x}_j) \geq \min_{1 \leq j \leq K} A_{jj}(\hat{x}_j) = y_l.$$
we first claim that thus if we denote are constant functions so they always satisfy the Lipschitz bound when clearly we have and however the minimax is always greater than or equal to the maximin, so in fact $A^\ast(\hat{x}_i) = A^- (\hat{x}_i) = y_l$, that is, $A^\ast$ 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 $\min \{ y_j, y_k \} \leq A_{jk}(x) \leq \max \{ y_j, y_k \}$. 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 $j \neq k$, fix $x, x' \in X$, and we denote

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

Then

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

$$= (y_j - y_k) \left( \frac{d_{xk} d_{x'j} + d_{xk} d_{x'k} - d_{xj} d_{x'k} - d_{xj} d_{x'k}}{(d_{xj} + d_{xk})(d_{x'j} + d_{x'k})} \right)$$

$$= (y_j - y_k) \left( \frac{d_{xk} d_{x'j} - d_{xk} d_{xj} + d_{xj} d_{x'k} - d_{xj} d_{x'k}}{(d_{xj} + d_{xk})(d_{x'j} + d_{x'k})} \right)$$

$$= (y_j - y_k) \left( \frac{d_{xk} d_{x'j} - d_{xj} d_{xk} + d_{xj} d_{x'k} - d_{xj} d_{x'k}}{(d_{xj} + d_{xk})(d_{x'j} + d_{x'k})} \right).$$

By triangular inequality,

$$|A_{jk}(x) - A_{jk}(x')| \leq |y_j - y_k| \left( \frac{d_{xk} d_{x'j} + d_{xj} d_{x'k}}{(d_{xj} + d_{xk})(d_{x'j} + d_{x'k})} \right)$$

$$\leq |y_j - y_k| \left( \frac{d_{x'j}}{d_{x'j} + d_{x'k}} \right)$$

$$\leq |y_j - y_k| \left( \frac{d_{x'k}}{d_{jk}} \right).$$

Thus if we denote $L := \max_{j \neq k} \frac{|y_j - y_k|}{||\hat{x}_j - \hat{x}_k||}$ 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 (4), which is to show that $y = A^\ast(x), A^- (x)$ satisfy the following condition for every $k$:

$$\Phi(y) - ||x - \hat{x}_k|| \leq \max_{j=1, \ldots, K} \left\{ \Phi(y_j) - ||\hat{x}_j - \hat{x}_k|| \right\} =: M_k.$$  

(Mk)

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

$$\Phi(A_{jk}(x)) - ||x - \hat{x}_k|| \leq \Phi \left( \frac{||x - \hat{x}_k||}{||x - \hat{x}_j|| + ||x - \hat{x}_k||} y_j + \frac{||x - \hat{x}_j||}{||x - \hat{x}_j|| + ||x - \hat{x}_k||} y_k \right) - ||x - \hat{x}_k||$$

$$\leq \frac{||x - \hat{x}_j||}{||x - \hat{x}_j|| + ||x - \hat{x}_k||} \Phi(y_j) + \frac{||x - \hat{x}_j||}{||x - \hat{x}_j|| + ||x - \hat{x}_k||} \Phi(y_k) - ||x - \hat{x}_k||.$$
Using definition of $M_k$ in (Mk),
\[ \Phi(y_j) \leq M_k + \|\hat{x}_j - \hat{x}_k\|, \quad \Phi(y_k) \leq M_k + \|\hat{x}_k - \hat{x}_k\| = M_k. \]

Plug into the above inequality,
\[ \Phi(A_{jk}(x)) - \|x - \hat{x}_k\| \leq M_k + \frac{\|x - \hat{x}_k\|}{\|x - \hat{x}_j\| + \|x - \hat{x}_k\|} \|\hat{x}_j - \hat{x}_k\| - \|x - \hat{x}_k\| = M_k + \|x - \hat{x}_k\| \left( \frac{\|\hat{x}_j - \hat{x}_k\|}{\|x - \hat{x}_j\| + \|x - \hat{x}_k\|} - 1 \right) \leq M_k. \]

In the last step we used the triangle inequality $\|\hat{x}_j - \hat{x}_k\| \leq \|x - \hat{x}_j\| + \|x - \hat{x}_k\|$. Fixing a $k$ in $1, \ldots, K$, then from the definition we can find $j_1, j_2$ such that $A_{j_1k}(x) \leq A^-(x) \leq A^+(x) \leq A_{j_2k}(x)$. Because both $y = A_{j_1k}(x)$ and $y = A_{j_2k}(x)$ satisfy (Mk), and $A^+, A^-$ are between $A_{j_1k}, A_{j_2k}$, we use the convexity of $\Phi$ again and conclude that $y = A^+(x), A^-(x)$ also satisfy (Mk), which implies (4).

\[ \square \]

\textit{Proof of Theorem 1.} To show the direction $v_D \geq v_D^*$, note that $v_D$ can be written with $f \in \mathcal{F}$ instead of $\hat{f} \in \hat{\mathcal{F}}$:
\[ v_D = \inf_{f \in \mathcal{F}} \inf_{\lambda \geq 0} \left\{ \lambda \rho + \mathbb{E}_\beta \left[ \sup_{x \in \mathcal{X}} \max_{z \in \mathcal{Z}} \{ \Psi(f(x), z) - \lambda \|x - \hat{X}\| - \lambda |z - \hat{Z}| \} \right] \right\} \]

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

To show $v_D \leq v_D^*$, we let $f = \operatorname{Ext}[\hat{f}]$ for a given $\hat{f} \in \hat{\mathcal{F}}$ and split into two cases: $\lambda > 0$ and $\lambda = 0$. If $\lambda > 0$, by Lemma 1 we know that $f$ satisfies (4). In particular, if we choose $\Phi(y) = \frac{1}{\lambda} \Psi(y, z) - |z - \hat{Z}|$, then
\[ \lambda \rho + \mathbb{E}_\beta \left[ \sup_{x \in \mathcal{X}} \max_{z \in \mathcal{Z}} \{ \Psi(f(x), z) - \lambda \|x - \hat{X}\| - \lambda |z - \hat{Z}| \} \right] \leq \lambda \rho + \mathbb{E}_\beta \left[ \sup_{x \in \mathcal{X}} \max_{z \in \mathcal{Z}} \{ \Psi(\hat{f}(x), z) - \lambda \|x - \hat{X}\| - \lambda |z - \hat{Z}| \} \right]. \]

If $\lambda = 0$, by Lemma 1 we know that the range of $f$ is $[\min \hat{f}, \max \hat{f}]$. Since $\Psi(\cdot, z)$ is convex, the supremum of $\Psi(f(\cdot), z)$ is attained at the extreme points, so
\[ \mathbb{E}_\beta \left[ \sup_{x \in \mathcal{X}} \max_{z \in \mathcal{Z}} \{ \Psi(f(x), z) \} \right] \leq \mathbb{E}_\beta \left[ \sup_{x \in \mathcal{X}} \max_{z \in \mathcal{Z}} \{ \Psi(\hat{f}(x), z) \} \right]. \]

Therefore, taking infimum in $\lambda \geq 0$ gives
\[ \inf_{\lambda \geq 0} \left\{ \lambda \rho + \mathbb{E}_\beta \left[ \sup_{x \in \mathcal{X}} \max_{z \in \mathcal{Z}} \{ \Psi(f(x), z) - \lambda \|x - \hat{X}\| - \lambda |z - \hat{Z}| \} \right] \right\} \leq \inf_{\lambda \geq 0} \left\{ \lambda \rho + \mathbb{E}_\beta \left[ \sup_{x \in \mathcal{X}} \max_{z \in \mathcal{Z}} \{ \Psi(\hat{f}(x), z) - \lambda \|x - \hat{X}\| - \lambda |z - \hat{Z}| \} \right] \right\}. \]

The left dominates $v_D$, so taking the inf over $\hat{f} \in \hat{\mathcal{F}}$ on the right gives $v_D \leq v_D^*$, which completes the proof of $v_D = v_D^*$. Note that the above also proves that if $f^*$ is a minimizer of $v_D^*$, then $f^* = \operatorname{Ext}[f^*]$ is a minimizer of $v_D$, also a minimizer of $v_D$ by Lemma 3.

Finally, existence of the saddle point is proved in Lemma 4 below.

\[ \square \]

\textbf{Lemma 4. With the same notation as in Proof of Lemma 1, $A^+ = A^-$.}

\textit{Proof.} We fix an $x$ and then omit the $x$ in $f_{jk}(x)$ to ease the notation. Denote the symmetric matrix $A = (f_{jk})_{jk}$, and we want to show that for this matrix, $\min_k \max_j f_{jk} = \max_j \min_k f_{jk}$, 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 $2 \times 2$ submatrix of $A$ has a saddle point.

For a general $2 \times 2$ matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we claim that if it has no saddle point then it has the \textit{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
\[ (a \land c) \lor (b \land d) \neq (a \lor b) \land (c \lor d). \]
Without loss of generality, assume $a$ is the smallest entry. Then the above inequality is simplified to

$$b \wedge d \neq b \wedge (c \lor d).$$

If $d \geq c$, then both sides equals to $b \wedge d$. If $d \geq b$, 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

$$B = \begin{pmatrix} A_{ik} & A_{jk} \\ A_{il} & A_{jl} \end{pmatrix}$$

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 $y_i, y_j, y_k$. Without loss of generality, assume $y_i \geq y_j, y_k \geq y_l$, and assume $y_i \geq y_k$ using the transpose symmetry. There are three possibilities:

(I) $y_i \geq y_k \geq y_j$.

(II) $y_i \geq y_k \geq y_j$.

(III) $y_i \geq y_j > y_k \geq y_l$.

In case (I), $A_{ik} \geq y_k \geq A_{jk}, A_{il} \leq y_l \leq A_{jl}$, so neither diagonal can dominate the other. In case (II), $A_{ik} \geq y_k \geq A_{jk} \geq y_j \geq A_{jl}$, again neither diagonal can dominate the other. The third case needs a further discussion.

Since $A^+, A^-$ are both extension of $\hat{f}$, they always agree on $\hat{X}$, so we may assume $x \neq \hat{x}, \hat{x}_j, \hat{x}_k, \hat{x}_l$ thus

$$d_i = \|x - \hat{x}_i\|, \quad d_j = \|x - \hat{x}_j\|, \quad d_k = \|x - \hat{x}_k\|, \quad d_l = \|x - \hat{x}_l\|$$

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

$$A_{ik}, A_{jl} > A_{jk}, A_{il}.$$

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

$$\frac{d_k}{d_i + d_k} y_j + \frac{d_i}{d_i + d_k} y_k > \frac{d_k}{d_j + d_k} y_j + \frac{d_j}{d_j + d_k} y_k$$

Similarly, from $A_{ik} > A_{il}, A_{jl} > A_{jk}, A_{jl} > A_{il}$ we conclude

$$(d_i + d_k)(y_j - y_l) > (d_i + d_l)(y_j - y_l),$$

$$(d_i + d_l)(y_j - y_l) > (d_i + d_k)(y_j - y_l),$$

$$(d_i + d_l)(y_j - y_l) > (d_i + d_l)(y_j - y_l).$$

Note that in case (III), every term in (9) 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 $A^+ = A^-$. □

Appendix B: Proofs for Section 3.2

**Lemma 5.** For each fixed $\lambda \geq b \lor h$ and $\hat{f} \in \hat{F}$, we can always find a $\frac{d\Psi_{\hat{f}}}{d\lambda}$-Lipschitz policy $\hat{g} \in \hat{F}$ satisfying $\hat{z}_k \leq \hat{g}(\hat{x}_k) \leq \hat{z}_k$, where

$$\hat{z}_k := \min_{(x, z) \in \text{supp} \hat{P}} \left\{ z + \|x - \hat{x}_k\| \right\}, \quad \hat{z}_k := \max_{(x, z) \in \text{supp} \hat{P}} \left\{ z - \|x - \hat{x}_k\| \right\},$$

such that for all $(\hat{x}, \hat{z}) \in \text{supp} \hat{P}$,

$$\max_{x \in \hat{X}} \left\{ \Psi(\hat{g}(x), \hat{z}) - \lambda \|x - \hat{x}\| \right\} \leq \max_{x \in \hat{X}} \left\{ \Psi(\hat{f}(x), \hat{z}) - \lambda \|x - \hat{x}\| \right\}. $$
Proof of Lemma 5. Denote \( y_k = \hat{f}(\hat{x}_k) \). First, we show that if \( \hat{f} \) 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 \( \hat{f} \). Suppose there exists \( j_0, k_0 \) such that

\[
y_{j_0} \geq y_{k_0} + L \| \hat{x}_{k_0} - \hat{x}_{j_0} \| := \tilde{y}_{j_0}
\]

for some constant \( L > 0 \) to be specified 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 \((x, \hat{z})\) in the support of \( P \),

\[
h(\tilde{y}_{j_0} - \hat{z}) + b(\hat{z} - \tilde{y}_{j_0}) - \lambda \| \hat{x}_{j_0} - \hat{x} \| \leq \max_{j=1,\ldots,K} \{ h(y_j - \hat{z}) + b(\hat{z} - y_j) - \lambda \| \hat{x}_j - \hat{x} \| \}.
\]

We consider two cases. If \( \tilde{y}_{j_0} \geq \hat{z} \), then

\[
h(\tilde{y}_{j_0} - \hat{z}) + b(\hat{z} - \tilde{y}_{j_0}) - \lambda \| \hat{x}_{j_0} - \hat{x} \| = h(\tilde{y}_{j_0} - \hat{z}) - \lambda \| \hat{x}_{j_0} - \hat{x} \| \leq h(y_{j_0} - \hat{z}) - \lambda \| \hat{x}_{j_0} - \hat{x} \| \leq \text{RHS}.
\]

If \( \tilde{y}_{j_0} < \hat{z} \), then

\[
h(\tilde{y}_{j_0} - \hat{z}) + b(\hat{z} - \tilde{y}_{j_0}) - \lambda \| \hat{x}_{j_0} - \hat{x} \| = b(\hat{z} - y_{k_0}) - \lambda \| \hat{x}_{j_0} - \hat{x} \|
\]

\[
= b(\hat{z} - y_{k_0}) - bL \| \hat{x}_{k_0} - \hat{x}_{j_0} \| - \lambda \| \hat{x}_{j_0} - \hat{x} \|
\]

\[
= b(\hat{z} - y_{k_0}) - bL \| \hat{x}_{k_0} - \hat{x}_{j_0} \| - \lambda \| \hat{x}_{j_0} - \hat{x} \|.
\]

If we choose any \( L \geq \frac{1}{\lambda} \), then

\[
h(\tilde{y}_{j_0} - \hat{z}) + b(\hat{z} - \tilde{y}_{j_0}) - \lambda \| \hat{x}_{j_0} - \hat{x} \| \leq b(\hat{z} - y_{k_0}) - \lambda \| \hat{x}_{k_0} - \hat{x}_{j_0} \| - \lambda \| \hat{x}_{j_0} - \hat{x} \|
\]

\[
\leq b(\hat{z} - y_{k_0}) - \lambda \| \hat{x}_{k_0} - \hat{x} \| \leq \text{RHS}.
\]

This completes the proof of the claim.

Now make use of the above claim iteratively.

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 \( y_j^{(2)} = y_j^{(1)} \land (y_k^{(1)} + L \| \hat{x}_j - \hat{x}_{k_1} \|) \) for every \( j \in [K] \).

Step 2. Pick \( k_2 \in \arg \min_{k \in [K] \setminus \{ k_1 \}} \{ y_k^{(2)} \} \), and define \( y_j^{(3)} = y_j^{(2)} \land (y_k^{(2)} + L \| \hat{x}_j - \hat{x}_{k_2} \|) \) 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 \( y_j^{(4)} = y_j^{(3)} \land (y_k^{(3)} + L \| \hat{x}_j - \hat{x}_{k_3} \|) \) for every \( j \in [K] \).

The above process terminates after Step \( K-1 \). We have a sequence of policies \( \hat{f}^{(m)}(\hat{x}_k) = y_k^{(m)} \).

According to the previous claim, each step does not deteriorate the objective value: for any \( 1 \leq m \leq K - 1 \),

\[
\max_{x \in \mathcal{X}} \{ \Psi(\hat{f}^{(m+1)}(x), \hat{z}) - \lambda \| x - \hat{x} \| \} \leq \max_{x \in \mathcal{X}} \{ \Psi(\hat{f}^{(m)}(x), \hat{z}) - \lambda \| x - \hat{x} \| \}.
\]

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

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

(II) The sequence \( y_k^{(m)} \) is increasing in \( m \), that is,

\[
y_k^{(1)} \leq y_k^{(2)} \leq y_k^{(3)} \leq \cdots \leq y_k^{(K)}.
\]

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

\[
y_k^{(m+1)} = y_k^{(m)} \land (y_k^{(m)} + L \| \hat{x}_{k_m} - \hat{x}_{k_{m+1}} \|) \geq y_k^{(m)}.
\]

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

\[
y_k^{(1)} = y_k^{(2)} = \cdots = y_k^{(K)}, \quad y_k^{(2)} = y_k^{(3)} = \cdots = y_k^{(K)}, \quad y_k^{(3)} = y_k^{(4)} = \cdots = y_k^{(K)}, \quad \ldots
\]

again following the definition. Therefore \( \hat{f}^{(K)}(\hat{x}_{k_m}) = y_k^{(m)} \).
Combine the above three properties, we have for any $m < n$,

$$y_{km}^{(m)} \leq y_{km}^{(n)} \leq y_{km}^{(m+1)} \leq y_{km}^{(m)} + L\|\hat{x}_n - \hat{x}_m\|.$$ 

Now we define $\hat{f} = \hat{f}^{(k)}$, then it is $L$-Lipschitz. A similar argument works for $L \geq \frac{1}{h}$, so we can pick $L = \frac{1}{b} \land \frac{1}{h} = \frac{1}{bvh}$, and by the above construction $\hat{f}$ is $\frac{1}{bvh}$-Lipschitz and satisfy

$$\max_{j=1,\ldots,K} \{h(\hat{f}(\hat{x}_j) - \hat{z}) + b(\hat{z} - \hat{f}(\hat{x}_j))_+ - \lambda\|\hat{x}_j - \hat{z}\|\} \leq \max_{j=1,\ldots,K} \{h(y_j - \hat{z}) + b(\hat{z} - y_j)_+ - \lambda\|\hat{x}_j - \hat{z}\|\}$$

for all $(\hat{x}, \hat{z}) \in \text{supp} \hat{P}$, that is,

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

Now we deal with upper and lower bound. By the first part of this proof we can assume without loss of generality that $\hat{f}$ is already $\frac{1}{bvh}$-Lipschitz to begin with. Define $\hat{y}_j = \hat{z}_j \land y_j$ for every $j$, we claim that for any $(\hat{x}, \hat{z})$ in the support of $\hat{P}$,

$$\max_{j=1,\ldots,K} \{h(\hat{y}_j - \hat{z}) + b(\hat{z} - \hat{y}_j)_+ - \lambda\|\hat{x}_j - \hat{z}\|\} \leq \max_{j=1,\ldots,K} \{h(y_j - \hat{z}) + b(\hat{z} - y_j)_+ - \lambda\|\hat{x}_j - \hat{z}\|\}.$$ 

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

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

When $\hat{y}_j = \hat{z} \leq y_j$, we split to two cases. On the one hand, if $\hat{z} \leq \hat{y}_j$, then

$$h(\hat{y}_j - \hat{z}) + b(\hat{z} - \hat{y}_j)_+ - \lambda\|\hat{x}_j - \hat{z}\| = h(\hat{y}_j - \hat{z}) - \lambda\|\hat{x}_j - \hat{z}\| \leq h(\hat{y}_j - \hat{z}) - \lambda\|\hat{x}_j - \hat{z}\| \leq \text{RHS}.$$ 

On the other hand, if $\hat{z} \geq \hat{y}_j$, using $\hat{y}_j = \hat{z} \geq \hat{z} - \|\hat{x} - \hat{z}\|$ by the definition of $\hat{z}$, so

$$h(\hat{y}_j - \hat{z}) + b(\hat{z} - \hat{y}_j)_+ - \lambda\|\hat{x}_j - \hat{z}\| = h(\hat{y}_j - \hat{z}) - \lambda\|\hat{x}_j - \hat{z}\| \leq h(\hat{y}_j - \hat{z}) - \lambda\|\hat{x}_j - \hat{z}\| \leq \text{RHS}.$$ 

Here we can for example, for any scenario we have

$$h(\hat{y}_j - \hat{z}) + b(\hat{z} - \hat{y}_j)_+ - \lambda\|\hat{x}_j - \hat{z}\| \leq \max_{j=1,\ldots,K} \{h(y_j - \hat{z}) + b(\hat{z} - y_j)_+ - \lambda\|\hat{x}_j - \hat{z}\|\}.$$ 

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

Similarly we can let $\tilde{y}_j = \hat{z}_j \lor y_j$, and $\tilde{y}_j$ will satisfy

$$\max_{j=1,\ldots,K} \{h(\tilde{y}_j - \hat{z}) + b(\hat{z} - \tilde{y}_j)_+ - \lambda\|\hat{x}_j - \hat{z}\|\} \leq \max_{j=1,\ldots,K} \{h(y_j - \hat{z}) + b(\hat{z} - y_j)_+ - \lambda\|\hat{x}_j - \hat{z}\|\}.$$ 

Now we define $\tilde{g}(\hat{x}_j) = \hat{z}_j \lor (\hat{z}_j \land \tilde{f}(\hat{x}_j))$ for every $j$, with $\tilde{f}$ being the $\frac{1}{bvh}$-Lipschitz function define in the first part of the proof. Then $\tilde{g}$ satisfies

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

Moreover, note that $\hat{x}_k \mapsto \hat{z}_k$ and $\hat{x}_k \mapsto \hat{z}_k$ are $1$-Lipschitz since they are the max and min of a family of $1$-Lipschitz function of $\hat{x}_k$, and $1 \leq \frac{1}{bvh}$, so $\tilde{g}$ is $\frac{1}{bvh}$-Lipschitz, and by definition $\hat{z}_k \leq \tilde{g}(\hat{x}_k) \leq \hat{z}_k$. 

**Proof of Theorem 2.** We start by proving that

$$v_D = \min_{f \in \mathcal{F}, \lambda \geq bvh} \left\{ \lambda \rho + \mathbb{E}_P \left[ \max_{x \in \mathcal{X}} \{\Psi_f(x, \hat{Z}) - \lambda\|x - \hat{X}\|\} \right]\right\}. \tag{10}$$

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

$$v_D = \inf_{f \in \mathcal{F}, \lambda \geq bvh} \inf_{z \in \mathcal{Z}} \left\{ \lambda \rho + \mathbb{E}_P \left[ \sup_{x \in \mathcal{Z}} \{\Psi_f(x, z) - \lambda\|z - \hat{Z}\|\} - \lambda\|x - \hat{X}\|\} \right]\right\}.$$ 

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

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

is already
will be infinity as \( z \to \infty \). Therefore, in order to find the minimum over \( \lambda \) we can disregard this case and constrain \( \lambda \geq b \lor h \). In this case, sup over \( z \) is \( \Psi ( f(x), \hat{Z} ) \) attained at \( z = \hat{Z} \), which proves (10).

For each \( \lambda \geq b \lor h \), denote

\[
\mathcal{F}_\lambda := \{ \hat{g} \in \mathcal{F} : \mathcal{Z}_k \leq \hat{g}(\hat{x}_k) \leq \mathcal{Z}_k, \forall k \}, \quad \mathcal{F}_\lambda := \{ \hat{g} \in \mathcal{F} : \| \hat{g} \|_{\text{Lip}} \leq \frac{\lambda}{b \lor h} \}.
\]

Then

\[
u_D = \inf_{\lambda \geq b \lor h} \inf_{f \in \mathcal{F}_\lambda} \left\{ \lambda \rho + \mathbb{E}_{\hat{p}} \left[ \max_{x \in \mathcal{X}} \{ \Psi ( f(x), \hat{Z} ) - \lambda \| x - \hat{x} \| \} \right] \right\},
\]

where we replace \( \mathcal{F} \) by \( \mathcal{F}_\lambda \) in (10) using Lemma 5. For each \( f \in \mathcal{F}_\lambda \), because \( \| \Psi ( f(\cdot), \hat{Z} ) \|_{\text{Lip}} \leq \| f \|_{\text{Lip}} (b \lor h) \leq \lambda \), the max over \( x \) is attained at \( x = \hat{x} \). Therefore,

\[
u_D = \inf_{\lambda \geq b \lor h} \inf_{f \in \mathcal{F}_\lambda} \left\{ \lambda \rho + \mathbb{E}_{\hat{p}} \left[ \Psi ( f(\hat{x}), \hat{Z} ) \right] \right\}.
\]

Now we switch the inf over \( \lambda \) and inf over \( \hat{f} \),

\[
u_D = \inf_{\lambda \geq b \lor h} \inf_{f \in \mathcal{F}} \left\{ \lambda \rho + \mathbb{E}_{\hat{p}} \left[ \Psi ( f(\hat{x}), \hat{Z} ) \right] \right\}
\]

Using the change of variables \( y_k = \hat{f}(x_k) \), \( k = 1, \ldots, K \), the above is equivalent to

\[
u_D = \inf_{y_k \in \{ \mathcal{Z}_k, \mathcal{Z}_k \}, 1 \leq k \leq K} \left\{ (b \lor h) \left( 1 \lor \max_{i,j} \frac{|y_i - y_j|}{\| \hat{x}_i - \hat{x}_j \|} \right) \rho + \frac{1}{n} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \Psi ( y_k, \hat{z}_k ) \right\}.
\]

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 \( \mathcal{F} \) in place of \( \mathcal{F} \), we conclude that \( \nu_D = \nu_R \) and (\( D \)) is equivalent to (\( R \)).

\section*{Appendix C: Proofs for Section 3.3}

We recall and define \( L = \| f^* \|_{\text{Lip}} \) and

\[
\hat{x}^*_c := \{ \hat{x} \in \hat{x} : \hat{q}(\hat{x}) < \hat{f}^*(\hat{x}) \}, \quad \hat{x}^*_c := \{ \hat{x} \in \hat{x} : \hat{q}(\hat{x}) > \hat{f}^*(\hat{x}) \},
\]

\[
\hat{x}^*_e := \{ \hat{x} \in \hat{x} : \hat{q}(\hat{x}) \leq \hat{f}^*(\hat{x}) \leq \hat{q}(\hat{x}) \}, \quad \hat{x}^* := \hat{x}^*_c \cup \hat{x}^*_e, \quad \hat{x}^*_c := \hat{x}^*_c \cup \hat{x}^*_e.
\]

Proof of Lemma 1. 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.

\begin{itemize}
  \item \textbf{Step 1.} Define \( y_k^* \leftarrow \hat{q}(\hat{x}_k) \), \( \forall k \in [K] \). By (I), we know that \( y_k^* \) satisfies
  \begin{equation}
  \hat{q}(\hat{x}_k) - y_k^* \leq \| \hat{x}_j - \hat{x}_k \|, \quad \forall j, k \in [K].
  \end{equation}
  Note that this also implies \( \hat{q}(\hat{x}_j) \leq y_k^* \) for all \( j \) by setting \( k = j \).
  \item \textbf{Step 2.} Choose \( k_1 \in \arg \min_{k \neq k_1} y_k^* \). For any \( k \neq k_1 \), denote \( y_k^{(1)} = y_{k_1}^* + \| \hat{x}_k - \hat{x}_{k_1} \| \). Then for all \( j, k \),
  \[
  \hat{q}(\hat{x}_j) - y_k^{(1)} = \hat{q}(\hat{x}_j) - y_k^* - \| \hat{x}_k - \hat{x}_{k_1} \| \leq \| \hat{x}_j - \hat{x}_{k_1} \| - \| \hat{x}_k - \hat{x}_{k_1} \| \leq \| \hat{x}_j - \hat{x}_k \|.
  \]
  This means that if we reassign values to \( y_k' \leftarrow y_k^* \land y_k^{(1)} \), \( \forall k \neq k_1 \), then (12) would still hold. Note that since \( y_k^{(1)} \geq y_k^* \), after reassignment \( y_k' \) is still the smallest among all \( y_k' \).
\end{itemize}
Step 3. Choose \( k_2 \in \arg\min_{k\neq k_1} y_k' \). For any \( k \notin \{k_1, k_2\} \), denote
\[
y_k^{(2)} = y_{k_2} + \|\hat{x}_k - \hat{x}_{k_2}\|,
\]
and reassign \( y_k' = y_k^* \wedge y_k^{(2)} \). Same as Step 2, (12) still holds, and \( y_k^* \leq y_{k_2}^* \leq y_k^* \) for all \( k \notin \{k_1, k_2\} \).

Step 4. Repeat Step 3. Eventually we would have \( y_{k_1}^* \leq y_{k_2}^* \leq \cdots \leq y_{k_f}^* \), with (12) still holds. Now for every \( i < j \), we have
\[
0 \leq y_{i_j}^* - y_{i_k}^* \leq y_{i_k}^* - y_{i_{k_j}}^* = \|\hat{x}_{i_j} - \hat{x}_{i_k}\|.
\]
Therefore, define \( \hat{f}(\hat{x}_k) := y_k^* \), then \( \hat{f} \) is a 1-Lipschitz function. In the above process \( y_k^* \) is decreasing its value, so \( y_k^* \leq q(\hat{x}_k) \) which is its initial value, and (12) ensures \( y_k^* \geq q(\hat{x}_k) \) by setting \( j = k \).

Since \( q \leq \hat{f} \leq q \), \( \hat{f} \) is a conditional \( \frac{\hat{f}}{q} \) quantile, with \( \|\hat{f}\|_{\text{Lip}} \leq 1 \). Then it must be a minimizer of \( (\hat{R}) \) for any \( \rho \geq 0 \), because it minimizes both terms. Any other optimal policy \( f^* \) 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 \( f^* \) must has Lipschitz constant \( \|f^*\|_{\text{Lip}} = L \geq 1 \). Indeed, if \( L < 1 \), we can always adjust the value of \( f^* \) to reduce costs in the second term of \( v_R \) without paying more cost in the first term. So, the only possibility that \( \|f^*\|_{\text{Lip}} < 1 \) is that the second term is already optimized, that is \( q \leq f^* \leq \tilde{q} \). However, this would imply (I) holds, which is a contradiction.

Now we partition \( X \) according to (11). First we fix \( \hat{x}_k \in \hat{X}_k \). Indeed, there must be \( \hat{x}_{j_1} \in \hat{X}_k \setminus \{\hat{x}_k\} \) such that \( y_k - y_{j_1} = L\|\hat{x}_k - \hat{x}_{j_1}\| \), otherwise we can increase the value of \( y_k \) and optimize the second term in (10) without jeopardizing the first term. If \( \hat{x}_{j_1} \in \hat{X}_k \), then the claim is proved. For the same reason, if \( \hat{x}_{j_1} \in \hat{X}_k \), then we can find \( \hat{x}_{j_2} \in \hat{X}_l \setminus \{\hat{x}_k, \hat{x}_{j_1}\} \) such that \( y_{j_1} - y_{j_2} = L\|\hat{x}_{j_1} - \hat{x}_{j_2}\| \), thus \( y_k - y_{j_2} = L\|\hat{x}_k - \hat{x}_{j_2}\| + L\|\hat{x}_{j_1} - \hat{x}_{j_2}\| \geq L\|\hat{x}_k - \hat{x}_{j_2}\| \), and here inequality sign must be equality because \( f^* \) is \( L \)-Lipschitz. Note that this also shows that \( (\hat{x}_k, y_k), (\hat{x}_{j_1}, y_{j_1}) \) and \( (\hat{x}_{j_2}, y_{j_2}) \) are on the same straight line if \( \|\cdot\| = \|\cdot\|_2 \). If \( \hat{x}_{j_2} \in \hat{X}_k \), then we finish the proof of the claim, otherwise \( \hat{x}_{j_1} \) can be found. Note that \( y_k > y_{j_1} > y_{j_2} > \cdots \) is strictly decreasing, thus \( \hat{x}_k, \hat{x}_{j_1}, \hat{x}_{j_2}, \cdots \) are distinct. Finitely many steps later, we must have \( \hat{x}_j \in \hat{X}_k \) and \( y_k - y_j = L\|\hat{x}_k - \hat{x}_j\| \) before we run out of points.

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

**Proposition 2 (Worst-case Distribution).** Let \( \hat{f}^* \) be an in-sample optimal robust policy. Then there exists a worst-case distribution \( P^* \) of \( \hat{P} \) such that the following holds.

(I) If \( \|\hat{f}^*\|_{\text{Lip}} \leq 1 \), then \( P^* \) perturbs \( \hat{P} \) by moving \( (\hat{x}, \hat{z}) \) with \( \hat{z} \leq \hat{f}^*(\hat{x}) \) toward \( (\hat{x}, z') \) for some \( z' > \hat{z} \).

(II) If \( \|\hat{f}^*\|_{\text{Lip}} > 1 \), then \( P^* \) perturbs \( \hat{P} \) by moving each \( (\hat{x}, \hat{z}) \) with \( \hat{x} \in \hat{X}_k \) and \( \hat{z} \geq \hat{f}^*(\hat{x}) \) toward \( (x', \hat{z}) \) for some \( x' \) in \( \hat{X}_k \setminus \{\hat{x}\} \) and \( f^*(\hat{x}) - f^*(x') = \|\hat{f}^*\|_{\text{Lip}}\|\hat{x} - x'\| \).

In Figure 7, we have the identical setting as in Figure 4. Above the graph of \( f^* \) is a blue shadow region representing \( \{(x, z) : z \geq f^*(x)\} \), and \( P^* \) moves the probability mass in this region when backorder costs more than holding. In the left figure \( \|\hat{f}^*\|_{\text{Lip}} < 1 \), so it is more cost-efficient to move along the direction of \( z \). In the right figure \( \|\hat{f}^*\|_{\text{Lip}} > 1 \), 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.

![Figure 7](image-url) 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).
Proof of Proposition 2. To ease the notation we use $f$ to represent $f^*$ in this proof.

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

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

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

$$\mathbb{E}_{\mathbb{P}}[\Psi_f] - \mathbb{E}_{\hat{\mathbb{P}}}[\Psi_f] = \mathbb{E}_{\hat{\mathbb{P}}}[\Psi_f \circ T(X, Z)] - \mathbb{E}_{\hat{\mathbb{P}}}[\Psi_f(X, Z)]$$

$$= \mathbb{E}_{\hat{\mathbb{P}}}[1_{\{Z \geq f(X)\}}(\Psi_f(X, Z + t) - \Psi_f(f(X), Z))]$$

$$= b t \mathbb{E}_{\hat{\mathbb{P}}}[Z \geq f(X)]$$

By choosing $t = \rho / (\hat{\mathbb{P}}[Z \geq f(X)])$, we have $\mathbb{W}(\mathbb{P}, \hat{\mathbb{P}}) = \rho$, so $\mathbb{P}$ is feasible, and $\mathbb{E}_{\mathbb{P}}[\Psi_f] = \mathbb{E}_{\hat{\mathbb{P}}}[\Psi_f] + b \rho$ is a worst case distribution. Note that the denominator $\hat{\mathbb{P}}[Z \geq f(X)]$ is never zero, otherwise $\hat{X} \subset \hat{X}_\subset$ and $\hat{X}_\subset = \emptyset$ which contradicts with Lemma 1.

(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 $\hat{x}_j \in \hat{X}_\subset$ can be from $\hat{x}_k \in \hat{X}_\subset$ in Lemma 1 (II). For every $k$ we define

$$\tau(k) = \arg \min_{j \neq k} \{y_j : y_k - y_j = L(\hat{x}_k - \hat{x}_j)\}, \quad \Delta(\hat{x}_k) := \|\hat{x}_k - \hat{x}_{\tau(k)}\| = \frac{1}{L}(y_k - y_{\tau(k)})$$

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

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

$$T(\hat{x}_k, z) := \begin{cases} (\hat{x}_{\tau(k)}, z) & x \in \hat{X}_\subset, z \geq f(x) \\ (\hat{x}_k, z) & x \in \hat{X}_\subset \text{ or } z < f(x) \end{cases}$$

This implies that for every $k$,

$$\Psi_f \circ T(\hat{x}_k, z) - \Psi_f(\hat{x}_k, z) = 1_{\{\hat{x}_k \in \hat{X}_\subset, z \geq f(\hat{x}_k)\}}(b(z - y_{\tau(k)}) - b(z - y_k))$$

$$= b 1_{\{\hat{x}_k \in \hat{X}_\subset, z \geq y_k\}} L \|\hat{x}_k - \hat{x}_{\tau(k)}\|$$

$$= b L 1_{\{\hat{x}_k \in \hat{X}_\subset, z \geq y_k\}} \Delta(\hat{x}_k)$$

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

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

We will show that

$$\mathbb{E}_{\hat{\mathbb{P}}}[\Psi_f] - \mathbb{E}_{\hat{\mathbb{P}}}[\Psi_f] \geq b L \rho.$$

If this is true, we can construct a feasible $\mathbb{P}$ by convex combination (let $\mathbb{P} = \theta \hat{\mathbb{P}} + (1 - \theta) \hat{\mathbb{P}}$), such that

$$\mathbb{E}_{\mathbb{P}}[\Psi_f] - \mathbb{E}_{\hat{\mathbb{P}}}[\Psi_f] = b L \rho \geq b L \mathbb{W}(\mathbb{P}, \hat{\mathbb{P}}),$$

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

Recall $\Delta$ is defined in (13). To prove (14), we construct a competitor $f^*(\hat{x}_k) := y_k - \epsilon \Delta(\hat{x}_k)$ which means “ordering less” than $f$. We claim when $\epsilon$ is small, $\|f^*\|_{\text{Lip}} = L - \epsilon$. To see why this is true, we can consider
only the pairs of points \( \hat{x}_k \) and \( \hat{x}_j \) with \( y_k - y_j = L\|\hat{x}_k - \hat{x}_j\| \) since \( \epsilon \) can be chosen sufficiently small. In this situation,
\[
f^\varepsilon(\hat{x}_k) - f^\varepsilon(\hat{x}_j) = y_k - y_j - \epsilon(\Delta(\hat{x}_k) - \Delta(\hat{x}_j)) = y_k - y_j - \frac{\epsilon}{L}((y_k - y_{r(k)}) - (y_j - y_{r(j)})).
\]
It can be seen that \( y_{r(k)} \leq y_{r(j)} \) by the minimality of \( r(k) \) (see the last paragraph in the proof of Lemma 1), hence
\[
f^\varepsilon(\hat{x}_k) - f^\varepsilon(\hat{x}_j) \leq y_k - y_j - \frac{\epsilon}{L}(y_k - y_j) = \left(1 - \frac{\epsilon}{L}\right)(y_k - y_j) = (L - \epsilon)\|\hat{x}_k - \hat{x}_j\|.
\]

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

Since \( f \) minimizes \( u_p \), we have
\[
(b \vee h)L\rho + \mathbb{E}_p[\Psi_f] \leq (b \vee h)(L - \epsilon)\rho + \mathbb{E}_p[\Psi_{f^\varepsilon}]
\]
\[
(b \vee h)\epsilon\rho \leq \mathbb{E}_p[\Psi_{f^\varepsilon} - \Psi_f].
\]

Here, “ordering less” means we need to pay more backorder cost and less holding cost, so
\[
\Psi(f^\varepsilon(\hat{x}_k), z) - \Psi(f(\hat{x}_k), z) = \begin{cases} \]b\epsilon\Delta(\hat{x}_k), & y_k \leq z, \\ -h\epsilon\Delta(\hat{x}_k), & y_k > z.\end{cases}
\]

Here we choose \( \epsilon \) small such that \( y_k > z \) implies \( y_k > z + \epsilon\Delta(\hat{x}_k) \) for every \((\hat{x}_k, z) \in \text{supp} \hat{\mathbb{P}}\). Now take the conditional expectation,
\[
\mathbb{E}_{\hat{\mathbb{P}}^{f^\varepsilon}_X}[\Psi(f^\varepsilon(X), Z) - \Psi(f(X), Z)|X = \hat{x}_k] = \epsilon\Delta(\hat{x}_k)\left[b\hat{\mathbb{P}}[Z \geq y_k|X = \hat{x}_k] - h\hat{\mathbb{P}}[Z < y_k|X = \hat{x}_k]\right].
\]

If \( \hat{x}_k \in X_\leq, y_k \) is no less than the conditional quantile, so the above will be nonpositive. Thus
\[
\mathbb{E}_{\hat{\mathbb{P}}^{f^\varepsilon}_X}[\Psi(f^\varepsilon(X), Z) - \Psi(f(X), Z)|X = \hat{x}_k] \leq \begin{cases} b\epsilon\Delta(\hat{x}_k)\hat{\mathbb{P}}[Z \geq y_k|X = \hat{x}_k], & \hat{x}_k \in X_\leq, \\ 0, & \hat{x}_k \in X_\geq.\end{cases}
\]

Finally, take expectation in \( X \), we have
\[
(b \vee h)\epsilon\rho \leq \mathbb{E}_{\hat{\mathbb{P}}}[\Psi_{f^\varepsilon} - \Psi_f] \leq b\epsilon\mathbb{E}_{\hat{\mathbb{P}}}[\Delta(X)1_{\{X \in X_\leq, Z \geq f^\varepsilon(X)\}}].
\]

In particular, \( \mathbb{E}_{\hat{\mathbb{P}}}[\Psi_f \circ T - \Psi_f] = bL\mathbb{E}_{\hat{\mathbb{P}}}[\Delta(X)1_{\{X \in X_\leq, Z \geq f^\varepsilon(X)\}}] \geq bL\rho \), which proves (14). \( \square \)

Remark 1. Similarly, when \( b < h \), the transport map should move \( X \in \hat{X}_\leq \) in \( \{Z \leq f^\varepsilon(X)\} \) when \( L > 1 \), and should decrease \( Z \) in \( \{Z \leq f^\varepsilon(X)\} \) when \( L \leq 1 \). Considering the demand must be nonnegative, in the case \( L \leq 1 \) we should use the transport map
\[
T(x, z) := \begin{cases} (x, (z - t)_+) & z \geq f^\varepsilon(x) \\ (x, z) & z \leq f^\varepsilon(x).\end{cases}
\]

for some \( t \geq 0 \) if \( \rho \leq \mathbb{E}_{\hat{\mathbb{P}}}[Z1_{\{Z < f^\varepsilon(X)\}}] \). From this proposition we can see that, when \( \rho \) is sufficiently small, the worst case distribution \( \hat{\mathbb{P}} \) is still supported in \( \hat{X} \times Z \).

Essentially the only place where \( b \geq h \) really matters is the proof of (10), 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 \( Z = \mathbb{R} \) instead of \( \mathbb{R}_+ \) in (P):
\[
u_p = \min_{f: \hat{X} \to \mathbb{R}} \sup_{\mathbb{P} \in \Pi_{\hat{X} \times \mathbb{R}}} \mathbb{E}_{(X, Z), -\mathbb{P}}[\Psi_f(X, Z)]: \forall \mathbb{V}(\hat{\mathbb{P}}) \leq \rho \] ,
we would have the same worst case distribution supported over \( \{Z \geq 0\} \), so the value of \( \nu_p \) is going to be the same. Since the proof of \( \nu_p = \nu_p = \nu_p = \nu_p \) in Theorem 1 doesn’t rely on whether \( Z = \mathbb{R} \) or \( \mathbb{R}_+ \) because of the property (iii) in the Lemma 1, \( Z \) in (P), (D), (P), (D) can all be replaced by \( \mathbb{R} \), and thus we can send \( z \to -\infty \) to fix the proof for (10).
References


