informs
https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 70, No. 4, April 2024, pp. 2315–2329
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Optimal Robust Policy for Feature-Based Newsvendor

Luhao Zhang,[a] Jincheng Yang,[a] Rui Gao[b,*]

[a] Department of Mathematics, The University of Texas at Austin, Austin, Texas 78712; [b] Department of Information, Risk and Operations Management, The University of Texas at Austin, Austin, Texas 78712

*Corresponding author

Contact: luhaozhang@utexas.edu, https://orcid.org/0000-0001-8568-3581 (LZ); jincheng@utexas.edu, https://orcid.org/0000-0002-3581-9425 (JY); rui.gao@mccombs.utexas.edu, https://orcid.org/0000-0003-0145-8577 (RG)

Received: May 9, 2022
Revised: September 3, 2022
Accepted: October 13, 2022
Published Online in Articles in Advance: June 2, 2023

https://doi.org/10.1287/mnsc.2023.4810

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

History: Accepted by J. George Shanthikumar, data science.
Supplemental Material: The online appendix and data are available at https://doi.org/10.1287/mnsc.2023.4810.

Keywords: side information • contextual decision making • inventory management • adjustable robust optimization

1. Introduction

The newsvendor model is a classic and fundamental problem in operations management that faces new challenges in the era of big data. More often than not, much feature information—temporal, spatial, social, or economical—is available prior to decision making and reveals partial information on the product demand. The feature information reduces the uncertainty and helps the decision maker customize ordering decisions to each individual feature realization. It is crucial to involve feature information in decision making. Otherwise, the decision may be inconsistent, namely, not converging to the true optimal policy even with an infinite amount of data (Ban and Rudin 2019). The ordering decision is often made for a population of features but not for a single feature realization. For example, ordering decisions are made for multiple shelves at different locations in selected time windows or for a number of customers with different demographics. In these cases, it would make sense to consider the average performance over the entire distribution of features. In this work, we are interested in the policy optimization decision-making problem. It seeks a policy (aka decision rule) that outputs an ordering decision for every feature value to minimize the overall average cost. We refer to the paragraph on 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 equals the conditional critical fractile of the demand distribution under each feature value. Unfortunately, many real-world problems comprise a potentially large set of feature values that historic data cannot exhaust. Thus, the true underlying conditional distribution of the demand under a new feature value is likely unknown. In this case, a natural way is to replace the unknown underlying distribution of demands and features with its empirical counterpart. However, the resulting empirical risk-minimization problem produces a pathological policy that can take arbitrary values on unseen feature values; see a more detailed discussion in Section 2.2.

The pathological behavior of an empirical feature-based newsvendor motivates the development of methods to generalize ordering decisions to unseen feature values. The most common approach is parameterization, namely, restricting the search to a parametric policy class. For example, section 2.3.1 of Ban and Rudin (2019) studies affine policies, which can be efficiently solved using linear or convex optimization methods. The affine class can be restrictive and suboptimal. Indeed, numerical experiments in Ban and Rudin (2019) show that affine policies are outperformed by their proposed kernel optimization method in the same paper. One possible remedy is to consider nonlinear transformations of features (basis functions). Thereby, one can enlarge the policy search space to an arbitrarily complex class with coefficients affinely dependent on the basis functions (Bertsimas and Koduri 2022). However, specifying nonlinear transformations with good interpretability is a fundamentally challenging question. Similarly, neural-network polices (Oroojlooyjadid et al. 2020, Meng et al. 2022) may have nice empirical performance but are often hard to interpret and data demanding. In summary, in existing methods, there is a trade-off between the richness of the policy class and its interpretability/tractability. As such, the following question remains open: can we find an optimal policy without restricting onto a parametric family and still maintaining computational efficiency and interpretability?

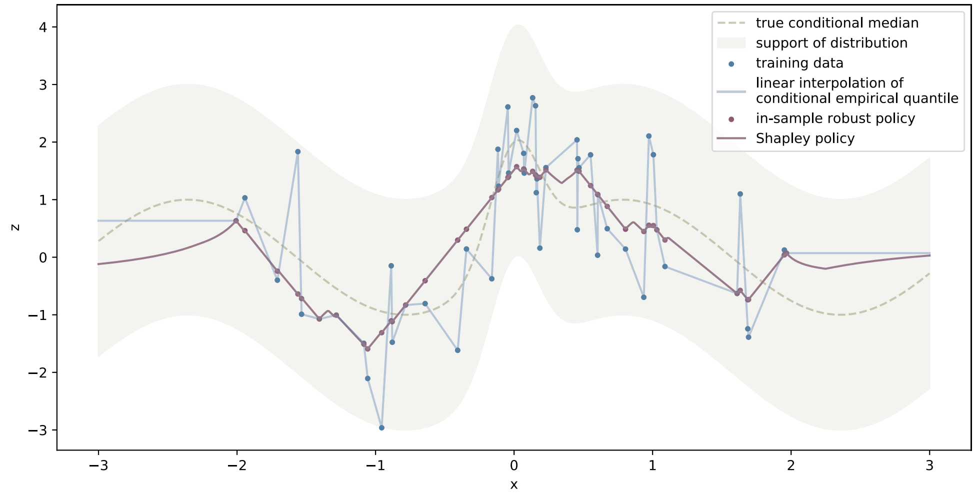
To answer this question, we consider a distributionally robust policy optimization framework. More specifically, it optimizes over all policies that are measurable functions of the features and does not parameterize the policy class. Moreover, it involves a minimax Wasserstein distributionally robust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

Our formulation belongs to the class of adjustable robust optimization (Yanıkoğlu et al. 2019). Computationally, this class of problems involves a challenging infinite-dimensional functional optimization over the space of policies, which is "typically severely computationally intractable" (Ben-Tal et al. 2009, p. 365, chapter 14.2.3). Without parameteriztion, the optimal solutions are generally unknown except for a few notable cases. Perhaps surprisingly, by utilizing the structure of the problem, we are able to identify a new class of policies that are proven to be optimal and can be computed efficiently. More specifically,

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

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

Figure 1. (Color online) An Illustration of Policies for Which Training Data Are Generated from a Two-Dimensional Continuous Distribution of Feature-Demand Pairs (x, z)



Note. Our proposed Shapley policy has a smaller variation (Lipschitz norm) compared with the linear interpolation of conditional quantiles of the empirical distribution.

of the Shapley policy with various benchmarks using synthetic and real data, which demonstrate its superior empirical performance.

1.1. Related Literature

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

Traditionally, a widely used approach to this problem is a two-step predict-then-optimize process that first predicts the conditional demand distribution given a new feature value and then optimizes for the ordering quantity (e.g., Toktay and Wein 2001, Zhu and Thonemann 2004). There are some theoretical guarantees in this approach discussed in El Balghiti et al. (2019) and Hu et al. (2022). However, as pointed out by Liyanage and Shanthikumar (2005) and Ban and Rudin (2019) among others, for high-dimensional problems, the demand model specification may be challenging, and the first step estimation error can be amplified in the second step optimization.

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

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

1.1.3. 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. 2015b, Natarajan et al. 2018), tail information (Das et al. 2021), temporal dependence (See and Sim 2010, Carrizosa et al. 2016, Xin and Goldberg 2022), total variation distance (Rahimian et al. 2019a, b), phi-divergence (Ben-Tal et al. 2013, Bayraksan and Love 2015, Wang et al. 2016, Fu et al. 2021), Wasserstein distance (Lee et al. 2012, Mohajerin Esfahani and Kuhn 2018, Lee et al. 2020, Chen and Xie 2021, Gao and Kleywegt 2022), etc. Except for See and Sim (2010), most of these works do not consider feature information. In our analysis, we use the duality results for Wasserstein distributionally robust optimization (Gao and Kleywegt 2022) to obtain an equivalent reformulation of the worst case newsvendor cost for a fixed policy. Nevertheless, we emphasize that the main challenge and focus of this paper is on the policy optimization that is not studied by existing distributionally robust optimization literature.

1.1.4. On Adjustable Robust Optimization. Various decision-rule approaches are studied extensively in the literature, including affine families (Chen et al. 2008; Bertsimas et al. 2010, 2011, 2023; Bertsimas and Goyal 2012; Iancu et al. 2013; El Housni and Goyal 2021; Georghiou et al. 2021), binary decision rules (Bertsimas and Georghiou 2015), *k*-adaptability (Hanasusanto et al. 2015a, 2016; Subramanyam et al. 2019), iterative splitting of uncertainty sets (Postek and den Hertog 2016), nonparametric Markovian stopping rules (Sturt 2021), etc. Most of these works do not consider feature information except for Bertsimas et al. (2023), who consider dynamic decision making with side information using affine decision rules. Unlike this work, we consider general decision rules in a static setting. Moreover, policy classes with a provable zero suboptimality gap are rare commodities (Bertsimas et al. 2010, Bertsimas and

Goyal 2012, Iancu et al. 2013, Georghiou et al. 2021, Sturt 2021). Our result complements the literature by providing a new class of provably optimal policies for adjustable robust optimization.

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

2. Model Setup and Main Results

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

2.1. Policy Optimization for Feature-Based Newsvendor

Consider a company selling a perishable product that 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 and backorder costs, respectively. The total cost Ψ is then computed as

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

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

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

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

where the expectation is taken with respect to the conditional demand distribution of Z given $X = x$. Such an objective is considered in the pioneer work of Ban and Rudin (2019) on the big data newsvendor. If the true underlying demand distribution is known, the true optimal ordering quantity equals the critical fractile of the true conditional demand distribution.

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

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

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

2.2. Distributionally Robust Formulation

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

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

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

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

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

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

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

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

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

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

where ρ is the radius of the 1-Wasserstein ball that we use to construct the uncertainty set. Note that (P) can be viewed as a special case of policy optimization for two-stage Wasserstein distributionally robust optimization (Bertsimas et al. 2023). Throughout the paper, we assume $b > 0$ and $0 \le h \le b$. We remark that when $b < h$, all results in the paper still hold for sufficiently small ρ (see Remark EC.1 in Online Appendix D).

2.3. Main Results

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

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

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

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

II. *(Shapley extension) With $\hat{f}^*$ being the optimal solution to the preceding linear programming problem, an optimal policy f^* for problem* (P) *is defined by the following minimax matrix saddle point:*

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

Here,

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

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

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

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

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

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

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

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

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

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

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

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

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

3. Proof of Main Results

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

3.1. Shapley Policy and Its Optimality

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

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

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

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

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

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

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*

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

where $$A_{jk}(x) := \frac{\|x-\hat{x}_k\|}{\|x-\hat{x}_j\| + \|x-\hat{x}_k\|}\hat{f}(\hat{x}_j) + \frac{\|x-\hat{x}_j\|}{\|x-\hat{x}_j\| + \|x-\hat{x}_k\|}\hat{f}(\hat{x}_k) \text{ when } j \ne k$$

and $$A_{kk}(x) := \hat{f}(\hat{x}_k). \tag{S}$$

Here, 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 $\Phi : \mathbb{R} \to \mathbb{R}$,*

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

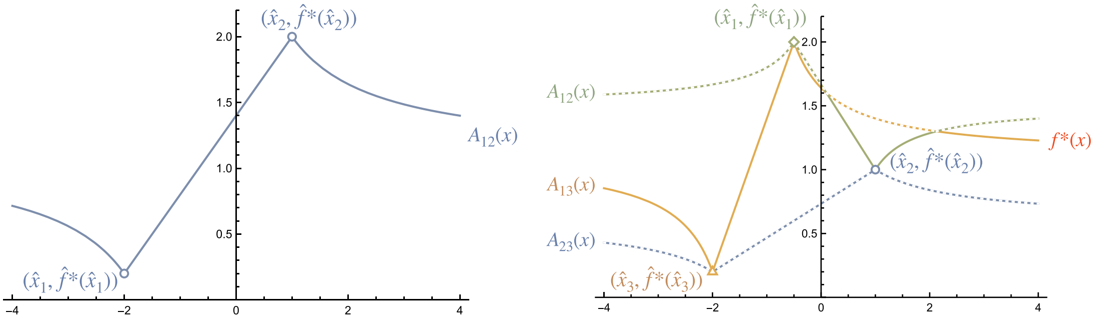
iii. *(Boundedness) $\min_{k\in[K]} \hat{f}(\hat{x}_k) \le f \le \max_{k\in[K]} \hat{f}(\hat{x}_k)$.*

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

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 the inverse distance weighting with two points (Shepard 1968). The extension $f(x)$ is given by the saddle point (pure Nash equilibrium) of a matrix $A_{jk}(x)$, whose existence is because of Shapley's theorem (Lemma EC.2 in Online 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 (4) and is the key to the proof of Theorem 1. The third and fourth properties indicate that the bound and Lipschitz norm of the extended policy is controlled by those of the in-sample policy.

As an illustration, in the two plots of Figure 2, we plot the Shapley extension when $K = 2, 3$ and when the feature space $\mathcal{X} = \mathbb{R}$. The horizontal axis represents the feature X, and the vertical axis represents the policy value (ordering quantity). The points represent an in-sample robust optimal ordering policy. When $K = 2$, the extension $f^*(x) = A_{12}(x)$. On the line segment connecting two points $\hat{x}_1$ and $\hat{x}_2$, the interpolation is linear and is curved elsewhere. As $|x| \to \infty$, the policy converges to a "noninformative" ordering quantity $(\hat{f}^*(\hat{x}_1) + \hat{f}^*(\hat{x}_2))/2$, which, intuitively, means that the historic observations of features provides little guidance on a faraway new feature value x, and thus, the policy simply takes the average of the two in-sample policy values.

Figure 2. (Color online) Graph of the Shapley Policy $y = f(x)$ When $K = 2, 3, x \in \mathbb{R}$



When $K = 3$, for each group of three historic observations of features $\hat{x}_1, \hat{x}_2, \hat{x}_3$, we plot three curves $A_{12}(x)$, $A_{13}(x)$, and $A_{23}(x)$. By solving the minimax saddle point problem, the extended policy $f^*(x)$ is 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 historic observations of features, $f^*(x)$ converges to $(\min_{k\in[K]}\hat{f}^*(\hat{x}_k) + \max_{k\in[K]}\hat{f}^*(\hat{x}_k))/2$. The case of two-dimensional feature space $\mathcal{X} = \mathbb{R}^2$ is similar as shown in Figure EC.1 in Online Appendix A.

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

3.2. Further Intuition Behind the Shapley Policy

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

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

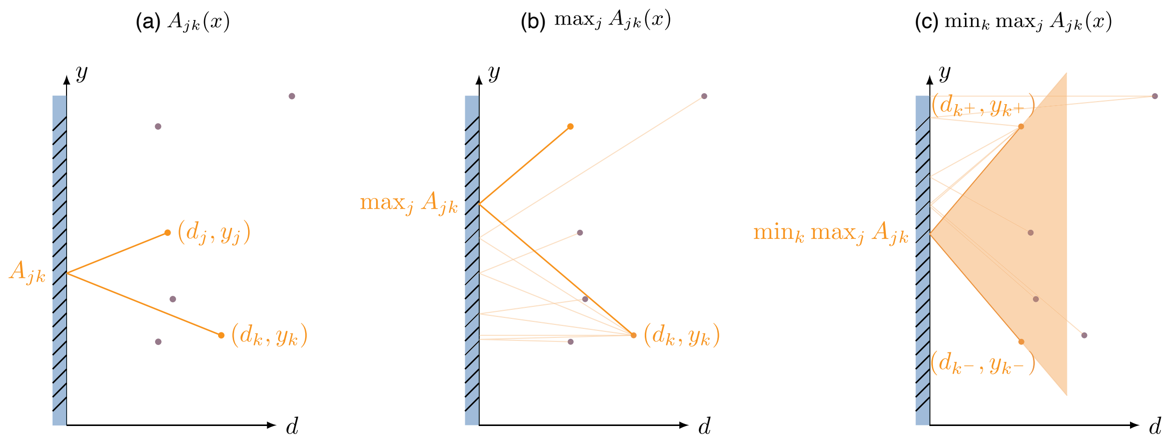
Pick any $x \in \mathcal{X} \setminus \hat{\mathcal{X}}$ and $z \in \mathcal{Z}$. Denote $y_k = \hat{f}(\hat{x}_k)$, $y = f(x)$, $d_k = \|\hat{x}_k - x\|$, and the slope of a secant line connecting x and $\hat{x}_k$ by $L_k := (y_k - y)/d_k$, and define $L := \max_{k\in[K]} |L_k|$. We claim that L can 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 $y_j \le y$, then there exists some $\delta \in [0, 1]$ such that $y = \delta y_j + (1 - \delta)y_{k^+}$ because $y_{k^+} = y + Ld_{k^+} \ge y$. Now, we can propose another transport plan, which instead moves δ fraction of the mass on $(\hat{x}_j, \hat{z}_j)$ to $(\hat{x}_j, z)$ 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 because of the convexity of $\Psi(\cdot, z)$: $\delta\Psi(y_j, z) + (1 - \delta)\Psi(y_{k^+}, z) \ge \Psi(y, z)$.
- The new plan has a smaller transport cost: distance in the z-direction remains the same, whereas in the x-direction, the distance is shorter: $\delta\|\hat{x}_j - \hat{x}_j\| + (1 - \delta)\|\hat{x}_j - \hat{x}_{k^+}\| \le 0 + (1 - \delta)(d_j + d_{k^+}) \le d_j = \|\hat{x}_j - x\|$. Here, in the first inequality, we use the triangle inequality $\|\hat{x}_j - \hat{x}_{k^+}\| \le \|\hat{x}_j - x\| + \|x - \hat{x}_{k^+}\|$, and the second inequality is because

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

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

Figure 3. (Color online) Visualization of the Shapley Saddle Point



3.2.2. Shapley Policy as the Solution to Slope Minimization. Next, we show geometrically in Figure 3 that the Shapley extension (S) is the closed-form solution to (L). To visualize $A_{jk}(x)$, we plot the d–y plane on which a point (d_k, y_k) means $\hat{x}_k$ 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 of the point kth horizon as shown in Figure 3(b). Minimizing over k gives the Shapley saddle point $f(x) = \min_k \max_j A_{jk}(x)$ (Figure 3(c)). Geometrically, because $L_{k_+} = \ \ L_{k_-}$, the shadow region in Figure 3(c) is a symmetric cone with the smallest opening that covers all the points $\{(d_k, y_k)\}_{k\in[K]}$. Thus, it is apparent that the minimax theorem holds by symmetry, and the vertex of such a smallest opening is precisely determined by the Lipschitz-minimization problem (L).

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

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

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

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

4. Discussions

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

4.1. Interpretation as Lipschitz Regularization

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

Consider the following Lipschitz regularization problem defined as

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

where we denote by $a_1 \vee a_2$ the maximum between a_1 and a_2 and by $\|\cdot\|_{\text{Lip}}$ 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 $1 \vee \|f\|_{\text{Lip}}$) and its expected in-sample cost. If we set $\rho = 0$, (R) degenerates to the nonrobust empirical risk minimization. In this case, the optimal policy f^* is defined only on the in-sample data, which equals the critical fractile of the empirical conditional distribution and can take any value outside $\hat{\mathcal{X}}$. If we prohibit perturbing any data in the z-direction, then the lower cutoff one of the Lipschitz norm $\|f\|_{\text{Lip}}$ 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 classic newsvendor problem without feature information

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

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

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

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

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

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

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

4.2. Linear Programming Reformulation for In-Sample Robust Problem

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

Proposition 2. *Problem ($\hat{\text{R}}$) is equivalent to*

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

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

4.3. Generalization Error Bound

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

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

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

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

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

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

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

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

4.4. Structure of the Optimal Robust Policy

In this section, 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 Online Appendix D.

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

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

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

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

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

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

Proposition 4.

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

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

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

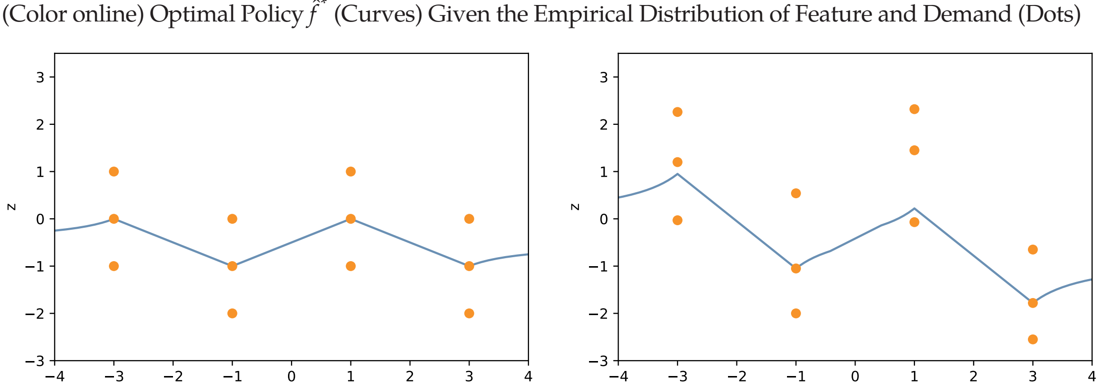
This discussion is illustrated in Figure 4. We have $b = 3$, $h = 2$, $\rho = 10$, $\mathcal{X} = \mathbb{R}$, $\hat{\mathcal{X}} = \{-3, -1, 1, 3\}$, and $\hat{\mathbb{P}}_n$ is supported on $n = 12$ points (as indicated by the dots in the x–z plot), in which each $\hat{x} \in \hat{\mathcal{X}}$ is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

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

5. Numerical Experiments

In this section, using synthetic and real data in Sections 5.1 and 5.2, respectively, we compare the empirical performance of our proposed `Shapley` policy against several

Figure 4. (Color online) Optimal Policy $\hat{f}^*$ (Curves) Given the Empirical Distribution of Feature and Demand (Dots)



Notes. If the empirical conditional critical fractile q is one-Lipschitz, then $\hat{f}^* = q$ (left). Otherwise, $\hat{f}^*$ regularizes q by reducing its Lipschitz norm (right).

benchmarks, including empirical risk minimization using an affine policy with ℓ^1 and ℓ^2 regularization (ERM2 (ℓ^1/ℓ^2)) and kernel-weights optimization (KO) in Ban and Rudin (2019), conditional stochastic optimization using random forests (RandForest) and k-nearest neighbors (kNN) in Bertsimas and Kallus (2020), and stochastic optimization forests with different splitting criteria (StochOptForest (apx-soln/apx-risk)) in Kallus and Mao (2022).

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

5.1. Synthetic Data

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

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

where the feature vector $X \in \mathbb{R}^d$ is a multivariate Gaussian random variable with mean zero and covariance matrix $(\sigma_{ij})_{d\times d}$ with $\sigma_{ij} = 0.995^{|i-j|}$; the coefficient parameter $\beta_0 = c(200, -200, 199, -199, \ldots, 1, -1, 0, 0, \ldots, 0)^\top$ is a d-dimensional vector with $d = 5{,}000$, where $c = 25/(\sum_{k=1}^{200} 2k^2)^{1/2}$ is chosen such that $\|\beta_0\|_2 = 25$ and $X^\top \beta_0$ is normally distributed with mean zero and roughly standard deviation one; and ε is standard Gaussian.

To study the impact of the extremity of the quantile, we vary $h \in \{0.2, 0.5, 1\}$, fixing $b = 1$. To study the impact of the sample size n, we vary the number of data points $n \in \{10, 100, 1{,}000\}$. The testing data size is fixed as 10,000. We use Euclidean norm for all distance-based approaches, namely, Shapley, KO, and kNN.

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

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

II. For benchmark methods, kNN is outperformed by the Shapley policy in most situations (eight out of nine box plots in terms of the median performance). The affine policy ERM2 (ℓ^1) and KO have a clear advantage over our Shapley when $n = 1{,}000$, $h = 1$. But for a small sample size ($n = 10$), they are outperformed by our Shapley. Forest-based methods (RandForest, StochOptForest (apx-soln)) have a relatively good performance when

Figure 5. (Color online) Box Plots of the Relative Differences in the Out-of-Sample Performance Between Shapley and Other Benchmarks (Arranged from Left to Right as in Legends)

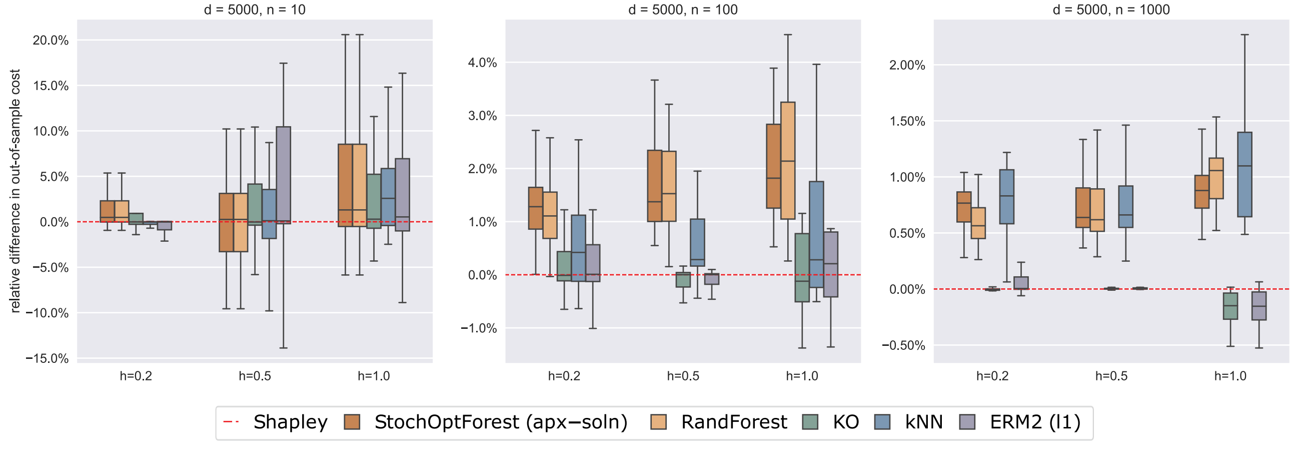


Table 1. Average Computational Time (in Seconds) per Problem Instance

	Average training time			Average testing time		
Model	$n = 10$	$n = 100$	$n = 1{,}000$	$n = 10$	$n = 100$	$n = 1{,}000$
`Shapley`	0.0089	0.2591	20.9073	0.8290	2.5210	28.8658
k`NN`	0.0003	0.0010	0.0092	98.4019	134.6318	496.1717
`KO`	0.0003	0.0010	0.0101	100.7559	136.6974	500.7450
`ERM2` (ℓ^1)	1.6626	3.0285	23.7068	8.6050	8.6662	8.5541
`RandForest`[a]	0.0877	21.1517	479.5995	19.8414	25.0150	28.6125
`StochOptForest (apx-soln)`[a]	0.0880	24.1451	536.7108	19.7084	24.9899	28.4322

[a]The algorithm runs with a reduced version (see (II)).

the sample size is small. We note, however, that the performance plotted in Figure 5 may not show the full strength of the forest-based methods because we run a reduced version of the algorithm because of the time limit; see (II) as follows. Moreover, even under this case, the running time of `StochOptForest (apx-risk)` still exceeds the time threshold, so we are unable to show its performance under this experiment, but do so later using a real data set.

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

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

II. For the average testing time, affine policy `ERM2` (ℓ^1) takes the least time for testing when the testing size is large ($n = 1{,}000$) and is independent of the training sample size by construction. The average testing time of `KO` and k`NN` roughly grows linearly in the training sample size n (subtracting some overhead time) as their theoretical bounds are $O(n)$. We find empirically that the testing time of `Shapley` grows sublinearly in the sample size using our equivalent formulation (6), which is faster than computing the matrix saddle point naïvely. In this experiment, forest-based methods (`RandForest`, `StochOptForest (apx-soln)`) are run with a reduced version, which reduces the number of trees by randomly selecting one 10th of the features for the candidate splits as we set a time threshold of one hour for each instance. Comparisons with the full version of the algorithms are conducted in Section 5.2.

5.2. Real Data

We test the performance using a real-world data set considered in Oroojlooyjadid et al. (2020) modified from Pentaho (2008), which consists of historic 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 (`ERM2` (ℓ^1/ℓ^2)) and forest-based (`RandForest`, `StochOptForest (apx-soln/apx-risk)`) models, these feature variables are converted into 43 dummy variables, so the dimension is $d = 40$ (43 dummy variables minus 3 reference categories). For distance-based methods (`Shapley`, k`NN`, `KO`), the metric on $\mathcal{X}$ is defined as the following. Let $x_k = (i_k, m_k, d_k) \in \mathcal{X}$, $k = 1, 2$, where $i_k = 1, \ldots, 24$, $m_k = 1, \ldots, 12$, $d_k = 1, \ldots, 7$ represent department ID, month, and weekday, respectively, and we define the distance on $\mathcal{X}$ as

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

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

The raw data set contains 9,877 data for training and 3,293 data for testing. To generate the training data sets for our purpose, in each of the repeated experiments, we randomly draw n distinct samples from the raw training data set to train the model. The out-of-sample cost is defined as the average newsvendor cost over the full set of the testing data. To investigate in different

Table 2. Average Out-of-Sample Costs with 95% Confidence Intervals (Model with Lowest Average Cost is Bolded)

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

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 $h \in \{0.2, 0.5, 1\}$, fixing the unit back-order cost $b = 1$.

The average out-of-sample costs and their 95% confidence intervals are shown in Table 2. We observe that our `Shapley` policy has a superior performance compared 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, $h/b = 0.2$.

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

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

II. Compared with nonparametric distance-based approaches `KO` and *k*`NN`, in terms of the median out-of-sample costs, `Shapley` outperforms both of them in eight out of nine 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 nine cases. Note that, unlike the synthetic experiments, here, we use the full version of these methods.

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

6. Concluding Remarks

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

Figure 6. (Color online) Box Plots of the Differences in the Out-of-Sample Performance Between `Shapley` and Other Benchmarks

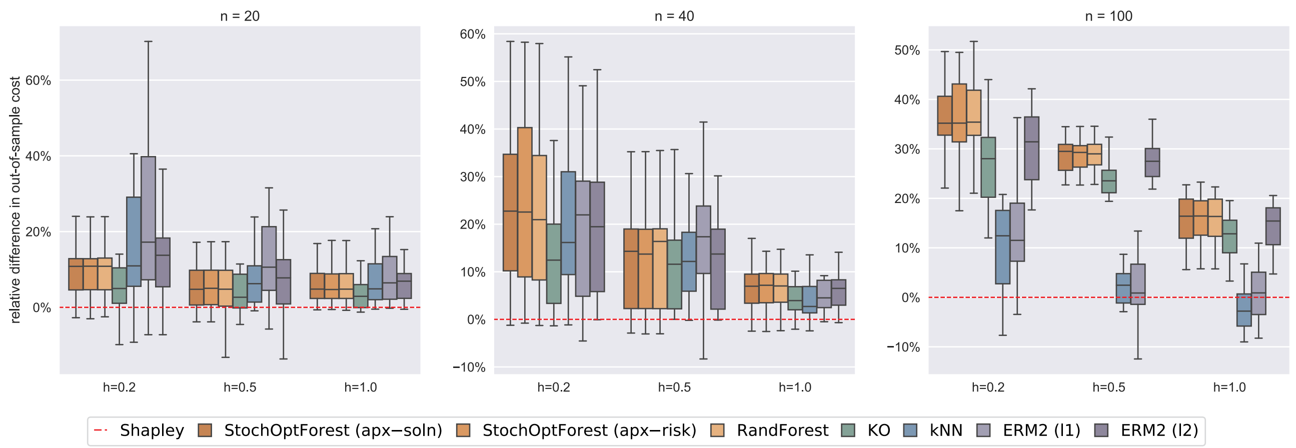


Table 3. Average Computational Time (in Seconds) per Problem Instance

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

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

References

An Y, Gao R (2021) Generalization bounds for (Wasserstein) robust optimization. *Adv. Neural Inform. Processing Systems* 34: 10382–10392.

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

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

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

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) *Robust Optimization*, vol. 28 (Princeton University Press, Princeton, NJ), 365–370.

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

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

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

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

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

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

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

Bertsimas D, McCord C, Sturt B (2023) Dynamic optimization with side information. *Eur. J. Oper. Res.* 304(2):634–651.

Carrizosa E, Olivares-Nadal AV, Ramírez-Cobo P (2016) Robust newsvendor problem with autoregressive demand. *Comput. Oper. Res.* 68:123–133.

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

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

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

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

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

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

Esteban-Pérez A, Morales JM (2022) Distributionally robust stochastic programs with side information based on trimmings. *Math. Programming* 195(1–2):1069–1105.

Fu M, Li X, Zhang L (2021) Data-driven feature-based newsvendor: A distributionally robust approach. Preprint, submitted July 15, https://dx.doi.org/10.2139/ssrn.3885663.

Gallego G, Moon I (1993) The distribution free newsboy problem: Review and extensions. *J. Oper. Res. Soc.* 44(8):825–834.

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

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

Gao R, Kleywegt A (2022) Distributionally robust stochastic optimization with Wasserstein distance. *Math. Oper. Res.* 48(2):603–655.

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

Georghiou A, Tsoukalas A, Wiesemann W (2021) *On the Optimality of Affine Decision Rules in Robust and Distributionally Robust Optimization* (Optimization Online).

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

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

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

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

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

Hu Y, Kallus N, Mao X (2022) Fast rates for contextual linear optimization. *Management Sci.* 68(6):4236–4245.

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

Kallus N, Mao X (2022) Stochastic optimization forests. *Management Sci.* 69(4):1975–1994.

Kannan R, Bayraksan G, Luedtke JR (2020a) *Data-Driven Sample Average Approximation with Covariate Information* (Optimization Online).

Kannan R, Bayraksan G, Luedtke JR (2020b) Residuals-based distributionally robust optimization with covariate information. Preprint, submitted December 2, https://arxiv.org/abs/2012.01088.

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

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

Lee S, Kim H, Moon I (2020) A data-driven distributionally robust newsvendor model with a Wasserstein ambiguity set. *J. Oper. Res. Soc.* 72(8):1879–1897.

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

Loke GG, Tang Q, Xiao Y (2020) Decision-driven regularization: Harmonizing the predictive and prescriptive. Preprint, submitted June 17, https://dx.doi.org/10.2139/ssrn.3623006.

Luxburg UV, Bousquet O (2004) Distance-based classification with Lipschitz functions. *J. Machine Learn. Res.* 5:669–695.

Meng Q, Shi Y, Qi Y, Ma C, Yuan R, Wu D, Shen ZM (2022) A practical end-to-end inventory management model with deep learning. *Management Sci.* 69(2):759–773.

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

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

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

Pentaho (2008) Foodmart's database tables. Accessed January 1, 2021, http://pentaho.dlpage.phi-integration.com/mondrian/mysql-foodmart-database.

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

Postek K, den Hertog D (2016) Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS J. Comput.* 28(3):553–574.

Rahimian H, Bayraksan G, Homem-de Mello T (2019a) Controlling risk and demand ambiguity in newsvendor models. *Eur. J. Oper. Res.* 279(3):854–868.

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

Scarf H (1958) *A Min-Max Solution of an Inventory Problem* (RAND Corporation, Santa Monica, CA).

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

Shafieezadeh-Abadeh S, Kuhn D, Esfahani PM (2019) Regularization via mass transportation. *J. Machine Learn. Res.* 20(103):1–68.

Shapiro A, Dentcheva D, Ruszczyński A (2014) *Lectures on Stochastic Programming: Modeling and Theory* (SIAM, Philadelphia).

Shepard D (1968) A two-dimensional interpolation function for irregularly spaced data. *Proc. 23rd ACM National Conf.*, 517–524.

Srivastava PR, Wang Y, Hanasusanto GA, Ho CP (2021) On data-driven prescriptive analytics with side information: A regularized Nadaraya–Watson approach. Preprint, submitted October 10, https://arxiv.org/abs/2110.04855.

Sturt B (2021) A nonparametric algorithm for optimal stopping based on robust optimization. Preprint, submitted March 4, https://arxiv.org/abs/2103.03300.

Subramanyam A, Gounaris CE, Wiesemann W (2019) K-adaptability in two-stage mixed-integer robust optimization. *Math. Programming Comput.* 12:193–224.

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

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

Xin L, Goldberg DA (2021) Time (in) consistency of multistage distributionally robust inventory models with moment constraints. *Eur. J. Oper. Res.* 289(3):1127–1141.

Xin L, Goldberg DA (2022) Distributionally robust inventory control when demand is a martingale. *Math. Oper. Res.* 47(3):2387–2414.

Yanıkoğlu İ, Gorissen BL, den Hertog D (2019) A survey of adjustable robust optimization. *Eur. J. Oper. Res.* 277(3):799–813.

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

Zhu L, Huang M, Li R (2012) Semiparametric quantile regression with high-dimensional covariates. *Statistica Sinica* 22(4):1379–1401.

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