The Dissertation Committee for Luhao Zhang
certifies that this is the approved version of the following dissertation:

Quantitative and Modeling Aspects of Optimal Decision Making under Uncertainty

Committee:

Thaleia Zariphopoulou, Supervisor

Irene M. Gamba

Rui Gao

Mihai Sirbu

Efstathios Tompaidis

Quantitative and Modeling Aspects of Optimal Decision Making under Uncertainty

by

Luhao Zhang

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2023

Dedication

Dedicated to my soulmate Jincheng and my beloved family.

Acknowledgments

I would like to express my deepest gratitude and appreciation to the following individuals who have contributed immensely to the completion of this dissertation and my whole Ph.D. study:

My supervisor, Professor Thaleia Zariphopoulou, whose guidance, expertise, and unwavering support have been invaluable throughout this research journey. Her class in stochastic control is the start of my research. Her enthusiasm opens this new world to me. Her insightful guidance and encouragement have been instrumental in shaping this dissertation and my academic career.

I am also grateful to the members of my dissertation committee, Professors Irene Gamba, Rui Gao, Mihai Sirbu, and Stathis Tompaidis, for their time, expertise, and valuable input. Their thoughtful suggestions and critical review have significantly enhanced the quality of this dissertation.

Besides my thesis committee, I would like to thank the outstanding faculty members in the department of Mathematics. Their exceptional courses have provided me with the necessary tools and understanding to undertake this research. Their expertise, passion, and guidance have continuously inspired and challenged me to push my boundaries.

I also want to extend my heartfelt thanks to my coauthors, Professors Rui Gao and Renyuan Xu, for their support, patience, and understanding. Their belief in me and constant encouragement have provided me the strength to overcome challenges and persevere during this demanding process. They are like my big brother and sister, and I am truly fortunate to have them throughout my academic life. I am also grateful that I have met many supportive senior and junior researchers in this area, like Professors Ming Hu, Ningyuan Chen, Junyu Cao, Jiashuo Jiang and Xian Yu. Their passion, brilliance, and collaborative spirit have been invaluable in my journey.

My family and friends have been my pillars of strength all the time. To my parents and young little sister, despite the physical separation caused by the pandemic, their love and encouragement have transcended distance and time. Their constant presence in my heart has given me the strength to keep moving. I am lucky to have my dearest friends here, Chenlu Wang, Dao Reyes, Shiyu Liang and Yuying Wang. Their words of encouragement, and willingness to listen to my ideas, frustrations, and triumphs have helped me survive every hard times I had.

To my dearest love, Jincheng Yang, thank you for being in my life. You are the best thing that's ever been mine.

Abstract

Quantitative and Modeling Aspects of Optimal Decision Making under Uncertainty

Luhao Zhang, Ph.D.
The University of Texas at Austin, 2023

SUPERVISOR: Thaleia Zariphopoulou

This dissertation focuses on the problem of decision making under uncertainty, more precisely, the quantitative and modeling aspects of "*how to acquire and, in turn, exploit information optimally for decision-making in stochastic environments*". To address the challenges posed by different types of uncertainty, a range of methods have been developed in the fields of stochastic control under partial information, dynamic information acquisition, data-driven optimization, model uncertainty, and robust optimization.

Specifically, this dissertation is composed by two parts:

The first part focuses on an *offline data-driven decision-making problem with side information.* With abundant data routinely collected in many industries to support decision-making, historical data with numerous side information–temporal, spatial, social, or economical–are available prior to the decision making and reveals partial information on the randomness of the problem. The challenge for these high-dimensional problems is that the empirical distribution constructed from the observed

data is not representative of the underlying true distribution between contextual information and decisions, and strategies solely based on the empirical data can lead to poor performance when implemented. Therefore, a fundamental problem in data-driven decision-making under uncertainty, as well as in statistical learning, is finding solutions that perform well not only on the observed data but also on new and previously unseen data. To hedge against the distributional uncertainty of the offline dataset, this dissertation provides an end-to-end learning framework, based on *distributionally robust stochastic optimization* (DRSO), that prescribes a non-parametric policy with certified robustness, provable optimality, and efficient implementation. Specifically, we study policy optimization for a series of feature-based decision-making problems, which seeks an end-to-end policy that renders an explicit mapping from features to decisions.

In this dissertation, we first consider a Wasserstein robust optimization framework, where we highlight our contribution in finding an optimal robust policy without restricting onto a parametric family while still maintaining computational efficiency and interpretability. More specifically, by exploiting the structure of the optimal policy, we identify a new class of policies that are proven to be robust optimal and can be computed by linear programming. We apply our work in newsvendor problem.

Furthermore, we propose a new uncertainty set based on causal transport distance which contains distributions that share a similar conditional information structure with the nominal distribution. We derive a tractable dual reformulation for evaluating the worst-case expected cost and show that the worst-case distribution has a similar conditional information structure as the nominal distribution. We identify tractable cases to find the optimal decision rules over an affine class or the entire nonparametric class, and apply our work in conditional regression, incumbent pricing and portfolio selection.

The second part is concerned with *dynamic information acquisition with sequential decision-making and differential information sources.* When involving dynamic learning to facilitate decision making, since the decision makers often have imperfect and costly information, they encounter a trade-off between the information learning and the expected payoff, given the limited information. For example, when comparing new technologies, the firm often spends a considerable amount of funds and time on research and development (R&D) to identify the best technology to adopt. Other examples include investors designing algorithms to learn about the return of different assets, scientists conducting research to investigate the validity of different hypotheses, etc. From the viewpoint of dynamic information acquisition, the practically important features are the choice of "what to learn", as well as "when to learn and stop learning".

Most of the decision-making problems considered in this line of work are static (i.e. a single, irreversible decision) problems which, however, over-simplify the structure of many real-world applications that require dynamic or sequential decisions. Moreover, the information acquisition source in these studies typically remains constant (e.g. a single noise signal) throughout the decision process, failing to capture the adaptive nature of decision makers in response to stochastically changing environments.

Herein, we introduce a general framework in which we allow for *both* sequential (possibly reversible) decisions and dynamically changing information sources (distinct signals), and it also includes the cost of acquiring information across time. We analyze a benchmark example, motivated by the return/exchange policies in e-commerce platforms. Specifically, we introduce a sequential decision-making problem that allows decision makers to reverse their initial decisions and their costly information acquisition setting to change accordingly. We investigate the optimal strategies for information acquisition and decision reversal, and carry out a complete sensitivity and asymptotic analysis on how decision makers can effectively adapt their learning behavior to ultimately achieve the best decision-making outcomes.

In what follows, we describe each approach separately. For each part, we introduce the corresponding model, construct solutions, and provide a detailed analytical methodology.

Table of Contents

Chapter 1: Data-driven Decision Making and Distributionally Robust Stochastic Optimization

This chapter is based on Zhang et al. (2023), where I made significant contributions by initiating and defining the research problem, conducting the research itself that resulted in the paper, and composing the manuscript.

1.1 Distributionally Robust Stochastic Optimization (DRSO) with Side Information

Uncertainty is a common challenge faced in various fields such as science, engineering, and management, which has generated significant theoretical and practical interest among research communities including operations research, statistics, and machine learning. Over time, various methods such as stochastic optimization, robust optimization, and dynamic programming have been developed to formulate, analyze, and solve these problems. Typically, these optimization models represent uncertainty through probability distributions that are assumed to be accurately estimated from learning and observation.

With the advance in big data technology, abundant feature values are routinely collected in many industries to support decision-making. More often than not, numerous side information–temporal, spatial, social, or economical–are available prior to the decision-making and reveals partial information on the randomness of the problem. The side information reduces the uncertainty and helps the decision-maker customize decisions. It is crucial to involve the feature information in decision-making. Otherwise, the decision may be inconsistent, namely, not converging to the true optimal policy even with an infinite amount of data Ban and Rudin (2019). The ordering decision is often made for a population of features but not for a single feature realization. For example, ordering decisions are made for multiple shelves at different locations in selected time windows, or for a number of customers with different demographics. In these cases, it would make sense to consider the average performance over the entire distribution of features. Thus in this thesis, we consider the stochastic optimization problem with side information, which also known as contextual optimization or conditional stochastic optimization, shown as follows:

$$\min_{w \in \mathcal{D}} \mathbb{E}[\Psi(w, Z) \mid X = x]. \tag{1.1}$$

It attempts to find a decision w from the feasible region $\mathcal{D}$ to minimize the conditional expectation of the cost $\Psi(w, Z)$ dependent on the decision w and a random variable Z, given some side information, represented by a covariate X. More informed or personalized decisions can be made with the side information revealed from the covariate data. This problem has received increasing attention nowadays as more side information becomes available to assist the decision making in e-commerce, online platform, etc. Quite often, while the decision is made based on the covariate, the performance is evaluated for the covariate population — for example, the manager in an e-commerce company cares about the overall performance across all customer types. By averaging over these covariate values, we are interested in finding a policy that minimizes the expected cost over the joint distribution of the covariate X and the random variable Z:

$$\min_{f \in \mathcal{F}} \mathbb{E}[\Psi(f(X), Z)]. \tag{1.2}$$

The policy offers an end-to-end map from the covariate space $\mathcal{X}$ to the decision space $\mathcal{D}$, chosen from a family $\mathcal{F}$ of functions—parametric or non-parametric—on $\mathcal{X}$. The choice of $\mathcal{F}$ can vary from small parametric classes like affine policies, to large non-parametric classes and even all measurable functions.

The formulation (1.2) covers many contextual optimization problems in operations research and machine learning. For instance, suppose $\Psi(w, z) = h(w - z)_+ + b(z - w)_+$, where z represents the demand of a product and $h, b \geq 0$ represent the holding cost and the backorder cost respectively, then (1.2) is known as the big-data newsvendor model Ban and Rudin (2019). If $\mathcal{F}$ is the set of all measurable functions on $\mathcal{X}$, then the optimal order quantity equals the conditional critical fractile $f^\star(x) = F_x^{-1}(\frac{b}{h+b})$, where F_x is the conditional cumulative distribution function of demand Z given $X = x$; and if $\mathcal{F}$ is the set of affine functions on $\mathcal{X}$, then (1.2) finds the optimal affine policy for the big-data newsvendor. As another example, when $\Psi(w, z) = (w - z)^2$ and $\mathcal{F}$ is the set of all measurable functions on $\mathcal{X}$, the optimal solution to (1.2) is $f^\star(x) = \mathbb{E}[Z|X = x]$ and thus the formulation (1.2) finds the conditional mean of Z given X.

Similar to the classical stochastic optimization, the underlying joint distribution $\mathbb{P}_{\text{true}}$ of (X, Z) is often not known exactly, but instead, historical data from the underlying distribution are available. As such, it is reasonable to consider a data-driven distributionally robust contextual decision-making framework

$$\min_{f \in \mathcal{F}} \max_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X), Z)], \tag{1.3}$$

a minimax formulation that hedges the data uncertainty. At the core of the distributionally robust formulation is the choice of the uncertainty set, and the presence of the side information adds new challenges beyond those for classic stochastic optimization. Below, in Section 1.2, we review some existing choices of uncertainty sets.

1.2 Discussion on Some Existing Uncertainty Sets

A central question for the above framework is: how to choose a good set of distributions $\mathfrak{M}$ to hedge against? A good choice of $\mathfrak{M}$ should take into account the properties of the practical application as well as the tractability of problem (1.3). Two typical ways of constructing $\mathfrak{M}$ are moment-based and distance-based.

The moment-based approach considers distributions whose moments (such as mean and covariance) satisfy certain conditions Scarf (1958); Zymler et al. (2013); Popescu (2007); Delage and Ye (2010). In various cases, it has been demonstrated that the resulting DRO (Distributionally Robust Optimization) problem can be expressed as a semi-definite or conic quadratic program. However, the moment-based approach rests on the assumption that certain moment conditions are precisely known, while no other information about the relevant distribution is available. This assumption is rarely applicable in practice, as in most situations, one either has access to data from repeated observations of the variable Z or no data at all, and in both cases, the moment conditions do not provide a complete description of what is known about Z. Furthermore, the worst-case distributions obtained by this approach may result in overly conservative decisions.

The distance-based approach considers distributions that are close, in the sense of a chosen statistical distance, to a nominal distribution, such as an empirical distribution or a fitted Gaussian distribution. Two classes of distance-based uncertainty sets have been studied frequently in the literature. The first class is the divergence family, deeply rooted in statistics, information theory, and physics. Consider the following example.

Example 1 (KL robust solution is degenerate)**.** *Suppose* $\mathfrak{M}$ *is a Kullback-Leibler*

(KL) divergence ball, centered at the empirical distribution $\widehat{\mathbb{P}}$ constructed from K independently and identically distributed (i.i.d.) samples from a continuous underlying distribution. Then with probability one, $\widehat{\mathbb{P}}$ can be represented as $\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{\delta}_{(x_k,z_k)}$, where K is the sample size and all $(\widehat{x}_k,\widehat{z}_k)$'s are different from each other. Let $\mathcal{F}$ be the set of all measurable functions on $\mathcal{X}$. Then we claim that the KL robust optimal solution would satisfy

$$f_{\mathsf{kl}}(x) = \begin{cases} \arg\min_{w\in\mathcal{D}} \ \Psi(w,\widehat{z}_k), & \text{if } x = \widehat{x}_k,\ k = 1,\ldots,K, \\ \text{arbitrary value}, & \text{otherwise.} \end{cases}$$

Indeed, every distribution in the KL ball can be supported only on in-sample data, but differ from $\widehat{\mathbb{P}}$ in the probability weights. On an in-sample data point $\widehat{x}_k$, regardless of its weight, the optimal decision would always be the minimizer of $\Psi(\cdot,\widehat{z}_k)$ due to interchangeability principle Shapiro et al. (2014). Furthermore, since the KL robust cost depends only on the function values on the in-sample data, the robust optimal solution can take any value on out-of-sample data without changing the objective value.

♣

Example 1 shows that the KL robust optimal policy is degenerate with probability one when the underlying distribution is continuous, regardless of the size of the uncertainty set, the sample size, or the objective function. A similar phenomenon also holds for all other divergence measures, due to the structure of the worst-case distribution Bayraksan and Love (2015).

In this thesis, we would like to focus on distance-based uncertainty sets, more precisely, the distributional distance defined through optimal transport. A popular choice is Wasserstein, or transport cost distance, family. It is well-known that the resulting uncertainty set avoids some degeneracy issues of the divergence sets in stochastic optimization Kuhn et al. (2019); Gao and Kleywegt (2016).

1.2.1 Wasserstein DRSO

Consider any underlying metric d on Ξ which measures the closeness of any two points in Ξ. Let $p \geq 1$, and let $\mathcal{P}(\Xi)$ denote the set of Borel probability measures on Ξ. The Wasserstein distance of order p between two distributions $\mu, \nu \in \mathcal{P}(\Xi)$ is defined as

$$\mathsf{W}_p(\mu, \nu) := \inf_{\gamma \in \mathcal{P}(\Xi^2)} \mathbb{E}^{1/p}_{(\xi, \zeta \sim \gamma)} [d^p(\xi, \zeta) : \gamma \text{ has marginal distributions } \mu, \nu],$$

Given a nominal distribution ν and a radius $\rho > 0$, we are interested in solving

$$\min_{f \in \mathcal{F}} \max_{\mu \in \mathcal{P}(\Xi)} \mathbb{E}_\mu [\Psi(f(X), Z)], \tag{1.4}$$

In the contextual decision-making formulation (1.3), let $\mathcal{P}_1(\mathcal{X} \times \mathcal{Z})$ be the set of probability distributions on $\mathcal{X} \times \mathcal{Z}$ with finite first moment. The *Wasserstein distance* (of order p) is defined as

$$\mathsf{W}_p(\mathbb{P}, \mathbb{Q}) := \left(\inf_{\gamma \in \Gamma(\mathbb{P}, \mathbb{Q})} \mathbb{E}_{(\tilde{X}, \tilde{Z}), (X, Z) \sim \gamma} \left[\|\tilde{X} - X\|^p_{\mathcal{X}} + \|\tilde{Z} - Z\|^p_{\mathcal{Z}} \right] \right)^{1/p},$$

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

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

where ρ is the radius of the 1-Wasserstein ball where we use to construct the uncertainty set. More detailed explanation and discussion on the above formulation will be presented in Section 1.4.2.

Wasserstein distance and the related field of optimal transport, which is a generalization of the transportation problem, have been studied in depth. In the stochastic optimization literature, Wasserstein distance has been used for single stage stochastic optimization Wozabal (2012, 2014), and for multistage stochastic optimization Pflug and Pichler (2014). The challenge for solving (1.3) is that, the inner maximization involves a supremum over possibly an infinite dimensional space of distributions. To tackle this problem, existing works focus on the setup when the nominal distribution is the empirical distribution on a finite-dimensional space. Recently, using duality theory showed that under certain conditions, the inner maximization problem of (Wasserstein-DRSO) is actually equivalent to a finite-dimensional convex problem. Constructive proofs Esfahani and Kuhn (2018); Blanchet and Murthy (2019); Zhao and Guan (2018); Sinha et al. (2018); Zhen et al. (2021) rely on advanced convex duality theory. More specifically, Esfahani and Kuhn (2018); Zhao and Guan (2018); Zhen et al. (2021) exploit advanced conic duality Shapiro (2001) for the problem of moments that requires the nominal distribution $\widehat{\mathbb{P}}$ to be finitely supported and the space $\mathcal{X}$ to be convex, along with some other assumptions on the transport cost c and the loss function f; Blanchet and Murthy (2019) use an approximation argument that represents the Polish space $\mathcal{X}$ as an increasing sequence of compact subsets, on which the duality holds for any Borel distribution $\widehat{\mathbb{P}}$ thanks to Fenchel conjugate on vector spaces Luenberger (1997), under certain semicontinutiy assumptions on the transport cost c and loss function f; using the same infinite dimensional convex duality, (Sinha et al., 2018, Theorem 5) provide a simplified analysis by assuming the function $(\widehat{X}, X) \mapsto \lambda c(\widehat{X}, X) - f(X)$ is a normal integrand Rockafellar and Wets (2009). Compared with these non-constructive duality proofs, our (non-constructive) proof uses only Legendre transform, namely, the convex duality for univariate real-valued functions. The constructive proof developed by Gao and Kleywegt (2022) provides a result at a similar level of generality as Blanchet and Murthy (2019) without using convex duality theory, by constructing an approximately worst-case distribution using the first-order optimality condition of the weak dual problem. In Zhang et al. (2022), their analysis is shorter and more elementary.

1.2.2 Causal DRSO

Causal transport distance and its associated optimal transport problem were introduced in Lassalle (2013), whose main motivation is to investigate optimal transportation problems with filtrations and their applications to stochastic calculus. The discrete-time counterpart was investigated in Backhoff et al. (2017). The definition of causal transport distance, specialized to our considered setting, is as follows.

Definition 1 (Causal Transport Distance). *A joint distribution* $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$ *is called a* causal transport plan *if for* $((\widehat{X}, \widehat{Z}), (X, Z)) \sim \gamma$, X *and* $\widehat{Z}$ *are conditionally independent given* $\widehat{X}$*:*

$$X \perp\!\!\!\perp \widehat{Z} \mid \widehat{X}.$$

We denote by $\Gamma_c(\widehat{\mathbb{P}}, \mathbb{P})$ *the set of all transport plans* $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$ *that are causal. Let* $p \in [1, \infty)$*. The* p-causal transport distance *between* $\widehat{\mathbb{P}}$ *and* $\mathbb{P}$ *is defined as*

$$\mathcal{C}_p(\widehat{\mathbb{P}}, \mathbb{P}) := \left(\inf_{\gamma \in \Gamma_c(\widehat{\mathbb{P}}, \mathbb{P})} \mathbb{E}_{((X,Z),(\widehat{X},\widehat{Z}))\sim\gamma} \left[\|X - \widehat{X}\|_{\mathcal{X}}^p + \|Z - \widehat{Z}\|_{\mathcal{Z}}^p \right] \right)^{1/p}. \qquad \diamond$$

Like Wasserstein distance, causal transport distance finds the minimal transport cost between two distributions, where norms capture the geometry of the data space and similarity between samples. Nevertheless, causal transport distance differs from Wasserstein distance in the involved class of transport plans: Wasserstein distance considers all transport plans with given marginals while causal transport distance restrict on causal transport plans as defined in Definition 1. A further discussion and interpretation of Causal transport distance will be discussed in Chapter 2.

Here is a toy example to show the difference between Causal distance and Wasserstein distance.

Example 2 (Wasserstein set cannot capture conditional information)**.** *In Figure 1.1, $\widehat{\mathbb{P}}$ and $\mathbb{P}$ are two uniform distributions supported respectively on the blue and green line segments with a common endpoint with x-entry being $\widehat{x}$. The angle between the two line segments is ε radian. Notably, the conditional distribution $\mathbb{P}^{\varepsilon}_{Z|X=x}$ is a Dirac*

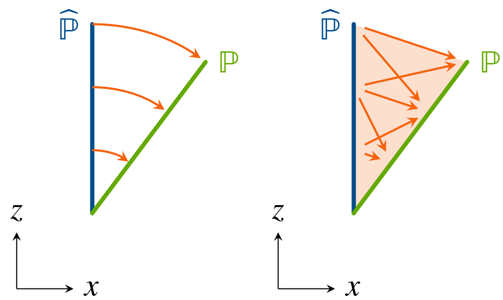


Figure 1.1: $\widehat{\mathbb{P}}$ and $\mathbb{P}$ has completely different conditional information structure but with $O(\varepsilon)$ Wasserstein distance.

measure for $x > \widehat{x}$, which is apparently very different from the conditional distribution $\widehat{\mathbb{P}}_{Z|X=\widehat{x}}$. As will be calculated in Section 2.1, Wasserstein distance between $\widehat{\mathbb{P}}$ and $\mathbb{P}$ is $O(\varepsilon)$, and the optimal transport maps is a rotation. This means a Wasserstein ball centered at $\widehat{\mathbb{P}}$ would always contain a distribution that has a different conditional information structure than that of $\widehat{\mathbb{P}}$ regardless of the value of ε. But when considering causal transport plan (shown in the right plot), which is the independent joint distribution $\widehat{\mathbb{P}} \otimes \mathbb{P}^{\varepsilon}$, it will lead to a large causal transport distance. ♣

In practice, the following situation is often seen from data: the conditional distribution can be estimated accurately under a number of covariate values, but is largely unobserved for other values. For example, historical data may reveal an accurate estimate of the conditional demand distribution of the product sold at deployed vending machines, but the demand at some new location is unexplored. Nonetheless, it is conceivable that the conditional demand distribution should share some resemblance among similar locations. In such cases, it would be reasonable to expect that the conditional distributions $\mathbb{P}_{Z|X=x}$ and $\mathbb{P}_{Z|X=\widehat{x}}$ corresponding to two similar values x and $\widehat{x}$ should be close in a certain way. Therefore, we would like to choose an uncertainty set containing distributions that share a similar conditional information structure with the nominal distribution. Example 2 demonstrates that the Wasserstein uncertainty set fails to preserve the conditional information structure, and in fact, we will show in Chapter 2 that this phenomenon also holds for the worst-case distribution. This raises the concern of overly conservativeness of Wasserstein robust solutions.

1.3 Related Literature

On stochastic optimization with side information. In the literature, the frameworks for contextual optimization (with an offline data set) can be broadly classified into three categories: *separate prediction and optimization*, *conditional stochastic optimization*, and *optimization over policys*.

(I) Separate prediction and optimization is a classical two-step process that first estimates a conditional distribution of Z given a new context $X = x$, and then optimizes for the conditional expectation $\min_{w\in\mathcal{D}} \mathbb{E}[\Psi(w, Z)|X = x]$ (e.g., Toktay and Wein (2001); Zhu and Thonemann (2004)). There are some theoretical guarantees in this approach discussed in El Balghiti et al. (2019); Hu et al. (2022). One main issue of this framework, as discussed in Liyanage and Shanthikumar (2005); Ban and Rudin (2019), is that the statistical estimation error and model mis-specification error may propagate to the decision optimization model and thus lead to a sub-optimal performance. Recent developments in contextual decision-making highlights the need for integrating the prediction and optimization.

(II) Conditional stochastic optimization avoids estimating the conditional distribution by directly estimating the conditional expected objective $\mathbb{E}[\Psi(w, Z)|X = x]$. Various estimation approaches have been studied, for example, based on Dirichlet process Hannah et al. (2010), Nadaraya-Watson kernel regression Hanasusanto and Kuhn (2013); Ban and Rudin (2019); Srivastava et al. (2021), local regression and classification Bertsimas and Kallus (2020); Bertsimas and McCord (2019), smart prediction-then-optimization Elmachtoub and Grigas (2021); El Balghiti et al. (2019); Elmachtoub et al. (2020); Ho-Nguyen and Kılınç-Karzan (2022), trees and forests Ban et al. (2019); Kallus and Mao (2020), robustness optimization and regularization Tulabandhula and Rudin (2013); Zhu et al. (2021); Bertsimas and Van Parys (2021); Loke et al. (2020); Esteban-Pérez and Morales (2020), regret minimization Estes (2021), empirical residuals Kannan et al. (2020a,b), bilevel optimization Muñoz et al. (2022); Cao and Gao (2021), etc. This approach requires solving a decision optimization problem for each individual context, which could be computationally prohibitive when numerous contexts are presented.

(III) Optimization over policys is an end-to-end formulation which finds a policy prescribing the decision for every possible context. Due to the computational difficulty of this infinite-dimensional optimization, typically the policies are parameterized by a finite dimensional vector, such as coefficients in an affine function of features Brandt et al. (2009); Ban and Rudin (2019); Bazier-Matte and Delage (2020); Bertsimas et al. (2022) or in a reproduce kernel Hilbert space Bertsimas and Koduri (2022) and weight matrices in a neural network Oroojlooyjadid et al. (2020); Qi et al. (2021); Liu et al. (2021). Our formulation falls into this category, but our results in Section 2.3 do not restrict the class of policys on a parametric family. In this respect, the closest work to ours is Zhang et al. (2023), which considers robust optimization over policys with Wasserstein uncertainty set; see the last paragraph of this subsection for a detailed comparison.

We remark that in online setting, stochastic optimization with side information has also been considered under the umbrella of contextual bandits and reinforcement learning, which are beyond the scope of this paper.

On distributionally robust optimization. Distributionally robust optimization (DRO) has received significant attentions recently as a tool for decision-making under uncertainty, and different approaches mainly differ in how the uncertainty set is constructed. Our choice of uncertainty set is aligned with DRO with transport distance, such as Wasserstein distance Pflug and Wozabal (2007); Wozabal (2012); Esfahani and Kuhn (2018); Blanchet and Murthy (2019); Blanchet et al. (2019); Gao and Kleywegt (2016); Gao et al. (2017); Gao (2020) and nested distance Analui and Pflug (2014); Pichler and Shapiro (2021); Rüschendorf (1985). To our best knowledge, our distributionally robust formulation based on the causal transport distance has not been studied in the literature. We refer to Rahimian et al. (2017) for a thorough review on other choices of uncertainty set.

On decision-rule approach in adjustable robust optimization. In the literature for adjustable robust optimization, different choices of policys have been thoroughly investigated, 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. (2022); Georghiou et al. (2021), k-adaptability Hanasusanto et al. (2015b, 2016); Subramanyam et al. (2019), iterative splitting of uncertainty sets Postek and Hertog (2016), binary policys Bertsimas and Georghiou (2015), non-parametric Markovian stopping rules Sturt (2021), etc. Most of these works do not consider covariate in their problem. Bertsimas et al. (2022) consider dynamic decision-making with side information using affine policys where as we consider general policys in a static setting; and Zhang et al. (2023) consider the newsvendor problem with Wasserstein distance, whereas we consider a different uncertainty set, and we adopt a completely different proof strategy and obtain a broader class of optimal policies for adjustable robust optimization that encapsulates the Shapely policy proposed therein.

1.4 Optimal Robust Policy for Feature-based Newsvendor

In the rest of the chapter, we delve into the data-driven decision-making problem under uncertainty using the Wasserstein distributionally robust framework. Specifically, we focus on its application of the feature-based newsvendor model.

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

If the underlying distribution is known, then the true optimal policy would equal the conditional critical fractile of the demand distribution under each feature value. Unfortunately, many real-world problems comprise a potentially large set of feature values that historical data cannot exhaust. Thus, the true underlying conditional distribution of the demand under a new feature value is likely unknown. In this case, a natural way is to replace the unknown underlying distribution of demands and features with its empirical counterpart. However, the resulting empirical risk minimization problem produces a pathological policy that can take arbitrary values on unseen feature values; see more detailed discussion in Section 1.4.2.

The pathological behavior of empirical feature-based newsvendor motivates the development of methods to generalize ordering decisions to unseen feature values. The most common approach is parameterization, namely, restricting the search to a parametric policy class. For example, Section 2.3.1 of Ban and Rudin (2019) studies affine policies, which can be efficiently solved using linear or convex optimization methods. The affine class can be restrictive and sub-optimal. Indeed, numerical experiments in Ban and Rudin (2019) show that affine policies are outperformed by their proposed kernel optimization method in the same paper. One possible remedy is to consider nonlinear transformations of features (basis functions). Thereby, one can enlarge the policy search space to an arbitrarily complex class with coefficients affinely dependent on the basis functions Bertsimas and Koduri (2022). However, specifying nonlinear transformations with good interpretability is a fundamentally challenging question. Similarly, neural-network polices Oroojlooyjadid et al. (2020); Meng et al. (2021) may have nice empirical performance but are often hard to interpret and data-demanding. In summary, in existing methods, there is a trade-off between the richness of the policy class and its interpretability/tractability. As such, the following question remains open: *Can we find an optimal policy without restricting onto a parametric family while still maintaining computational efficiency and interpretability*?

To answer this question, we consider a distributionally robust policy optimization framework. More specifically, it optimizes over all policies that are measurable functions of the features and does not parameterize the policy class. Moreover, it involves a minimax Wasserstein distributionally robust formulation Kuhn et al. (2019) that hedges against data uncertainty on the demand and features and helps to resolve the pathological issue of the empirical feature-based newsvendor. We remark that most literature on Wasserstein distributionally robust optimization has been focusing on deriving tractable reformulations when the decision variable is a finite-dimensional vector. Our main challenge here, however, is on the infinite-dimensional policy optimization. This distinguishes our model from most existing works. To ensure good generalization capability, many existing works exploit robust formulation for inventory models by considering various uncertainty sets based on moments Scarf (1958); Gallego and Moon (1993); Perakis and Roels (2008); Han et al. (2014); Xin and Goldberg (2021), precentiles Gallego et al. (2001), shape information Perakis and Roels (2008); Hanasusanto et al. (2015a); Natarajan et al. (2018), tail information Das et al. (2021), temporal dependence See and Sim (2010); Xin and Goldberg

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

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

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

(II) We further show that the optimal robust policy can be interpreted as a regularized critical fractile that regularizes the variation (measured by its Lipschitz norm with respect to the features) of the policy; see Figure 1.2 as an illustration. Based on this connection, the optimal robust policy optimization can be solved by linear programming. We compare the out-of-sample cost of the Shapley policy with various benchmarks using synthetic and real data, which demonstrate its superior empirical performance.

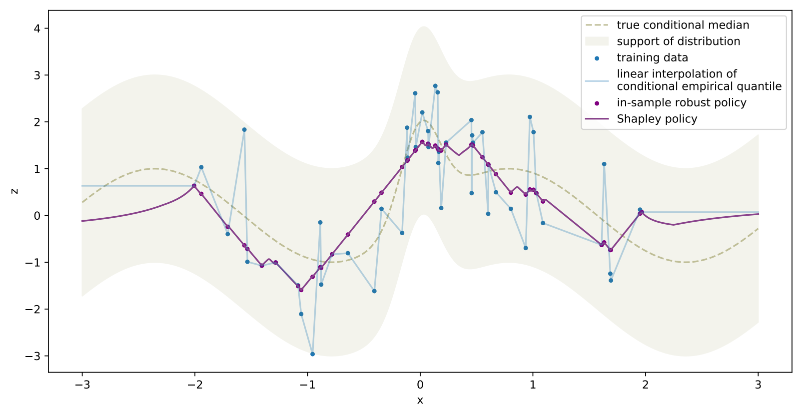


Figure 1.2: An illustration of policies where training data are generated from a two-dimensional continuous distribution of feature-demand pairs (x, z). Our proposed Shapley policy has a smaller variation (Lipschitz norm) compared to the linear interpolation of conditional quantiles of the empirical distribution.

1.4.1 Policy Optimization

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

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

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

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

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

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

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

$$\mathbb{E}\left[\inf_{y \in \mathcal{Z}} \mathbb{E}[\Psi(y, Z) \mid X]\right] = \inf_{f:\mathcal{X} \to \mathcal{Z}} \mathbb{E}[\Psi(f(X), Z)]. \tag{1.5}$$

Here, on the left-hand side of (1.5), 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 (1.5) is termed the *optimal true risk* in the literature Ban and Rudin (2019). Whereas on the right-hand side of (1.5), 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 (1.5) 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 (1.5) is the numerical performance measure considered by Ban and Rudin (2019); Kallus and Mao (2022) among other literature on decision-making with feature information.

1.4.2 Distributionally Robust Formulation

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

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

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

To solve the right-hand side of (1.5), a conventional wisdom is to consider the empirical risk minimization by replacing the true distribution with the empirical distribution $\widehat{\mathbb{P}}$

$$\inf_{f:\mathcal{X}\to\mathcal{Z}} \mathbb{E}_{(X,Z)\sim\widehat{\mathbb{P}}}[\Psi(f(X), Z)]. \tag{1.6}$$

Unfortunately, this would yield a degenerate solution that is only defined on the set of historical observations of features $\widehat{\mathcal{X}} := \{\widehat{x}_k : k \in [K]\}$, but can take arbitrary values elsewhere. In addition, suppose $\widehat{f}$ is the optimal policy of (1.6). Then for every $\widehat{x} \in \widehat{\mathcal{X}}$, $\widehat{f}(\widehat{x})$ is the critical fractile of the empirical conditional distribution. When the historical samples are generated from some continuous underlying distribution, then with probability one we have $n_k = 1$ for all k and $\widehat{f}(\widehat{x}_k) = \widehat{z}_{k1}$, which could be far away from the critical fractile of true conditional distribution.

Motivated by the above degeneracy of empirical risk minimization, we consider a minimax distributionally robust formulation which finds a decision hedging against a set $\mathfrak{M}$ of relevant probability distributions

$$\inf_{f \in \mathcal{F}} \sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X), Z)],$$

where $\mathcal{F}$ is the set of all measurable functions on $\mathcal{X}$. In our formulation, we choose $\mathfrak{M}$ to be a ball of distributions that are within ρ Wasserstein distance to a nominal distribution ($\rho \geq 0$). It is a natural choice since such distributional uncertainty set is data-driven and incorporates distributions on unseen feature values (see, e.g., Kuhn et al. (2019)). Let $\|\cdot\|_*$ denote the dual norm of the norm $\|\cdot\|$ on $\mathcal{X}$. Let $\mathcal{P}_1(\mathcal{X}\times\mathcal{Z})$ be the set of probability distributions on $\mathcal{X}\times\mathcal{Z}$ with finite first moment. The *Wasserstein distance* (of order 1) is defined as

$$\mathsf{W}(\mathbb{P}, \mathbb{Q}) := \inf_{\gamma \in \Gamma(\mathbb{P},\mathbb{Q})} \mathbb{E}_{(\tilde{X},\tilde{Z}),(X,Z)\sim\gamma}[\|\tilde{X} - X\| + |\tilde{Z} - Z|], \tag{1.7}$$

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

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

where ρ is the radius of the 1-Wasserstein ball where we use to construct the uncertainty set. Note that (P) can be viewed as a special case of policy optimization for two-stage Wasserstein distributionally robust optimization Bertsimas et al. (2022). Throughout the paper, we assume $b > 0$ and $0 \le h \le b$. We remark that when $b < h$, all results in the paper still hold for sufficiently small ρ (see Remark 4 in Appendix A.4).

1.4.3 Main Results

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

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

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

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

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

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

where

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

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

(i) solve the problem that is solely based on the in-sample data. Here in-sample means that instead of considering the full feature space, we restrict our attention to historical observations of features only, which is a finite subset. Observe that

the set of in-sample policies is $\widehat{\mathcal{F}} := \{\widehat{f} : \widehat{\mathcal{X}} \to \mathcal{Z}\} \subset \mathbb{R}_+^K$*, hence the in-sample problem, is a finite-dimensional optimization problem;*

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

In Section 1.5, we will prove Theorem 1 through the in-sample dual problem ($\widehat{\mathsf{D}}$)

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

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

In Section 1.6.1, we will show that the in-sample dual problem ($\widehat{\mathsf{D}}$) is also equivalent to the following in-sample Lipschitz regularization problem

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

which turns out to be equivalent to a finite-dimensional linear program that will be defined in Section 1.6.2.

In Section 1.5.2, we show that the matrix saddle point defined in Theorem 1(II) is the closed-form solution to the following Lipschitz constant minimization problem

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

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

1.5 Proof of Main Results

In this section, we prove Theorem 1. Problem (P) is an infinite-dimensional optimization whose main difficulty is that, we need not only to assign an ordering quantity for every historical observation of features in $\widehat{\mathcal{X}}$ but also to each unseen feature values in $\mathcal{X} \setminus \widehat{\mathcal{X}}$. In Section 1.5.1, we prove that the Shapely extension from the in-sample robust problem ($\widehat{\mathsf{P}}$) renders an optimal policy to the primal problem (P). In Section 1.5.2, we provide further intuition that drives behind the Shapely policy.

1.5.1 Shapley Policy and its Optimality

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

To begin with, applying strong duality for Wasserstein distributionally robust optimization (Lemma 6 in Appendix A.1) on the inner maximization of (P) and ($\widehat{\mathsf{P}}$) respectively yields their strong dual problems

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

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

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

$$\sup_{z\in\mathcal{Z}, x\in\mathcal{X}} \left\{\Psi(f^*(x), z) - \lambda\|x-\widehat{X}\| - \lambda|z-\widehat{Z}|\right\} \le \sup_{z\in\mathcal{Z}} \max_{x\in\widehat{\mathcal{X}}} \left\{\Psi(\widehat{f}^*(x), z) - \lambda\|x-\widehat{X}\| - \lambda|z-\widehat{Z}|\right\}. \tag{1.8}$$

To this end, we establish the following key lemma.

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

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

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

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

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

$$\sup_{x\in\mathcal{X}} \left\{\Phi(f(x)) - \|x-\widehat{x}_k\|\right\} \le \max_{j=1,\dots,K} \left\{\Phi(\widehat{f}(\widehat{x}_j)) - \|\widehat{x}_j - \widehat{x}_k\|\right\}. \tag{1.9}$$

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

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

For any x on the line segment connecting $\widehat{x}_j$ and $\widehat{x}_k$, $A_{jk}(x)$ is simply a linear interpolation, elsewhere $A_{jk}(x)$ is a distance-dependent weighted average, which is equivalent to the inverse distance weighting with two points Shepard (1968). The extension $f(x)$ is given by the saddle point (pure Nash equilibrium) of a matrix

$A_{jk}(x)$, whose existence is due to Shapley's theorem (Lemma 7 in Appendix A.1), thus we call it the *Shapley policy*. The first property states that f and $\widehat{f}$ coincide on in-sample data, thus f is indeed an extension. The second property implies (1.8) 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 1.3, we plot the Shapley extension when $K = 2, 3$ and when the feature space $\mathcal{X} = \mathbb{R}$. The horizontal axis represents the feature X and the vertical axis represents the policy value (ordering quantity). The points represent an in-sample robust optimal ordering policy. When $K = 2$, the extension $f^*(x) = A_{12}(x)$. On the line segment connecting two points $\widehat{x}_1$ and $\widehat{x}_2$, the interpolation is linear, and is curved elsewhere. As $|x| \to \infty$, the policy converges to a "non-informative" ordering quantity $(\widehat{f}^*(\widehat{x}_1) + \widehat{f}^*(\widehat{x}_2))/2$ which, intuitively, means that the historical observations of features provides little guidance on a faraway new feature value x and thus the policy simply takes the average of the two in-sample policy values. When $K = 3$, for each pair of three historical observations of features $\widehat{x}_1, \widehat{x}_2, \widehat{x}_3$, we plot three curves $A_{12}(x)$ (green), $A_{13}(x)$ (orange), and $A_{23}(x)$ (blue). By solving the minimax saddle point problem, the extended policy $f^*(x)$ would be the middle one among the three curves, as marked with the solid line. Thereby, the saddle point curve $f^*(x)$ is a balanced choice among all pairwise weighted averages. Generally, when x is close to $\widehat{x}_k$, $f^*(x)$ is close to $\widehat{f}^*(\widehat{x}_k)$. When x is away from all historical observations of features, $f^*(x)$ converges to $\frac{1}{2}\big(\min_{k\in[K]} \widehat{f}^*(\widehat{x}_k) + \max_{k\in[K]} \widehat{f}^*(\widehat{x}_k)\big)$. The case of two-dimensional feature space $\mathcal{X} = \mathbb{R}^2$ is similar, as shown in Figure A.1 in the Appendix A.1.

With Lemma 1, we can prove our main result easily by setting $\Phi(y) = \frac{1}{\lambda} \sup_{z\in\mathcal{Z}}\{\Psi(y, z)-$

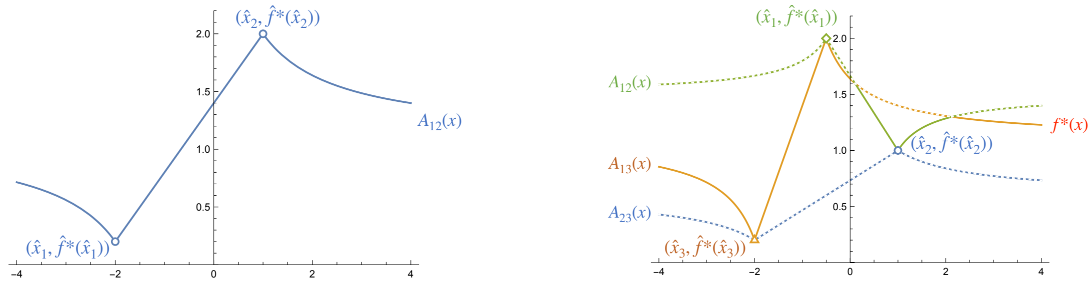


Figure 1.3: Graph of the Shapley policy $y = f(x)$ when $K = 2, 3$, $x \in \mathbb{R}$.

$\lambda|z - \widehat{Z}|\}$ (the degenerate case $\lambda = 0$ is trivial) and $\widehat{f} = \widehat{f}^*$, we can prove (1.8) and thereby, the dual problem (D) is equivalent to the in-sample dual problem $(\widehat{\mathsf{D}})$ in the following sense: $v_D = v_{\hat{D}}$, and every optimal policy $\widehat{f}^* \in \widehat{\mathfrak{F}}$ of $(\widehat{\mathsf{D}})$ can be extended to an optimal policy of (D) via the Shapley extension (S). Since the primal problems (P) and $(\widehat{\mathsf{P}})$ are equivalent to their dual problems (D) and $(\widehat{\mathsf{D}})$, respectively, the theorem is proved. We refer to Appendix A.1 for a complete proof.

1.5.2 Further Intuition Behind the Shapley Policy

Here we provide some intuition to explain the structural property of the Shapley policy that makes it robust optimal. We would like to emphasize that the discussion below argues from the perspective of the primal problem, whose main purpose is to offer some insights of the structural properties of the problem, possibly at the sacrifice of mathematical rigor. A complete proof will be established from the dual perspective in Appendix A.1.

1.5.2.1 Extension based on Slope Minimization

As discussed in Section 1.5.1, the solution to the problem (P) is related to the solution to the problem $(\widehat{\mathsf{P}})$. On the one hand, since we restrict the uncertainty set in $(\widehat{\mathsf{P}})$, the worst-case cost of a policy f in problem (P) is always greater or equal than the worst-case cost of its restriction policy $\widehat{f} = f|_{\widehat{X}}$ in problem ($\widehat{\mathsf{P}}$). On the other hand, if for any in-sample policy $\widehat{f} : \widehat{\mathcal{X}} \to \mathcal{Z}$ we can find an extended policy $f : \mathcal{X} \to \mathcal{Z}$ such that a worst case distribution is guaranteed to be supported on $\widehat{\mathcal{X}}$, then f has the same worst-case expected cost in (P) as $\widehat{f}$ in ($\widehat{\mathsf{P}}$), which makes f an optimal extension, thus the two problems become equivalent. Below we show that the extension defined by the slope minimization problem (L) indeed leads to a worst-case distribution supported on $\widehat{\mathcal{X}}$ and thus is optimal.

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

(I) The new transport plan incurs a higher cost due to the convexity of $\Psi(\cdot, z)$: $\delta \Psi(y_j, z) + (1 - \delta)\Psi(y_{k^+}, z) \ge \Psi(y, z)$.

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

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

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

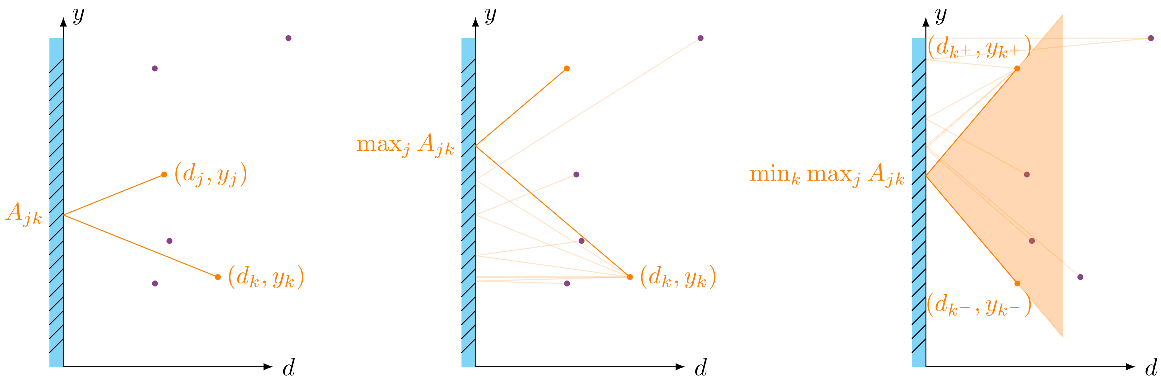
1.5.2.2 Shapley Policy as the Solution to Slope Minimization

Next, we show geometrically in Figure 1.4 that the Shapley extension (S) is the closed-form solution to (L). To visualize $A_{jk}(x)$, we plot the d–y plane on which a point (d_k, y_k) means $\widehat{x}_k$ is of distance d_k away from x and is assigned a policy value y_k. Imagine a mirror at $d = 0$ facing right. It is not hard to see that $A_{jk}(x)$ is the reflection point of the point j in the mirror from the point k's viewpoint (Figure 1.4(a)). Thereby $\max_j A_{jk}(x)$ corresponds to the highest reflection points among all points in the mirror of point k'th horizon, as shown in Figure 1.4(b). Minimizing over k gives the Shapley saddle point $f(x) = \min_k \max_j A_{jk}(x)$ (Figure 1.4(c)). Geometrically, since $L_{k_+} = -L_{k_-}$, the shadow region in 1.4(c) is a symmetric cone with the smallest opening that covers all the points $\{(d_k, y_k)\}_{k\in[K]}$. Thus, it is apparent that the minimax theorem holds by symmetry, and the vertex of such smallest opening is precisely determined by the Lipschitz-minimization problem (L).

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

$$\begin{aligned} \min_{L\geq 0, y\in\mathbb{R}} \quad & L \\ \text{subject to} \quad & y - Ld_k \leq y_k \leq y + Ld_k, \qquad \text{for all } k \in [K] \end{aligned} \tag{1.10}$$

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



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

Figure 1.4: Visualization of the Shapley saddle point

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

1.6 Discussions

In this section, we provide additional properties of our distributionally robust formulation (P) or its dual (D). Unlike Section 1.5, results in this section relies on the specific form of the newsvendor cost beyond convexity. In Section 1.6.1, we show that the problems (D) ($\widehat{\mathsf{D}}$) 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 1.6.2, and derive the generalization bound of the Shapley policy in Section 1.6.3. Finally, in Section 1.6.4, we discuss the structure of the optimal robust policy.

1.6.1 Interpretation as Lipschitz Regularization

In this subsection, we establish an equivalence between our robust formulation and Lipschitz regularization, as already hinted in Section 1.5.2.

Consider the following Lipschitz regularization problem defined as

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

where we denote by $a_1 \vee a_2$ the maximum between a_1 and a_2, and by $\|\cdot\|_{\text{Lip}}$ 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) would degenerate to the non-robust empirical risk minimization. In this case, the optimal policy f^* is defined only on the in-sample data, which equals to the critical fractile of the empirical conditional distribution; and can take any value outside $\widehat{\mathcal{X}}$. If we prohibit perturbing any data in z-direction, then the lower cut off 1 of the Lipschitz norm $\|f\|_{\text{Lip}}$ 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 classical newsvendor problem without feature information

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

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

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

The following result enables us to further reformulate the problem as Lipschitz regularization. The proof is provided in Appendix A.2.

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

In the literature, it is known that 1-Wasserstein distributionally robust optimization is upper bounded by Lipschitz regularization Esfahani and Kuhn (2018); Shafieezadeh-Abadeh et al. (2019); Gao et al. (2017), and the two problems are equivalent under certain assumptions. That said, Proposition 1 is arguably surprising because our considered problem does not satisfy the assumptions imposed in the references above that ensure the equivalence. The key observation in the proof (Lemma 8 in Appendix A.2) is that there exists a robust optimal in-sample policy $\widehat{f}^*$ with sufficiently small Lipschitz norm, which gives a direct restriction on the range of the dual multiplier λ and transforms ($\widehat{\mathsf{D}}$) to ($\widehat{\mathsf{R}}$).

Thus far, combining all results in Sections 1.5.1 and 1.6.1, we have shown that problems ($\widehat{\mathsf{P}}$), ($\widehat{\mathsf{D}}$) and ($\widehat{\mathsf{R}}$) share an in-sample optimal robust policy $\widehat{f}^* \in \widehat{\mathcal{F}}$, which can be extended to an optimal robust policy $f^* \in \mathcal{F}$ for problems (P), (D) and (R).

1.6.2 Linear Programming Reformulation for In-sample Robust Problem

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

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

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

This linear program has $N+K+1$ variables with K^2+2N+1 constraints. The variable L stands for $1\vee\|\widehat{f}\|_{\mathrm{Lip}}$; and ψ_{ki} represents the cost of ordering quantity y_k when the demand is $\widehat{z}_{ki}$ and the feature value is $\widehat{x}_k$. In practice, we may impose different norm scaling parameter $\beta>0$ in $\|(x,z)\| = \|x\| + \frac{1}{\beta}|z|$ that balances between the uncertainty in the feature information and in the demand. In this new setting, $v_{\hat{R}}$ becomes $v_{\hat{R}} = (b\vee h)(\beta\vee\|\widehat{f}\|_{\mathrm{Lip}})\rho + \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi_{\widehat{f}}]$, and in the linear programming $L\geq 1$ becomes $L\geq\beta$. In Appendix A.3, we show in Figure A.2 how the performance of the algorithm is affected by the choice of parameters ρ and β.

1.6.3 Generalization Error Bound

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

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

Proposition 2, we can see that

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

where λ^* is the optimal dual variable in $(\widehat{\mathsf{D}})$. Hence, using the Lipschitz composition property of the Radamacher complexity, the expected generalization gap of Ψ_{f^*} is dominated by λ^* times the expected Radamacher complexity of a 1-Lipschitz ball in $\mathcal{X}$, $\mathfrak{R}_n(\text{Lip}_1(\mathcal{X})) = O(n^{-\frac{1}{d}})$, 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 $\bar{D} > 0$. Then the expected generalization gap of the optimal robust policy f^* is upper bounded by*

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

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

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

1.6.4 Structure of the Optimal Robust Policy

In this subsection, we discuss the structure of the optimal robust policy, which provides further interpretation of the in-sample optimal robust policy. The proofs are provided in Appendix A.4.

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

$$\overline{q}(\widehat{x}) := \max\left\{z \in \mathcal{Z} : \widehat{\mathbb{P}}\{Z < z | X = \widehat{x}\} \leq \frac{b}{b+h}\right\}, \tag{1.11}$$

$$\underline{q}(\widehat{x}) := \min\left\{z \in \mathcal{Z} : \widehat{\mathbb{P}}\{Z \leq z | X = \widehat{x}\} \geq \frac{b}{b+h}\right\}. \tag{1.12}$$

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

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

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

Proposition 4.

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

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

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

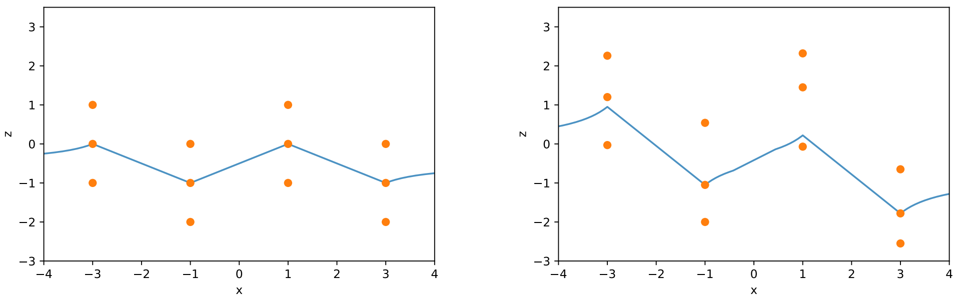


Figure 1.5: Optimal policy $\widehat{f}^*$ (blue curves) given the empirical distribution of feature and demand (orange dots). If the empirical conditional critical fractile q is 1-Lipschitz, then $\widehat{f}^* = q$ (left). Otherwise, $\widehat{f}^*$ regularizes q by reducing its Lipschitz norm (right).

than the empirical critical fractile $\underline{q}(\widehat{x}_k)$ at the cost of back ordering more, resulting in a set $\widehat{\mathcal{X}}_>$.

The discussion above is illustrated in Figure 1.5. We have $b = 3$, $h = 2$, $\rho = 10$, $\mathcal{X} = \mathbb{R}$, $\widehat{\mathcal{X}} = \{-3, -1, 1, 3\}$, and $\widehat{\mathbb{P}}_n$ is supported on $n = 12$ points (as indicated by the orange dots in x–z plot), where each $\widehat{x} \in \widehat{\mathcal{X}}$ are associated to three demand realizations with the middle level corresponding to the empirical conditional critical fractile $q = \overline{q} = \underline{q}$. In the left example, q is 1-Lipschitz, hence the optimal policy $\widehat{f}^*$ (represented by the blue curve) passes through all empirical conditional critical fractiles; on the right, $\widehat{f}^*$ regularizes q by ordering less than q on $\widehat{x} = -3, 1$ so as to reduce the variation of the policy.

Given the structure of the optimal policy, in Proposition 13 and Figure A.3 in Appendix A.4, we investigate the worst-case distribution $\mathbb{P}^*$, which sheds light on the (non-)conservativeness of our formulation.

Chapter 2: Distributionally Robust Stochastic Optimization with Causal Transport Distance

2.1 Introduction

This chapter is based on Yang et al. (2022), where we investigate the distributionally robust stochastic optimization (DRSO) with the causal transport distance. The rest of the chapter proceeds as follows. We first introduce the causal transport distance and corresponding robust model. In Section 2.2, we develop a duality result for evaluating the worst-case expected cost by exploiting the structure of the worst-case distribution. In Section 2.3, we consider the outer optimization over affine decision rules and over all decision rules. Proofs and additional results are deferred to Appendices.

We first briefly provide some background on distributionally robust optimization with causal transport distance.

We revisit the definition of causal transport distance introduced in Definition 1, Section 1.2.2. A joint distribution $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$ is called a *causal transport plan* if for $((\widehat{X}, \widehat{Z}), (X, Z)) \sim \gamma$, X and $\widehat{Z}$ are conditionally independent given $\widehat{X}$:

$$X \perp\!\!\!\perp \widehat{Z} \mid \widehat{X}.$$

We denote by $\Gamma_c(\widehat{\mathbb{P}}, \mathbb{P})$ the set of all transport plans $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$ that are causal. Let $p \in [1, \infty)$. The *p-causal transport distance* between $\widehat{\mathbb{P}}$ and $\mathbb{P}$ is defined as

$$\mathcal{C}_p(\widehat{\mathbb{P}}, \mathbb{P}) := \left(\inf_{\gamma \in \Gamma_c(\widehat{\mathbb{P}}, \mathbb{P})} \mathbb{E}_{((X,Z),(\widehat{X},\widehat{Z})) \sim \gamma} \left[\|X - \widehat{X}\|_{\mathcal{X}}^p + \|Z - \widehat{Z}\|_{\mathcal{Z}}^p \right] \right)^{1/p}.$$

Like Wasserstein distance, causal transport distance finds the minimal transport cost between two distributions, where norms capture the geometry of the data space and similarity between samples. The conditional independence condition in Definition 1 basically means that the destination X of a sample in a causal transport plan should depend only on the origin $\widehat{X}$ but not on the associated information of $\widehat{Z}$. There are other equivalent definitions of a causal transport plan, which are provided in Appendix B.1. Let us use the following example to explain a causal transport plan.

Example 3 (Causal Transport between Color Images). *Let $\mathcal{X} = \{1, 2, \ldots, H\}^2$, where H represents the width of a squared image, and let $\mathcal{Z} = \{R, G, B\}$, representing the three color channels, red (R), green (G), and blue(B). A bitmap image stores the position-color information of an image via a $H \times H \times 3$ tensor $A = (A_{ijk})_{i,j \in \{1,2,\ldots,H\}, k \in \{1,2,3\}}$. Its (i, j, k)-th entry $A_{ijk} \in \{0, 1, \ldots, 255\}$ represents the 8-bit indexed color at pixel position (i, j) in the k-th channel. With a normalizing constant $M = \sum_{i,j,k} A_{ijk}$, the tensor A/M represents a probability mass function on $\mathcal{X} \times \mathcal{Z}$. Let us equip norms $\|\cdot\|_{\mathcal{X}} = \|\cdot\|_1$ and $\|\cdot\|_{\mathcal{Z}} = c\mathbf{1}\{\cdot = 0\}$, where c is a scaling parameter.*

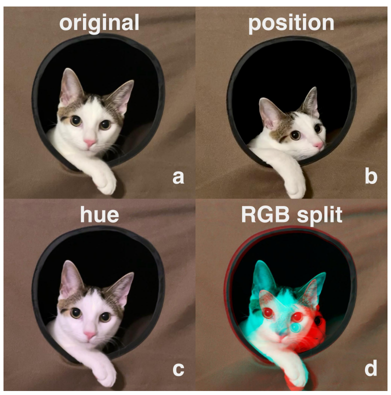


Figure 2.1: A color image (a) and its variations by shifting the position (b), adjusting the hue (c), or splitting the RGB channels (d)

Figure 2.1 contains four images of a cat: (a)(b)(c) can be viewed as real natural images with different poses or color portions, whereas (d) can be viewed as a fake *image in which the pose exhibited via the red channel is different from that via the green/blue channel.*

(I) *The movement of the cat yields a causal transport plan from (a) to (b), as under such movement, the destination* (X, Z) *in (b) of a position-channel pair* $(\widehat{X}, \widehat{Z})$ *in (a) depends only on its original position* $\widehat{X}$ *but not on the channel information* $\widehat{Z}$*, or put it differently, all channels are moved in the same way from* $\widehat{X}$ *to* X *without changing the channel value* $\widehat{Z}$*. This matches precisely the definition of a causal transport.*

(II) *The cats in (a) and (c) have identical poses but different hue values. Changing hue values of an image would affect its RGB values and thus the distribution on* $\mathcal{Z}$*. Such color adjustment (changing RGB values while fixing the position) defines a causal transport plan from (a) to (c). Indeed, under such movement, a position-channel pair* $(\widehat{X}, \widehat{Z})$ *in (a) keeps its position in* $\boldsymbol{c}$*, namely,* $X = \widehat{X}$*, regardless of the value of* $\widehat{Z}$*. Note that in a causal transport plan, we allow the destination* Z *of* $\widehat{Z}$ *to be dependent on both* $\widehat{X}$ *and* $\widehat{Z}$*, that is, at each position of the image, changes in the color portions are permitted.*

(III) *The green and blue channels of (d) has the same pose as (a), whereas the red channel of (d) has the same pose as (b). If we consider a transport plan that keeps a position-channel pair* $(\widehat{X}, \widehat{Z})$ *if* $\widehat{Z} \in \{\boldsymbol{G}, \boldsymbol{B}\}$*, and transport it according to the cat's movement if* $\widehat{Z} = \boldsymbol{R}$*, then such a transport plan is not causal, because given* $\widehat{X}$*, where this position-channel pair is transported depends on the channel information* $\widehat{Z}$*.*

In Table 2.1, we compute the Wasserstein distance and causal transport distance between Fig. 2.1(a) and the other three variations, with $\boldsymbol{H} = 32$ *and* $\boldsymbol{c} =$

Table 2.1: Distance between Figure 2.1(a) and the other three variations

Variations	(b)	(c)	(d)
Wasserstein distance	2.303	2.044	0.495
Causal transport distance	2.767	2.535	**6.388**

4. *We find that, under causal transport distance between Fig. 2.1(a) and the fake image Fig. 2.1(d) is much larger than that between Fig. (a) and the real images Fig. 2.1(b)(c). In contrast, Wasserstein distance fails to capture such an intuition.* ♣

As hinted in Example 3, one of the main advantage of causal transport distance over Wasserstein distance is that it captures the structure of the conditional distribution. To further illustrate this, let us revisit the toy Example 2.

Example 4 (Revisit of Example 2)**.** *We compute the causal transport distance and the Wasserstein distance between* $\widehat{\mathbb{P}}$ *and* $\mathbb{P}^{\varepsilon}$ *shown in Example 2. Since the conditional distribution of* $\mathbb{P}^{\varepsilon}$ *is a Dirac measure for every* x*, the causal transport distance between* $\widehat{\mathbb{P}}$ *and* $\mathbb{P}^{\varepsilon}$ *is uniformly bounded from below by a positive constant for all* $\varepsilon > 0$*. In fact, it is not hard to see that the only causal transport plan is the independent joint distribution* $\widehat{\mathbb{P}} \otimes \mathbb{P}^{\varepsilon}$*, so*

$$\begin{aligned}
\mathcal{C}_p(\widehat{\mathbb{P}}, \mathbb{P}^{\varepsilon})^p &= \frac{1}{\sin\varepsilon}\int_0^{\sin\varepsilon} |x-0|^p \,\mathrm{d}x + \frac{1}{\cos\varepsilon}\int_0^1\int_0^{\cos\varepsilon} |\hat{z}-z|^p \,\mathrm{d}z\,\mathrm{d}\hat{z} \\
&= \frac{\sin^p\varepsilon}{p+1} + \frac{1+\cos^{p+2}\varepsilon - (1-\cos\varepsilon)^{p+2}}{(p+1)(p+2)\cos\varepsilon} \\
&= \Big((1+p)(1+\frac{p}{2})\Big)^{-\frac{1}{p}} + O(\varepsilon).
\end{aligned}$$

As a result, $\mathbb{P}$ *would not belong to the uncertainty set induced from the causal transport distance with a small radius. This is consistent to our intuition. In contrast, for the Wasserstein distance, observe that the optimal transport plan is simply the rotation* *transform, thereby the Wasserstein distance is* $(p+1)^{-\frac{1}{p}}(\sin^p \varepsilon+(1-\cos\varepsilon)^p)^{\frac{1}{p}} = O(\varepsilon)$, *which is small whenever the angle between the two line segments is small. Consequently, any Wasserstein uncertainty set with a positive radius contains infinitely many distributions with dramatically different conditional information structure from the nominal one, and therefore may lead to an overly conservative solution.* ♣

Next we point out an important property of the uncertainty set constructed using the causal transport distance: for any $\widehat{\mathbb{P}} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z})$ and $\rho > 0$, the set $\mathfrak{M} = \{\mathbb{P} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z}) : \mathcal{C}_p(\widehat{\mathbb{P}}, \mathbb{P}) \leq \rho\}$ is convex, as indicated in the following lemma.

Lemma 2 (Convexity)**.** *If* $\gamma^{(0)}$ *and* $\gamma^{(1)}$ *are two causal transport plans from* $\widehat{\mathbb{P}}$ *to* $\mathbb{P}^{(0)}$ *and* $\mathbb{P}^{(1)}$ *respectively, then for any* $q \in [0,1]$, $\gamma^q := (1-q)\gamma^{(0)} + q\gamma^{(1)}$ *is also a causal transport plan from* $\widehat{\mathbb{P}}$ *to* $\mathbb{P}^{(q)} = (1-q)\mathbb{P}^{(0)} + q\mathbb{P}^{(1)}$. *Moreover, everything follows even if we replace* q *by any measurable function* $q : \mathcal{X} \to [0,1]$.

We remark that the direction of the transport plan matters: if $\gamma^{(0)}$ and $\gamma^{(1)}$ are two causal transport plans from $\widehat{\mathbb{P}}^{(0)}$ and $\widehat{\mathbb{P}}^{(1)}$ to $\mathbb{P}$ respectively, we cannot assert that their convex combination $\gamma^{(q)}$ is also a causal transport plan. For a counterexample, please refer to the Fig. 1.17 in Pflug and Pichler (2014).

We close this subsection by noting that although the notion of causal transport per se does not imply any causal relationship, the causal transport distance does indicate .

Example 5. *One interesting observation is that unlike Wasserstein distance, causal transport distance is asymmetric. For instance, the animation of mixing coffee and milk in Figure 2.2 is causal from left to right, but not from right to left if we reverse the time.*

Figure 2.2: Five frames of a video of coffee adding milk Mashed (2022)

This is because causal transport plan forbids color to split but allows colors to blend: two objects at different position $(\widehat{X}_1, \widehat{X}_2)$ *can converge to the same new location* X*, albeit they will have different colors* (Z_1, Z_2) *which are not independent from their old positions* $(\widehat{X}_1, \widehat{X}_2)$*.* ♣

2.1.1 Distributionally Robust Formulation

Based on the definition in the previous subsection, we study the following distributionally robust optimization problem with causal transport distance

$$v_{\mathsf{P}} := \inf_{f \in \mathcal{F}} \max_{\mathbb{P} \in \mathfrak{M}} \ \mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X), Z)], \text{ where } \mathfrak{M} = \left\{\mathbb{P} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z}) : \ \mathcal{C}_p(\widehat{\mathbb{P}}, \mathbb{P}) \leq \rho\right\}. \tag{P}$$

Below, we list a few examples.

Example 6 (Conditional mean estimation)**.** *The conditional mean of* Z *given* X *can be estimated by minimizing the square loss* $(f(X) - Z)^2$*. Thus we consider the following robust conditional mean estimation problem*

$$\inf_{f \in \mathcal{F}} \sup_{\mathbb{P} \in \mathfrak{M}} \ \mathbb{E}_{(X,Z)\sim\mathbb{P}}\left[(f(X) - Z)^2\right]. \qquad ♣$$

Example 7 (Feature-based Newsvendor)**.** *Let* h *and* b *represent the unit holding cost and the unit backordering cost, respectively, and let* Z *be the random demand and* X *be the covariate features. The goal is to minimize the newsvendor cost function*

$\Psi(w, z) = h(w - z)_+ + b(z - w)_+$. *Consider*

$$\inf_{f \in \mathcal{F}} \sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{(X,Z)\sim\mathbb{P}}\left[h(f(X) - Z)_+ + b(Z - f(X))_+\right].$$

Note that this model also serves as the conditional $\frac{b}{b+h}$*-quantile estimation. In particular, when* $h = b = 1$*, this is the conditional median estimation.* ♣

Example 8 (Personalized Pricing)**.** *Consider an affine demand model* $D(w) = z_1 w + z_2 = Z^\top \begin{pmatrix} w \\ 1 \end{pmatrix}$*, where* w *is the price and* z *are unknown coefficients, with* $z_2 > 0$ *representing the demand at zero price and* $z_1 < 0$ *representing the price sensitivity coefficient, which is the rate at which the price affects the demand. In practice, both coefficients* z_1 *and* z_2 *may exhibit heterogeneity among populations. As such, we model it as a two-dimensional random variable* Z*, which is affected by the contextual information* X*, based on which the decision maker can adjust the price directly or indirectly through personalized promotion. The revenue is calculated as* $w(Z_1 w + Z_2)$*. Consider a revenue maximization with personalized pricing*

$$\inf_{f \in \mathcal{F}} \sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{(X,Z)\sim\mathbb{P}}\left[-f(X) Z^\top \begin{pmatrix} f(X) \\ 1 \end{pmatrix}\right].$$ ♣

In the last example, we consider a portfolio optimization problem where the decision rule is restricted to be affine.

Example 9 (Portfolio Optimization with Affine Decision Rule)**.** *Consider a portfolio optimization involving* m *assets. The return rate of the* i*th asset is modeled as a random variable* Z_i*. Suppose a weight* $w \in \mathbb{R}^m$ *is allocated on the assets with the restriction* $\sum_{i=1}^m w_i = 1$*, thereby the random loss of a portfolio is given by* $w^\top Z$*. Again, the weight* w *can be chosen based on the contextual information* X*. Consider the portfolio optimization problem*

$$\inf_{f \in \mathcal{F}} \sup_{\mathbb{P} \in \mathfrak{M}} \mathbb{E}_{(X,Z)\sim\mathbb{P}}\left[f(X)^\top Z\right].$$

Here $\mathcal{F}$ is a class of functions $f : \mathcal{X} \to \mathcal{D}$, where

$$\mathcal{X} = \mathbb{R}^d, \qquad \mathcal{D} = \left\{ w \in \mathbb{R}^m : \mathbf{1}^\top w = 1 \right\}.$$

Here $\mathbf{1}$ is the m-dimensional all-one vector. In case of affine policies, we require $f \in \mathcal{F}$ to be affine. We can write

$$\mathcal{F} = \left\{ x \mapsto \left(\mathrm{Id} - \frac{1}{m} \mathbf{1}\mathbf{1}^\top \right) Ax + \frac{1}{m}\mathbf{1} : A \in \mathbb{R}^{m \times d} \right\}.$$

Here Id *is the m-dimensional identity matrix.* ♣

2.2 Evaluating the Worst-case Expectation

In this section, we develop a tractable reformulation for the inner maximization of (P) based on strong duality. As a byproduct of our proof, we also derive the structure of the worst-case distribution, which demonstrates how our choice of causal transport distance-based distributional uncertainty set helps to preserves the conditional information structure of the nominal distribution in the worst case.

Throughout this paper, we make the following assumption, which focuses on the data-driven setting where the nominal distribution is discrete. Our proof technique can be extended to a general metric space with additional technical treatment.

Assumption 1. *$\mathcal{X}$, $\mathcal{Z}$, $\mathcal{D}$ are normed vector spaces. The cost function $\mathbf{\Psi} : \mathcal{D} \times \mathcal{Z} \to \mathbb{R}$ is measurable. The nominal distribution $\widehat{\mathbb{P}} \in \mathcal{P}(\mathcal{X} \times \mathcal{Z})$ is a discrete probability measure*

$$\widehat{\mathbb{P}} = \sum_{k=1}^{K} \sum_{i=1}^{n_k} \hat{p}_{ki} \delta_{(\widehat{x}_k, \widehat{z}_{ki})}, \qquad \textit{with} \ \sum_{k=1}^{K} \sum_{i=1}^{n_k} \hat{p}_{ki} = 1.$$

2.2.1 Strong Duality Reformulation

We begin by developing a tractable reformulation through deriving its strong dual. For a fixed decision rule f, we define the primal problem as

$$v_{\mathsf{P}}^{f} := \max_{\mathbb{P}\in\mathfrak{M}} \mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)], \tag{P^f}$$

and the dual problem as

$$v_{\mathsf{D}}^{f} := \inf_{\lambda\geq 0}\left\{\lambda\rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\sup_{x\in\mathcal{X}}\left\{\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}}\left\{\Psi(f(x),z)-\lambda\|z-\widehat{Z}\|^p\right\}\mid\widehat{X}\right]-\lambda\|x-\widehat{X}\|^p\right\}\right]\right\}. \tag{D^f}$$

The dual variable λ corresponds to the Lagrangian multiplier of the causal constraint in the primal problem. We will show that (P^f) and (D^f) are equal, leading to the main result of Theorem 2 by taking the infimum over f.

To prove the strong duality, we first develop a relatively straightforward weak duality result.

Proposition 5 (Weak duality)**.** *Let $f:\mathcal{X}\to\mathcal{D}$ be a measurable function. Then $v_{\mathsf{P}}^{f}\leq v_{\mathsf{D}}^{f}$.*

Proof. Proof. The proof is based on an application of Lagrangian weak duality. First, we derive from the Lagrangian weak duality the following

$$\begin{aligned}
v_{\mathsf{P}}^{f} &= \sup_{\mathbb{P}}\left\{\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)]:\ \mathcal{C}_p(\widehat{\mathbb{P}},\mathbb{P})^p\leq\rho^p\right\}\\
&= \sup_{\mathbb{P}}\inf_{\lambda\geq 0}\left\{\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)]-\lambda\left(\mathcal{C}_p(\widehat{\mathbb{P}},\mathbb{P})^p-\rho^p\right)\right\}\\
&\leq \inf_{\lambda\geq 0}\sup_{\mathbb{P}}\left\{\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)]-\lambda\left(\mathcal{C}_p(\widehat{\mathbb{P}},\mathbb{P})^p-\rho^p\right)\right\}.
\end{aligned}$$

Since for any $\gamma\in\Gamma_c(\widehat{\mathbb{P}},\mathbb{P})$,

$$\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)] = \mathbb{E}_{((X,Z),(\widehat{X},\widehat{Z}))\sim\gamma}[\Psi(f(X),Z)],$$

so we can write

$$\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)] - \lambda\left(\mathcal{C}_p(\widehat{\mathbb{P}},\mathbb{P})^p - \rho^p\right) = \lambda\rho^p + \sup_{\gamma\in\Gamma_c(\widehat{\mathbb{P}},\mathbb{P})} \mathbb{E}_\gamma\left[\Psi(f(X),Z) - \lambda\|X-\widehat{X}\|^p - \lambda\|Z-\widehat{Z}\|^p\right]$$

By the tower property,

$$\begin{aligned}\mathbb{E}_\gamma[\cdot] &= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\gamma_{X|\widehat{X}}}\left[\mathbb{E}_{\gamma_{\widehat{Z}|(\widehat{X},X)}}\left[\mathbb{E}_{\gamma_{Z|(\widehat{X},\widehat{Z},X)}}\left[\cdot|\widehat{X},\widehat{Z},X\right]|\widehat{X},X\right]|\widehat{X}\right]\right]\\ &= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\gamma_{X|\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\mathbb{E}_{\gamma_{Z|(\widehat{X},\widehat{Z},X)}}\left[\cdot|\widehat{X},\widehat{Z},X\right]|\widehat{X},X\right]|\widehat{X}\right]\right]\end{aligned}$$

where we use $\gamma_{\widehat{Z}|(\widehat{X},X)} = \widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}$ for a.e.-$(\widehat{X},X)$ because γ is causal. Therefore we have

$$\begin{aligned}&\mathbb{E}_\gamma\left[\Psi(f(X),Z) - \lambda\|X-\widehat{X}\|^p - \lambda\|Z-\widehat{Z}\|^p\right]\\ &= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\gamma_{X|\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\mathbb{E}_{\gamma_{Z|(\widehat{X},\widehat{Z},X)}}\left[\Psi(f(X),Z) - \lambda\|X-\widehat{X}\|^p - \lambda\|Z-\widehat{Z}\|^p|\widehat{X},\widehat{Z},X\right]|\widehat{X},X\right]|\widehat{X}\right]\right]\\ &\le \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\sup_{x\in\mathcal{X}}\left\{\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}}\left\{\Psi(f(x),z) - \lambda\|z-\widehat{Z}\|^p\right\}|\widehat{X}\right] - \lambda\|x-\widehat{X}\|^p\right\}\right].\end{aligned}$$

This completes the proof for the weak duality. □ □

The main result for this section states as follows.

Theorem 2 (Strong Duality)**.** *Let $f:\mathcal{X}\to\mathcal{D}$ be a measurable function. Then* $v_{\mathsf{P}}^f = v_{\mathsf{D}}^f$.

The proof idea of Theorem 2 is to construct a nearly worst-case distribution of the primal problem based on the first-order optimality condition of the weak dual problem. Conceptually it shares some similar aspects to the duality proof for Wasserstein DRO Gao and Kleywegt (2016), but differs from it in terms of the construction of a nearly worst-case distribution. Specifically, the nearly worst-case distribution is obtained by moving $\widehat{z}_{ki}$ toward the maximizer of the innermost maximization problem $\Upsilon(\lambda;x,\widehat{z}_{ki}) := \sup_{z\in\mathcal{Z}}\{\Psi(f(x),z)-\lambda\|z-\widehat{z}_{ki}\|\}$, and moving $\widehat{x}_k$ toward the maximizer of the maximization problem $\sup_{x\in\mathcal{X}}\{\Upsilon(\lambda;x,\widehat{z}_{ki})-\lambda\|x-\widehat{x}_k\|\}$. One can see that such a transport plan is causal: where $\widehat{x}_k$ is transported depends only on $\widehat{x}_k$ but not on $\widehat{z}_{ki}$. If both maximizers over x and over z exist and are unique, then the transport plan would induce a worst-case distribution. If the maximizers do not exist or are not unique, we can still find two transport plans such that one induces a feasible yet suboptimal distribution, while the other induces an infeasible yet superoptimal distribution. By interpolating these two distributions, we can obtain a near-optimal feasible solution to the primal problem. We refer to the next subsection for a more detailed construction of a worst-case distribution and Appendix B.3 for a complete proof.

Remark 2 (Comparison with Wasserstein DRO)**.** *Recall the Wasserstein DRO problem*

$$\sup_{\mathbb{P}} \left\{ \mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)] : \ \mathsf{W}_p(\widehat{\mathbb{P}},\mathbb{P}) \le \rho \right\},$$

which has the following equivalent dual form Gao and Kleywegt (2016); Zhang et al. (2022)

$$\begin{aligned} &\inf_{\lambda\ge 0} \left\{ \lambda\rho^p + \mathbb{E}_{\widehat{\mathbb{P}}}\left[\sup_{\substack{x\in\mathcal{X} \\ z\in\mathcal{Z}}} \left\{ \Psi(f(x),z) - \lambda\|z-\widehat{Z}\|^p - \lambda\|x-\widehat{X}\|^p \right\} \right] \right\} \\ =&\inf_{\lambda\ge 0} \left\{ \lambda\rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{x\in\mathcal{X}} \left\{ \sup_{z\in\mathcal{Z}} \left\{ \Psi(f(x),z) - \lambda\|z-\widehat{Z}\|^p \right\} - \lambda\|x-\widehat{X}\|^p \right\} \mid \widehat{X} \right] \right] \right\}. \end{aligned}$$

Comparing it with the dual problem (D^f) *of causal transport distance DRO, the difference is the switching of supremum over* x *and the conditional expectation of* $\widehat{Z}$ *given* $\widehat{X}$*. Hence, if the switching does not change the objective value, which holds, for instance, when the conditional distribution* $\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}$ *is a Dirac measure for every* $\widehat{X}$*, then the Wasserstein DRO dual problem and causal transport distance DRO dual problems are equal. From a primal point of view, if* $\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}$ *is Dirac for every* $\widehat{X}$*, then every transport plan from* $\widehat{\mathbb{P}}$ *to* $\mathbb{P}$ *is causal. In this case, thus causal transport distance DRO* *and Wasserstein DRO coincide. Intuitively, if every conditional distribution* $\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}$ *is Dirac, then the nominal distribution does not have any meaningful conditional information structure to exploit, and thus the causal transport distance DRO reduces to Wasserstein DRO.* ◊

2.2.2 Worst-case Distribution

In this subsection, we investigate the structure of the worst-case distribution and its existence conditions. Compared with the results in Section 2.2.1, in the folowing result we require $\mathcal{X}$ and $\mathcal{Z}$ to be finite dimensional and thus locally compact, and require some continuity assumptions on $\mathbf{\Psi}$, so that the maximizers are attainable.

Theorem 3 (Worst-case distribution)**.** *Suppose* $\mathcal{X}, \mathcal{Z}$ *are finite dimensional, and* $\mathbf{\Psi}(f(\cdot), \cdot)$ *is upper semi-continuous. If the reformulation* (D^f) *is achieved at some* $\lambda^* > \kappa$ *for* κ *specified in Lemma 10, then a worst case distribution exists and has the following form*

$$\mathbb{P}^* = \sum_{k \neq k_0} \sum_{i=1}^{n_k} \hat{p}_{ki} \delta_{(x_k^*, z_{ki}^*)} + \sum_{i=1}^{n_{k_0}} \hat{p}_{k_0 i} \left(q \delta_{(\overline{x}_{k_0}, \overline{z}_{k_0 i})} + (1-q) \delta_{(\underline{x}_{k_0}, \underline{z}_{k_0 i})} \right),$$

where $1 \leq k_0 \leq K$*,* $0 \leq q \leq 1$*,* $(x_k^*, z_{ki}^*) = (\overline{x}_k, \overline{z}_{ki})$*, and for every* k *and* i*,*

$$\overline{x}_k, \underline{x}_k \in \arg\max_{x \in \mathcal{X}} \left\{ \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}} \left[\sup_{z \in \mathcal{Z}} \left\{ \mathbf{\Psi}(f(x), z) - \lambda^* \|z - \widehat{Z}\|^p \right\} | \widehat{X} = \widehat{x}_k \right] - \lambda^* \|x - \widehat{x}_k\|^p \right\},$$

$$\overline{z}_{ki} \in \arg\max_{z \in \mathcal{Z}} \left\{ \mathbf{\Psi}(f(\overline{x}_k), z) - \lambda^* \|z - \widehat{z}_{ki}\|^p \right\}, \qquad \underline{z}_{ki} \in \arg\max_{z \in \mathcal{Z}} \left\{ \mathbf{\Psi}(f(\underline{x}_k), z) - \lambda^* \|z - \widehat{z}_{ki}\|^p \right\}.$$

From Theorem 3 we see that there exists a worst-case distribution $\mathbb{P}^*$ supported on at most $N + n_{k_0}$ points, and its marginal $\mathbb{P}_X^*$ is supported on at most $K + 1$ points. We demonstrate the structure of the worst-case distribution in Figure 2.3 (left). In this plot, the support of $\widehat{\mathbb{P}}$ is represented by '•', and we have $K = 3$, $n_k = 3$, $k = 1, 2, 3$ and $k_0 = 2$. These points are transported to '⋆'s, which form the worst-case distribution $\mathbb{P}^*$. For $k = 1, 3$, we observe that $\widehat{x}_k$ is transported to x_k^*, and the conditional distribution $\mathbb{P}^*_{Z|X=x_k^*}$ has the same structure as the conditional distribution $\widehat{\mathbb{P}}_{\widehat{Z}|X=\widehat{x}_k}$, both supported on 3 points with identical probability mass function $(\hat{p}_{ki})_{i=1,2,3}$. Furthermore, $\widehat{x}_2$ is split into two values $\overline{x}_2$ and $\underline{x}_2$, and the conditional distributions $\mathbb{P}^*_{Z|X=\underline{x}_2}$, $\mathbb{P}^*_{Z|X=\overline{x}_2}$ have the same structure as the conditional distribution $\widehat{\mathbb{P}}_{\widehat{Z}|X=\widehat{x}_2}$, both supported on 3 points with identical probability mass function $(\hat{p}_{2i})_{i=1,2,3}$.

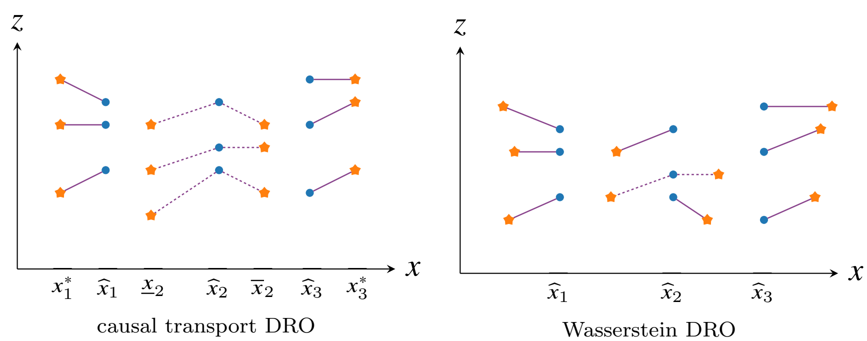


Figure 2.3: Structure of the worst-case distributions

As a comparison, on the right side of Figure 2.3, we plot a worst-case distribution resulting from Wasserstein DRO. According to Gao and Kleywegt (2016), the worst case distribution can be supported on $N + 1$ points, and points with the same x-value could have different x-values after transportation or splitting. The conditional distributions of the worst-case distribution change completely, each of which is a Dirac measure. This example illustrates that the worst-case distribution of the causal transport distance DRO preserves the conditional information structure of the nominal distribution, whereas the Wasserstein DRO fails to do so.

2.3 Finding the Optimal Decision Rule

In this section, we study the outer optimization over decision rules in (P). As a direct consequence of Theorem 2, problem (P) is equivalent to the following:

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

In particular, if we define $\|z-\widehat{z}\|_{\mathcal{Z}} := \infty\mathbf{1}\{z\neq\widehat{z}\}$, which is often used when the side information is relatively accurate, then (D) is simplified to

$$v_{\mathsf{D}} := \inf_{\substack{f\in\mathcal{F}\\ \lambda\geq 0}} \left\{ \lambda\rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\sup_{x\in\mathcal{X}} \left\{ \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Psi(f(x),\widehat{Z}) |\widehat{X} \right] - \lambda\|x-\widehat{X}\|^p \right\} \right] \right\}. \tag{2.1}$$

The tractability of the optimization over $f\in\mathcal{F}$ depends on the class of decision rules $\mathcal{F}$. If $\mathcal{F}$ admits a finite-dimensional parameterization, such as affine class, then the problem (D) is a finite-dimensional optimization and we identify cases where the overall problem can be solved by off-the-shelf convex programing solvers (Section 2.3.1). Otherwise if $\mathcal{F}$ is a non-parametric class, and particularly the class of all decision rules, then the optimization over $\mathcal{F}$ is an infinite-dimensional functional optimization, yet still, we identify cases where the overall problem can be solved efficiently (Section 2.3.2).

2.3.1 Optimizing over Affine Decision Rules

In this subsection, we provide tractable formulations when $\mathcal{F}$ is the affine class. Suppose affine functions in $\mathcal{F}$ are parametrized by Θ:

$$\mathcal{F}_\Theta = \left\{ x\mapsto B^\top x+\delta:\ (B,\delta)\in\Theta \right\} \tag{2.2}$$

where Θ is a finite-dimensional convex set.

Our first result shows that (2.1) is tractable when Ψ is affine in the decision variable w.

Corollary 1. *Suppose $\mathcal{F} = \mathcal{F}_\Theta$ defined in (2.2), and $\Psi(\cdot, z)$ is affine for every z:*

$$\Psi(w, z) = \ell^z(w) =: \beta(z)^\top w + b(z).$$

Set

$$\beta_k := \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}[\beta(\widehat{Z})|\widehat{X} = \widehat{x}_k], \quad b_k := \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}[b(\widehat{Z})|\widehat{X} = \widehat{x}_k].$$

Then the dual problem (2.1) is equivalent to the following convex program

$$\inf_{\lambda \geq 0, (B,\delta)\in\Theta} \left\{ \lambda\rho^p + \sum_{k=1}^{K} \Big(\sum_{i=1}^{n_k} \hat{p}_{ki} \Big) \cdot \Big(\beta_k^\top (B^\top \widehat{x}_k + \delta) + b_k + R_p\left(\lambda, |B\beta_k|\right) \Big) \right\},$$

where $R_p : \mathbb{R}_+^2 \to \mathbb{R} \cup \{\infty\}$ is a convex function defined as

$$R_p(\lambda, \mu) := \sup_{t \geq 0} \{\mu t - \lambda t^p\} = \begin{cases} \infty \mathbf{1}_{\{\lambda < \mu\}}, & p = 1, \\ \lambda(p-1)\left(\frac{\mu}{\lambda p}\right)^{\frac{p}{p-1}}, & p > 1. \end{cases} \tag{2.3}$$

Next, we consider the case when $\Psi(w, z)$ is bilinear. For the sake of tractability, we restrict ourselves to the case $p = 2$.

Corollary 2. *Suppose $\mathcal{F} = \mathcal{F}_\Theta$ as defined in (2.2) and $\Psi(w, z)$ is bilinear:*

$$\Psi(w, z) = w^\top A z + \beta^\top w + \alpha^\top z + b.$$

Set

$$\beta_k = \beta + A\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}[\widehat{Z}|\widehat{X} = \widehat{x}_k], \quad b_k = b + \alpha^\top \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}[\widehat{Z}|\widehat{X} = \widehat{x}_k].$$

Then (D) *with $p = 2$ is equivalent to the following convex program*

$$\inf_{\substack{(B,\delta)\in\Theta \\ \frac{1}{2}\|BA\|_2 < \lambda}} \Bigg\{ \lambda\rho^2 + \sum_{k=1}^{K} \Big(\sum_{i=1}^{n_k} \hat{p}_{ki} \Big) \cdot \Bigg(\beta_k^\top (B^\top \widehat{x}_k + \delta) + b_k + \frac{|A^\top (B^\top \widehat{x}_k + \delta) + \alpha|^2}{4\lambda} + \frac{1}{4}\Big[\frac{A(A^\top(B^\top \widehat{x}_k + \delta) + \alpha)}{2\lambda} + \beta_k\Big]^\top \Big[\frac{1}{4\lambda}(BA)(BA)^\top - \lambda\,\mathrm{Id}\Big]^{-1} \Big[\frac{A(A^\top(B^\top \widehat{x}_k + \delta) + \alpha)}{2\lambda} + \beta_k\Big] \Bigg) \Bigg\}.$$

Note that the problem above can be written as a semi-definite program.

Example 10 (CVaR Portfolio Optimization)**.** *Consider the portfolio optimization problem defined in Example 9 with $p = 2$. Rewrite $\mathcal{F}$ as*

$$\mathcal{F} = \left\{ x \mapsto Bx + w_0 : B = \left(\mathrm{Id} - \frac{1}{m} \mathbf{1}\mathbf{1}^\top \right) A \text{ for some } A \in \mathbb{R}^{m \times d}, \right\}$$

where $w_0 = \frac{1}{m}\mathbf{1}$. We denote $\mathscr{B} = \left\{ \left(\mathrm{Id} - \frac{1}{m}\mathbf{1}\mathbf{1}^\top \right) A : A \in \mathbb{R}^{m\times d} \right\}$. The dual problem can be written as

$$\inf_{B \in \mathscr{B}, \lambda > 0} \left\{ \lambda \rho^2 + \mathbb{E}_{\widehat{X}} \left[\sup_x \left\{ \mathbb{E}_{\widehat{Z}} \left[\sup_z \left\{ (Bx + w_0)^\top z - \lambda \|z - \widehat{Z}\|^2 \right\} \right] - \lambda \|x - \widehat{X}\|^2 \right\} \right] \right\}.$$

With Corollary 2, this is equivalent to the following semi-definite program as calculated in B.5:

$$\begin{aligned} \inf_{B \in \mathscr{B}, \lambda > 0, Y \succeq 0} \quad & \lambda \left(\rho^2 - \mathbb{E}[\|\widehat{X}\|^2] \right) + 2 w_0^\top \mathbb{E}[\widehat{Z}] + \frac{4}{\lambda} w_0^\top w_0 + \mathrm{tr}(Y) \\ \text{s.t.} \quad & B^\top B \preceq \mathrm{Id}, \\ & \begin{pmatrix} \lambda(\mathrm{Id} - B^\top B) & \begin{pmatrix} -\lambda \, \mathrm{Id} & \lambda B^\top & 2B^\top \end{pmatrix} S_1 \\ S_1^\top \begin{pmatrix} -\lambda \, \mathrm{Id} \\ \lambda B \\ 2B \end{pmatrix} & Y \end{pmatrix} \succeq 0 \end{aligned}$$

where $S_1^2 = \mathbb{C}\mathrm{ov} \begin{pmatrix} \widehat{X} \\ \widehat{Z} \\ w_0 \end{pmatrix}$. ♣

2.3.2 Optimizing over All (Non-parametric) Decision Rules

In this subsection, we consider $\mathcal{F}$ to be unrestricted and contains all measurable functions $\{f : \mathcal{X} \to \mathcal{D}\}$. In general, this infinite dimensional problem is hard to solve. Nonetheless, below we provide a tractable way to find the optimal decision rule for this problem in certain settings.

Recall that our dual reformulation in Theorem 2 states that

$$v_{\mathsf{D}} = \min_{f:\mathcal{X}\to\mathcal{D}} \min_{\lambda\geq 0} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\sup_{x\in\mathcal{X}} \left\{ \varphi(f(x);\lambda,\widehat{X}) - \lambda\|x-\widehat{X}\| \right\} \right] \right\}, \tag{2.4}$$

where $\varphi(w;\lambda,\widehat{x}) := \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}}\left\{\Psi(w,z)-\lambda\|z-\widehat{Z}\|\right\}|\widehat{X}=\widehat{x}\right]$. By replacing $\mathcal{X}$ with $\operatorname{supp}\widehat{\mathbb{P}}$, we define the in-sample dual problem as

$$\begin{aligned} v_{\widehat{\mathsf{D}}} &:= \min_{\substack{f:\mathcal{X}\to\mathcal{D}\\ \lambda\geq 0}} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\max_{1\leq k\leq K} \left\{ \varphi(f(x_k);\lambda,\widehat{X}) - \lambda\|x_k-\widehat{X}\| \right\} \right] \right\} && (2.5)\\ &= \min_{\substack{\widehat{w}\in\hat{\mathcal{F}}\\ \lambda\geq 0}} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\max_{1\leq k\leq K} \left\{ \varphi(\widehat{w}(x_k);\lambda,\widehat{X}) - \lambda\|x_k-\widehat{X}\| \right\} \right] \right\}, && (2.6) \end{aligned}$$

where the second equality holds because the objective value in (2.5) depends only on the value of f on $\operatorname{supp}\widehat{\mathbb{P}}$. Note that (2.6) is a finite-dimensional convex optimization problem with $K+1$ decision variables in the outer minimization.

Theorem 4. *Suppose $p=1$, $\mathcal{D}\subset\mathbb{R}$ is a convex subset, and $\Psi(w,z)$ is convex in w. Let $(\lambda^*,\widehat{w}^*)$ be a minimizer to the in-sample dual problem (2.6). Denote $\varphi_k(y) := \varphi(y;\lambda^*,\widehat{x}_k)$, $y_k := \widehat{w}^*(\widehat{x}_k)$, and $\phi_k := \max_j\{\varphi_k(y_j)-\lambda^*\|\widehat{x}_k-\widehat{x}_j\|\}$. For $x\in\mathcal{X}$, define*

$$I_k(x) := \{y\in\mathcal{D} : \varphi_k(y) \leq \lambda^*\|x-\widehat{x}_k\| + \phi_k\}.$$

Then the intersection of $I_k(x)$'s is nonempty, and every decision rule $f^*\in\mathcal{F}$ satisfying $f^*(x)\in\cap_k I_k(x)$ for all $x\in\mathcal{X}$ is a minimizer to (2.4). Moreover, let (λ^*,f^*) be a minimizer to the dual problem (D), then $(\lambda^*,\widehat{w}^*)$ is a minimizer to (2.6), and $f^*(x)\in\cap_k I_k(x)$ defined above.

Theorem 4 shows that problems (2.4) and (2.6) share the same optimal dual variable λ^*, and to solve the infinite-dimensional optimization over decision rules (2.4), it suffices to first solve a finite-dimensional robust in-sample optimization (2.6) and then extend the robust optimal in-sample decision rule to $\mathcal{X} \setminus \operatorname{supp} \widehat{\mathbb{P}}$ such that it is optimal to the original problem. Note that once the in-sample problem (2.6) is solved, the values y_k, ϕ_k are immediately available and the set I_k is defined precisely. There may be more than one way to extend the in-sample robust optimal decision rule $\widehat{w}$ to the entire space, as long as it belongs to the range of $\cap_k I_k(x)$.

The proof idea of Theorem 4 is as follows. Observe that $v_{\mathsf{D}} \geq v_{\hat{\mathsf{D}}}$, since the inner supremum in (2.4) is taken with respect to a larger set compared with the maximization in (2.5). To see the other direction, the main step is to show $I_k(x)$ has a nonempty intersection. Once this is shown, using simple algebra it is easy to verify that $f^*(x) \in \cap_k I_k(x)$ attains the value $v_{\hat{\mathsf{D}}}$, thereby v_{D} is dominated by the objective value of f^* which equals $v_{\hat{\mathsf{D}}}$. To show $I_k(x)$ has a nonempty intersection, it suffices to show they pairwise intersect because they are one-dimensional intervals. This can be established using the convexity of φ.

Remark 3 (Comparison with the Shapely policy in Zhang et al. (2023))**.** *In Zhang et al. (2023), the authors study* (1.3) *with Wasserstein uncertainty sets, focusing on the newsvendor cost. They show that when optimization over all decision rules, the optimal decision rule, called Shapely policy, can be found by first solving for the in-sample Wasserstein robust optimal decision rule $\widehat{w}_{\mathsf{W}}$, then extending to the entire space by solving*

$$f^*(x) = \arg\min_{y \in \mathbb{R}} \max_k \frac{|y - \widehat{w}_{\mathsf{W}}(\widehat{x}_k)|}{\|x - \widehat{x}_k\|},$$

which minimizes the maximal slope. Using the same idea, if we define

$$f^*_\infty(x) := \arg\min_{y \in \mathbb{R}} \max_k \frac{|y - \widehat{w}^*(\widehat{x}_k)|}{\|x - \widehat{x}_k\|}, \tag{2.7}$$

where $\widehat{w}^*(\widehat{x}_k)$'s are defined in Theorem 4, then it can be verified that $f^*_\infty(x) \in \cap_k I_k(x)$. Therefore, this shows that $f^*_\infty(x)$ defined a robust optimal decision rule for (2.4)*. Note* *that we use the subscript* ∞ *to indicate the* ∞*-norm (maximum) of the slope function* $k \mapsto \frac{|y-\widehat{w}^*(\widehat{x}_k)|}{\|x-\widehat{x}_k\|}$.

Differently, we can define another decision rule that minimizes the 1*-norm of the slope function,*

$$f_1^\dagger(x) := \arg\min_{y\in\mathbb{R}} \sum_k \frac{|y-y_k|}{\|x-\widehat{x}_k\|}.$$

The resulting decision rule may not necessarily optimal, but we can always truncate its values to force them falling into $\cap_k \boldsymbol{I}_k(x)$ *and thereby making it robust optimal. Namely, if we use* $\overline{\boldsymbol{I}}(\cdot)$ *and* $\underline{\boldsymbol{I}}(\cdot)$ *to represent the upper and lower bound of the region* $\cap_k \boldsymbol{I}_k(x)$*, the we define*

$$\bar{f}_1^\dagger(x) := \max\Big(\underline{\boldsymbol{I}}(x),\ \min\big(f_1^\dagger(x), \overline{\boldsymbol{I}}(x)\big)\Big). \tag{2.8}$$

We denote the truncated decision rule as $\bar{f}_1^\dagger(x)$*.*

We illustrate the two robust optimal decision rules defined above using a conditional median estimate problem with $Z = \mu(X) + \varepsilon$*,* $\varepsilon \sim \mathcal{N}(0,1)$*,* $\mu(x) = \sin(2x) + 2\exp(-16x^2)$*.*

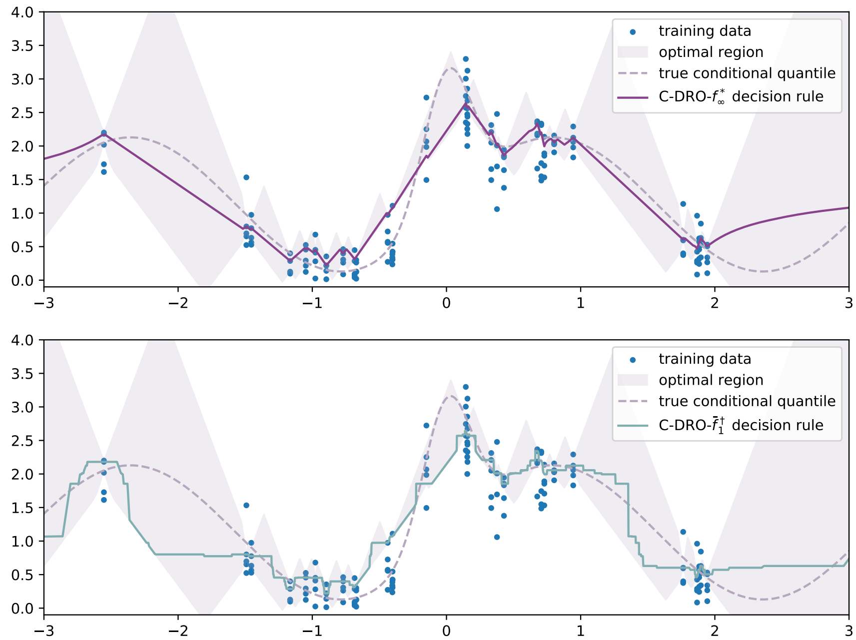


Figure 2.4: Two robust optimal decision rules f^*_∞ and $\overline{f}^\dagger_1$ of a median estimation problem

Example 11 (Conditional median estimation)**.** *Consider the feature-based newsvendor problem in Example 7. When* $h = b = 1$*, this is equivalent to conditional median estimation. As detailed in B.5, the in-sample dual problem* (2.6) *can be transformed into a linear programming problem*

$$
\begin{aligned}
\inf_{\{w_k\}_k, \lambda \geq 0} \; & \lambda\rho + \frac{1}{n}\sum_{k=1}^{K} c_j \\
\text{s.t.} \quad & c_j \geq \sum_{i=1}^{n_j} c_{kji} - \lambda n_j \|\widehat{x}_k - \widehat{x}_j\|, \forall j, k \\
& c_{kji} \geq w_k - \widehat{z}_{ji}, \forall k, j, i \\
& c_{kji} \geq \widehat{z}_{ji} - w_k, \forall k, j, i \\
& \lambda \geq 1
\end{aligned}
$$

Example 12 (Personalized Pricing)**.** *Consider the personalized pricing problem in*

Example 8. By Theorem 2, its strong dual problem can be written as

$$\inf_{\substack{f:\mathcal{X}\to\mathbb{R}\\ \lambda\geq 0}} \left\{\lambda\rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\sup_{x\in\mathcal{X}}\left\{\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}}\left\{-f(x)z^\top\begin{pmatrix}f(x)\\1\end{pmatrix} - \lambda\|z-\widehat{Z}\|^p\right\}|\widehat{X}\right] - \lambda\|x-\widehat{X}\|^p\right\}\right]\right\}.$$

In the case of $p = 1$, we notice that f is real-valued and $\mathbf{\Psi}$ is convex in w, so we may use Theorem 4 to reformulate the problem as

$$\inf_{\substack{\widehat{w}:\widehat{\mathcal{X}}\to\mathbb{R}\\ \lambda\geq 0}} \left\{\lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\max_{1\leq k\leq K}\left\{\varphi(\hat{f}(\widehat{x}_k);\lambda,\widehat{X}) - \lambda\|\widehat{x}_k - \widehat{X}\|\right\}\right]\right\}.$$

where

$$\varphi(w;\lambda;\widehat{x}) = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}}\left\{-wz^\top\begin{pmatrix}w\\1\end{pmatrix} - \lambda\|z-\widehat{Z}\|\right\}|\widehat{X}=\widehat{x}\right].$$

A detailed calculation could be found in B.5. In case when $\|z-\hat{z}\|_{\mathcal{Z}} = \infty\mathbf{1}_{\{z\neq\widehat{z}\}}$, by adding dummy variables, this can be transformed into a quadratic programming problem

$$\begin{aligned}
\underset{\substack{w_k\geq 0\\ \lambda\geq 0}}{\text{minimize}} \quad & \lambda\rho + \sum_{j\in[K]}\hat{p}_j c_j\\
\textit{subject to} \quad & c_j + w_k\bar{z}_k^\top\begin{pmatrix}w_k\\1\end{pmatrix} + \lambda\|\widehat{x}_k - \widehat{x}_j\| \geq 0, \qquad \forall k\in[K].
\end{aligned}$$

where $\hat{p}_k = \sum_{i=1}^{n_k}\hat{p}_{ki}$ and $\bar{z}_k = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\widehat{Z}|\widehat{X}=\widehat{x}_k\right]$.

Chapter 3: Dynamically Information Acquisition and Optimal Decision Making

3.1 Introduction

Many real-world analytics problems involve two significant challenges: estimation and optimization. Due to the typically complex nature of each challenge, the standard paradigm is estimate-then-optimize. By and large, machine learning or human learning tools are intended to minimize estimation error and do not account for how the estimations will be used in the downstream optimization problem (such as decision-making problems). In contrast, there is a line of literature in economics focusing on exploring the optimal way to acquire information and learn dynamically to facilitate decision-making Wald (1947); Arrow et al. (1949); Moscarini and Smith (2001); Zhong (2022); Fudenberg et al. (2018); Che and Mierendorff (2019). However, most of the decision-making problems considered in this line of work are static (i.e., one-shot) problems which over-simplify the structures of many real-world problems that require dynamic or sequential decisions.

Nowadays, real-world problems that require both learning and sequential decisions making have been emerging rapidly. One of the major playgrounds would be e-commerce platforms. E-commerce marketplaces provide easy access for sellers to compete and sell new products. These marketplaces are also platforms on which buyers publicize their reviews of purchased products. On these platforms, we often encounter situations where a new product of unknown quality is introduced to challenge existing products of known quality. A buyer who is choosing between the new product and several old products could collect information on the new product by reading the reviews, and decide whether to purchase the new product or not. Meanwhile, the buyer could also access implications for the pricing of old (existing) products on the platform. As most of the E-commerce platforms offer flexible return and exchange services (especially for the new product), buyers could make the first purchase of either the new product or the other old ones after a preliminary investigation. If the buyer is unsatisfied with the initial purchase, she can return the product and switch to the alternative (possibly with a small penalty). Similar decision-making problems with learning and opportunity to reverse can also be found in airline and hotel bookings. When a traveler is looking for an airline and flight tickets for her trip, she can make a refundable purchase after some investigation of different options. She can continue to collect information on these options and keep checking the prices after the initial purchase. If she realizes that there is a better option later on, she can pay a small fee to cancel the first ticket and switch to the one she prefers better.

Motivated by the above-mentioned examples, we study a decision-making problem with learning and an opportunity to reverse (referred to as "reversible decisions") over an infinite time horizon. In this problem, the decision maker (DM) can choose between two products, A and B. Product A is an established good that creates a known return μ for the DM. Product B is recently introduced, and its true value (expected return) $\Theta \in \{l, h\}$, is unknown to the DM. We consider a framework in which the DM learns Θ under a noisy signal Y_t over a period of time before making her initial decision. This noisy signal can be understood as the posted reviews of Product B on the E-commerce platform. In addition, if the DM chooses the new Product B as the first choice, she has an opportunity to switch to Product A later on at a cost. Mathematically, this can be formulated as a sequential decision-making problem in two periods. In the first period (up to a stopping time τ chosen by the

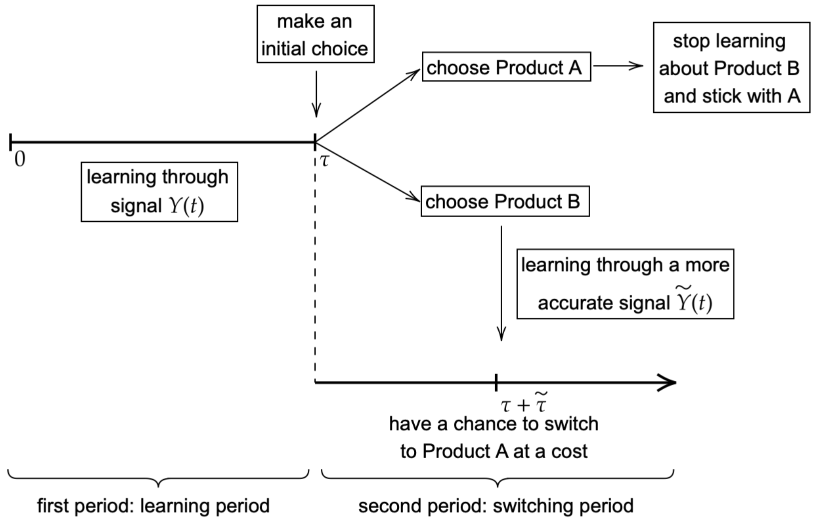


Figure 3.1: Timeline of the decisions.

DM), the DM constructs a posterior estimate for Θ using the observed noisy signals Y_t and the Bayesian formula. At the stopping time τ, the DM makes her first choice between Product A and Product B. If the DM decides to make a conservative decision by selecting the well-known Product A, the DM will stop observing any information about Product B and stick with Product A. However, if the DM decides to try out Product B at τ, the DM will have a more accurate observation process $\widetilde{Y}_t$ for Product B when start using it, and there will be no additional cost associated with this new observation, since they are already using Product B. In this scenario, the DM will have one opportunity to reverse her initial decision and switch to Product A at a fixed cost ($c_S > 0$). See Figure 3.1 for the timeline of the decisions.

As a benchmark, we also analyze a classic setting (referred to as "irreversible decision") with a static decision-making problem after learning, which is popular in the economic literature Moscarini and Smith (2001); Zhong (2022); Fudenberg et al. (2018); Che and Mierendorff (2019). Specifically, the DM chooses between Products

A and B at the stopping time and sticks to this choice for the rest of the horizon (regardless of the evolution of the posterior estimate).

Related Literature. Our work is related to two lines of literature, one on decision-making under information acquisition and one on stochastic control with filtering theory.

Decision-making under information acquisition has a long history which dates back to the seminal paper by Wald (1947), in which the flow of information is assumed to be fully exogenous and the decision-maker (DM) controls the decision time and action choice. Specifically, Wald (1947) formulates an optimal stopping problem where the whole space of beliefs can be partitioned into stopping region and continuation region. Early works along this direction have been focused on the *duration of search* when there is a cost per unit of time when searching for information. For example, Moscarini and Smith (2001) generalizes the framework in Wald (1947) to an information intensity control problem where information is modeled as the trajectory of a Brownian motion with drift representing the state and variance representing the intensity. A similar setting is used in Fudenberg et al. (2018) to study the trade-off between information speed and precision. Some recent works have been trying to address the *information selection issue* when there are multiple information sources and when the DM has limited capacity to process information. In this paradigm, Che and Mierendorff (2019); Mayskaya (2022) study the allocation of limited attention when there are multiple sources of information that are modeled by Poisson bandits. Liang et al. (2017) has a similar focus and assumes that the DM sequentially samples from a finite set of Gaussian signals, and wants to predict a persistent multi-dimensional state at an unknown final period. The authors show that optimal choice from Gaussian information sources is myopic. Recently, Zhong (2022) studies a setting where the DM has access to both Gaussian signals (which are available in continuous time but very noisy) and Poisson signals (which are less frequent but contain precise information). However, despite the importance of addressing the search duration and information selection questions, most of the final decisions considered in these papers are rather over-simplified (e.g., a one-shot decision between two products) or extremely abstract (a general form of the terminal cost with no further analysis). As the final decision is an integral part of the framework and has a non-negotiable impact on the learning process and information search behavior, it is crucial to discuss some realistic downstream decision-making scenarios and examine how these tasks affect learning behaviors. To the best of our knowledge, this aspect has been largely missing in the literature.

For stochastic control with filtering theory, the DM has partial information about the underlying system which is often modeled by a stochastic differential equation (SDE). The DM will first use the observation process to form an estimate of the state of the system and then, thanks to the separation principle Sirjaev (1973), construct the control signal as a function of this estimate. For this line of work, the observation process is often assumed to be obtained *at no cost* (i.e., for free). The main focus, on the other hand, is on the construction of the filtering process and the solvability of the associated control problem Mitter (1996); Sorenson (1976). Most of the studies have been focused on linear-quadratic problems Morris (1976); Anderson and Moore (2007); Stengel (1994). This is because a tractable finite-dimensional Kalman-Bucy filter could be derived in explicit form when the underlying SDE is linear and the associated control problem could be solved through the Riccati system when the cost function is quadratic. Although sharing some common ingredients with our framework such as using Bayesian formula to estimate unknown quantities, the partially observed quantities considered in this line of work are often more complex (i.e., unknown processes) than those considered in the literature on information acquisition (i.e., unknown variables). More importantly, the frameworks considered in stochastic control and filter theory can not be applied directly to the situation where the DM pays a cost to process information and to facilitate the understanding on how the cost of information affects the behavior of the DM.

Our Contributions. As the first attempt to understand information acquisition and sequential decision-making from an integrated perspective, we study a discrete choice model where the decision maker (DM) can make a first choice between two products, Product A with a deterministic return and Product B with an unknown return. The DM can collect information from a Gaussian signal on the unknown product B and decide when to stop and make a decision. We introduce flexibility to the framework by allowing the DM to switch back to Product A later if her initially choosing Product B, at a constant cost. Under this setting, we have the following main contributions:

(I) With the notion of viscosity solution, we establish the regularity property of the value function for a general class of decision-making problems that unifies reversible and irreversible decisions.

(II) For the benchmark of "irreversible decisions," we characterize the continuation and stopping regions to describe the policies from the DM. In addition, we conduct a sensitivity analysis for all model parameters to fully analyze the behaviors of the DM. We make novel contributions in using viscosity sub/super solutions to show asymptotic behaviors of the DM with respect to model parameters. Finally, when the running cost function is reduced to a constant, the optimal policy can be constructed semi-explicitly.

(III) In the "reversible decisions" setting, we consider two types of signal processes for the second period if the DM chooses B as the initial choice: a Poisson signal that reveals the true value of Product B and a Gaussian signal with a smaller variance (compared to the first period). We quantify the optimal policies for both cases and compare the results with the "irreversible decisions" setting. Finally, we also provide the monotonicity analysis on how the DM behaves when the cost of reverse is increasing.

We note that the set-up for the "irreversible decisions" part is akin to the set-up in Moscarini and Smith (2001), where the precision of the information signal can be further controlled at a cost. However, the mathematical tools used in our framework are remarkably different. Moscarini and Smith (2001) uses the smooth-fit principle to characterize the continuation and stopping regions under the assumption that the value function is $\mathcal{C}^2$. In contrast, we construct proper sub/super viscosity solutions and utilize the comparison principle to identify the continuation and stopping regions, and to establish the asymptotic and monotonicity behaviors with respect to model parameters. Some of our constructions are non-trivial and delicate, particularly in the sensitivity analysis with respect to the product values h and l. However, the discussion on sensitivity analysis in Moscarini and Smith (2001) is rather limited and lacks of mathematical details.

In addition, compared to the "irreversible decisions", we observe that:

- With the opportunity to pull back and reverse if choosing Product B as the initial choice, the DM will spend *less effort* to explore the value of Θ during the first period.

- When the cost to reverse is small, the DM will make her first choice *earlier* compared to the situation when the cost is expensive.

- When the cost of revising the initial decision is higher, the DM is more willing to choose the well-known Product A and less willing to choose Product B as the initial choice.

3.2 The Models for Irreversible and Reversible Decisions

We introduce the models for single (irreversible) decisions and for sequential decisions.

3.2.1 Single Decisions and Irreversible Choices

A decision maker (DM) acts in an infinite horizon and is concerned with purchasing one of two products, denoted by A and B. Product A has known performance $\mu > 0$ while product B is a new, not yet established product. Its performance is modeled by a random variable Θ that may take only two values, l and h, with $0 < l < \mu < h$.

Information acquisition and its cost: The levels l and h are known to the DM but not the actual performance of the new product. However, she learns indirectly about Θ through a signal process $(Y_t)_{t\geq 0}$, solving

$$dY_t = \Theta\, dt + \sigma dW_t, \quad Y_0 = 0, \tag{3.1}$$

with $(W_t)_{t\geq 0}$ being a standard Brownian motion in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. She, then, applies a Bayesian framework and dynamically updates her views, based on the information generated by Y. Specifically, she acquires the belief process $(q_t)_{t\geq 0}$,

$$q_t := \mathbb{P}\left[\Theta = h | \mathcal{F}_t^Y\right], \tag{3.2}$$

where $\mathcal{F}_t^Y := \sigma(Y_s; 0 \leq s \leq t)$. Classical results from filtering theory (see, for example, Karatzas and Zhao (2001), also Moscarini and Smith (2001); Zhong (2022) and others) yield that q_t is a martingale, solving

$$dq_t = \frac{h-l}{\sigma} q_t (1 - q_t)\, dZ_t, \quad q_0 = q \in [0, 1], \tag{3.3}$$

with $(Z_t)_{t\geq 0}$ being a standard Wiener process in $\left(\Omega, \mathcal{F}, \left\{\mathcal{F}_t^Y\right\}, \mathbb{P}\right)$. Note that the states 0 and 1 are absorbing, i.e. if the initial belief $q_0 = 0, 1$ then $q_t = 0, 1$, for $t > 0$, respectively.

To have access to the above signal process, the DM encounters *information acquisition costs*. It is assumed that they occur at rate $C(q)$ per unit of time, namely, the process $(C_t)_{t\geq 0}$, $C_t := \int_0^t C(q_s) ds$ is the cumulative cost of acquiring information in $[0, t]$. It is assumed that there are neither initial fixed costs nor cost jumps thereafter.

Examples: i) $C(q) = c > 0$. Then, $C_t = ct$, which essentially measures (up to a multiplicative constant) the time the DM spends in acquiring new information.

ii) $C(q) = \mathrm{Var}[\Theta|q]$. Then, the process $C_t := \int_0^t Var\left[\Theta | q_s\right] ds$, which expresses the cumulative uncertainty in DM's belief. Similarly, we may choose $C(q) = \sqrt{\mathrm{Var}[\Theta|q]}$.

The single (irreversible) decision problem: The DM seeks information about the new product B before she decides which product to choose. At a decision time, say τ, the exploration period ends, a product is chosen and she receives reward $G(q_\tau)$. The problem terminates once this decision is made and no follow-up choices are available.

The DM's value function $V : [0,1] \to \mathbb{R}_+$ is defined as

$$V(q) = \sup_{\tau \in \mathcal{T}_Y} \mathbb{E}\left[-\int_0^\tau e^{-\rho t} C(q_t) dt + e^{-\rho\tau} G(q_\tau) \middle| q_0 = q \right],$$

with $(q_t)_{t\geq 0}$ solving (3.3) and $\mathcal{T}_Y$ being the set of stopping times measurable with respect to $\mathcal{F}^Y$. This is the classical case of a single decision setting under costly information acquisition. It has been extensively analyzed in the literature (see, for example) and, primarily, under the specific reward

$$G(q) = \max(\mu, qh + (1-q)l), \tag{3.4}$$

the maximum of the known return μ of product A and the expected return $qh+(1-q)l$ of product B under belief q. The DM's aim is then given by

$$V(q) = \sup_{\tau \in \mathcal{T}_Y} \mathbb{E}\left[-\int_0^\tau e^{-\rho t} C(q_t) dt + e^{-\rho\tau} \max(\mu, q_\tau h + (1-q_\tau)l) \middle| q_0 = q \right]. \tag{3.5}$$

3.2.2 Reversible Choices and Sequential Decisions

Building on the previous classical irreversible decision setting, we introduce a general model in which the DM has the optionality to reverse her initial decision and, furthermore, to refine the information acquisition source. Namely, we assume that the unknown product product B is returnable for exchange with A. If exchanged, there is a return fee whose characteristics are known to the DM at initial time 0. To keep the extended model tractable, it is assumed that the known product is not exchangeable (this case can be easily incorporated herein). We denote, as before, the time of the DM's first decision - choose A or B - by τ. If she chooses product A, the problem terminates as A is not returnable.

If the DM chooses B at τ, she immediately commences its use. In parallel, she continues exploring B by using it and, also, by learning about it. The information acquisition is now done via a *new signal* process $\left(\widetilde{Y}_t\right)_{t\geq\tau}$ (see, (3.30) and (3.62)) which, in analogy to (3.1), generates an associated belief process, $(\widetilde{q}_t)_{t\geq\tau}$. The DM may decide to keep B or exchange it for A, say at time $\tau+\widetilde{\tau}$. If she returns it, she encounters penalty $R(\widetilde{q}_{\tau+\widetilde{\tau}})$, with the form of function $R(.)$ known at 0, and the overall decision process terminates. Therefore, if she chooses B at time τ, a new optimization problem is being generated that incorporates the optionality of its exchange,

$$U_B(q_\tau) \;:= \sup_{\widetilde{\tau}\in\mathcal{T}_{\widetilde{Y}}} \widetilde{\mathbb{E}}\left[\int_\tau^{\tau+\widetilde{\tau}} e^{-\rho t}\left(-\widetilde{C}\left(\widetilde{q}_t\right)+\widetilde{m}\left(\widetilde{q}_t\right)\right)dt + e^{-\rho(\tau+\widetilde{\tau})}\left(\mu - R(\widetilde{q}_{\tau+\widetilde{\tau}})\right)\middle|\, \widetilde{q}_0 = q_\tau\right]. \tag{3.6}$$

Herein, $\widetilde{C}\left(\widetilde{q}\right)$ is the information acquisition cost function for the modified signal $\widetilde{Y}$ and $\widetilde{m}(\widetilde{q})$ is the payoff the DM accumulates from using product B in $[\tau,\tau+\widetilde{\tau}]$. Furthermore, in analogy to $\mathcal{T}_Y$, $\mathcal{T}_{\widetilde{Y}}$ is the set of stopping times in $[\tau,\tau+\widetilde{\tau}]$ associated with the new filtration generated by the modified signal process. We make this all precise in Section 3.4.

We now formulate the *integrated* optimization problem in $[0,\infty]$,

$$V(q) = \sup_{\tau\in\mathcal{T}_Y}\mathbb{E}\left[-\int_0^\tau e^{-\rho t}C(q_t)dt + e^{-\rho\tau}\max\left(\mu, U_B(q_\tau)\right)\middle|\, q_0 = q\right]. \tag{3.7}$$

In other words, the DM solves an optimal stopping problem similar to the single-decision one, but with modified payoff

$$G(q_\tau) := \max\left(\mu, U_B(q_\tau)\right),$$

at the first decision time, τ.

3.3 The Core Optimal Stopping Problem: Regularity Results and Sensitivity Analysis

This section is dedicated to the analysis of the optimal stopping problem (3.7)

$$V(q) = \sup_{\tau \in \mathcal{T}_Y} \mathbb{E}\left[-\int_0^\tau e^{-\rho t} C(q_t) dt + e^{-\rho \tau} G(q_\tau) \middle| q_0 = q \right], \quad q \in [0,1], \tag{3.8}$$

with the belief process $(q_t)_{t \geq 0}$ solving (cf. (3.3)),

$$dq_t = \frac{h-l}{\sigma} q_t (1-q_t)\, dZ_t, \quad q_0 = q \in [0,1],$$

and for *general* information costs and payoff functions $C(.)$ and $G(.)$, respectively; as mentioned earlier, only the cases $C(q) = c$ and $G(q) = \max(\mu, qh + (1-q)l)$ have been so far analyzed.

Assumption 2. *i) The information cost function $C : [0,1] \to \mathbb{R}_+$ is Lipschitz with Lipschitz constant $L > 0$.*

ii) The function $G : [0,1] \to \mathbb{R}_+$ is Lipschitz with Lipschitz constant $K > 0$, convex and non-decreasing.

Classical results yield the associated (OP) problem

$$(OP) \begin{cases} \min\left(\rho V(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2 (1-q)^2 V''(q) + C(q),\ V(q) - G(q)\right) = 0, \quad q \in [0,1], \\ V(0) = G(0) \quad \text{and} \quad V(1) = G(1). \end{cases} \tag{3.9}$$

The boundary condition $V(0) = G(0)$ follows since, if $q = 0, 1$ then $q_t = 0, 1$, $t \geq 0$, as 0 and 1 are absorbing states (cf. (3.3)). Then, (3.8) becomes

$$V(0) = \sup_{\tau \in \mathcal{T}_Y} \mathbb{E}\left[-\int_0^\tau e^{-\rho t} C(q_t) dt + e^{-\rho \tau} G(0) \middle| q_0 = q \right],$$

and using that $C(q) > 0$ and $\rho > 0$, we deduce for the optimal time $\tau^* = 0$, and thus $V(0) = G(0)$. Similar arguments yield that $V(1) = G(1)$.

For the reader's convenience, we highlight below the key steps for the derivation of the variational inequality in (3.9). If $q \in (0, 1)$, there are two admissible, in general suboptimal, policies. Specifically, i) the DM may immediately choose product A or B, without seeking any information about the latter or ii) she may spend some time, say $(0, \varepsilon]$ with ε small, learning about B before deciding which product to choose. Choices (i) and (ii) give, respectfully,

$$V(q) \geq G(q), \tag{3.10}$$

and

$$V(q) \geq \mathbb{E}\left[-\int_0^{\varepsilon} e^{-\rho t} C(q_t) dt + e^{-\rho \varepsilon} V(q_\varepsilon) \middle| q_0 = q \right].$$

Assuming that V is smooth enough, Itô's formula would give

$$V(q_\varepsilon) = V(q) + \int_0^{\varepsilon} \frac{1}{2} \left(\frac{h - l}{\sigma} \right)^2 q_t^2 (1 - q_t)^2 V''(q_t)\, dt + \int_0^{\varepsilon} \sigma V'(q_t)\, dZ_t,$$

where we used (3.3). Diving by ε and passing to the limit $\varepsilon \to 0$, we deduce

$$\rho V(q) - \frac{1}{2} \left(\frac{h - l}{\sigma} \right)^2 q^2 (1 - q)^2 V''(q) + C(q) \geq 0. \tag{3.11}$$

Because one of these two choices must be optimal, (3.10) or (3.11) must hold as equality and (3.9) follows.

Lemma 3. *The value function V is Lipschitz continuous on* $[0, 1]$.

Proof. We show that there exists a positive constant $\rho_0 = \rho_0(h, l, \sigma)$ such that for $\rho > 2\rho_0$,

$$|V(q_1) - V(q_2)| \leq \left(\frac{L}{\rho_0} + K \right) |q_1 - q_2|, \quad q_1, q_2 \in [0, 1].$$

First note that, for $q_1, q_2 \in [0,1]$, the function $b(q) := \frac{h-l}{\sigma} q(1-q)$ satisfies $b(q) \le \frac{h-l}{\sigma} q$ and

$$|b(q_1) - b(q_2)| \le \frac{h-l}{\sigma}((1 + q_1 + q_2)|q_1 - q_2|) \le 3\frac{h-l}{\sigma}|q_1 - q_2|.$$

Therefore, (see, for example, (Pham, 2009, Theorem 1.3.16)), there exists a positive constant $\rho_0 = \rho_0(h, l, \sigma)$ such that

$$\mathbb{E}\left[\sup_{0 \le u \le t} |q_u^{q_1} - q_u^{q_2}|\right] \le e^{\rho_0 t}|q_1 - q_2|,$$

For $\rho > 2\rho_0$, we can, similarly, prove that $\mathbb{E}\left[\sup_{t\ge 0} e^{-\rho t}|q_t^{q_1} - q_t^{q_2}|\right] \le |q_1 - q_2|$.

In turn, the Lipschitz properties of functions C and G (see Assumption 1) yield

$$\begin{aligned} |V(q_1) - V(q_2)| &\le \sup_{\tau \in \mathfrak{T}_Y} \mathbb{E}_{\tau \in \mathfrak{T}_Y}\left[\int_0^\tau e^{-\rho t}\left|C(q_t^{q_1}) - C(q_t^{q_2})\right| dt + e^{-\rho\tau}\left|G(q_\tau^{q_1}) - G(q_\tau^{q_2})\right|\right] \\ &\le L\mathbb{E}\left[\int_0^\infty e^{-\rho t}\left|q_t^{q_1} - q_t^{q_2}\right| dt\right] + K\mathbb{E}\left[\sup_{t \ge 0} e^{-\rho t}\left|q_t^{q_1} - q_t^{q_2}\right|\right] \\ &\le \left(\frac{L}{\rho - \rho_0} + K\right)|q_1 - q_2|. \end{aligned}$$

□

The connection between the value function and viscosity solutions of optimal stopping problems was established in Reikvam (1998) (see, also, Pham (2009)). Throughout, we will be using throughout viscosity arguments to carry out an extensive analysis of the problem and, for completeness, we also highlight the key steps in the following characterization result.

Theorem 5. *The value function V* (3.8) *is a viscosity solution to* (3.9), *unique in the class of Lipschitz continuous functions.*

Proof. We establish that V is a viscosity subsolution of (3.9) in $(0,1)$. For this, let $x \in (0,1)$ and consider a test function $\varphi \in \mathcal{C}^2((0,1))$ such that $(V-\varphi)(x) = \max(V-\varphi)$ and $(V-\varphi)(x) = 0$. We need to show that

$$\min\left(\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 x^2(1-x)^2\varphi''(x) + C(x), V(x) - G(x)\right) \le 0.$$

We argue by contradiction, assuming that both inequalities

$$\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 x^2(1-x)^2\varphi''(x) + C(x) > 0 \quad \text{and} \quad V(x) - G(x) > 0 \tag{3.12}$$

hold. Using the continuity of the involved functions, there would exist $\delta > 0$ such that

$$\rho\varphi(\bar{q}_t) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 \bar{q}_t^2(1-\bar{q}_t)^2\varphi''(\bar{q}_t) + C(\bar{q}_t) \ge \delta \text{ and } V(\bar{q}_t) - G(\bar{q}_t) \ge \delta, \ 0 \le t \le \hat{\tau}, \tag{3.13}$$

where $\bar{q}_t := q_t^x$ and $\hat{\tau}$ is the exit time of $\bar{q}_t$ from $[x-\delta, x+\delta]$. Applying Itô's formula to $e^{-\rho t}\varphi(\bar{q}_t)$, $t \in [0, \hat{\tau} \wedge \tau]$, $\tau \in \mathcal{T}_Y$, yields

$$\begin{aligned} V(x) &= \varphi(x) = \mathbb{E}\left[\int_0^{\hat{\tau}\wedge\tau} e^{-\rho t}\left(\rho\varphi(\bar{q}_t) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 \bar{q}_t^2(1-\bar{q}_t)^2\varphi''(\bar{q}_t)\right)dt + e^{-\rho(\hat{\tau}\wedge\tau)}\varphi(\bar{q}_{\tau\wedge\hat{\tau}})\right] \\ &\ge \mathbb{E}\left[-\int_0^{\hat{\tau}\wedge\tau} e^{-\rho t}C(\bar{q}_t)dt + e^{-\rho\tau}G(\bar{q}_\tau)\mathbf{1}_{\tau<\hat{\tau}} + e^{-\rho\hat{\tau}}V(\bar{q}_{\hat{\tau}})\mathbf{1}_{\hat{\tau}\le\tau}\right] + \delta\mathbb{E}\left[\int_0^{\hat{\tau}\wedge\tau} e^{-\rho t}dt + e^{-\rho\tau}\mathbf{1}_{\tau<\hat{\tau}}\right. \end{aligned}$$

(3.1

where we use that $\varphi \ge V$ on $[x-\delta, x+\delta]$ and (3.13). Next, we claim that there exists $c_0 > 0$ such that

$$\mathbb{E}\left[\int_0^{\hat{\tau}\wedge\tau} e^{-\rho t}dt + e^{-\rho\tau}\mathbf{1}_{\tau<\hat{\tau}}\right] \ge c_0, \ \tau \in \mathcal{T}_Y.$$

To this end, let $w(q) := c_0\left(1 - \frac{1}{\delta^2}(q-x)^2\right)$, with $c_0 = \min\left((\rho + \frac{1}{16}\frac{(h-l)^2}{\sigma^2\delta^2})^{-1}, 1\right)$. Then,

$$\rho w(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 w''(q) \le 1, \quad w(q) \le c_0 \le 1, \quad q \in (x-\delta, x+\delta) \tag{3.15}$$

and

$$w(x-\delta)=0,\ w(x+\delta)=0 \ \text{ and } \ w(x)=c_0>0. \tag{3.16}$$

Applying Itô's formula to $e^{-\rho t}w(\bar{q}_t)$ gives

$$\begin{aligned} 0<c_0=w(x) &= \mathbb{E}\left[\int_0^{\hat{\tau}\wedge\tau} e^{-\rho t}\left(\rho w(\bar{q}_t)-\frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 \bar{q}_t^2(1-\bar{q}_t)^2 w''(\bar{q}_t)\right)dt + e^{-\rho(\hat{\tau}\wedge\tau)}w(\bar{q}_{\hat{\tau}\wedge\tau})\right] \\ &\leq \mathbb{E}\left[\int_0^{\hat{\tau}\wedge\tau} e^{-\rho t}dt + e^{-\rho\tau}\mathbf{1}_{\tau<\hat{\tau}}\right],\ \tau\in\mathfrak{T}_Y, \end{aligned} \tag{3.17}$$

the last inequality holds from (3.15) and (3.16). Plugging the last inequality into (3.14), and taking the supremum over $\tau \in \mathfrak{T}_Y$, we get a contradiction with (3.12) to the DDP (3.22). The supersolution property follows easily and the boundary conditions were justified earlier. For the uniqueness, we refer the reader to (Crandall et al., 1992, Theorem 3.3). □

3.3.1 The Exploration and the Product-selection Regions

We introduce the sets

$$\mathcal{S} := \{q\in[0,1] \mid V(q)=G(q)\} \quad \text{and} \quad \mathcal{E} := \{q\in[0,1] \mid V(q)>G(q)\} \tag{3.18}$$

with

$$G(q)=\max(\mu, qh+(1-q)l). \tag{3.19}$$

We will refer to $\mathcal{S}$ as the *product-selection*, or *stopping*, *region* since it is therein optimal to immediately stop and choose one of the products. Its complement $\mathcal{E}$ is the *continuation region*, as it is optimal in this region to keep exploring, acquiring information about the new product. We explore the structure of $\mathcal{S}$ and $\mathcal{E}$, and investigate their dependence on the model parameters.

The analysis is carried out using appropriate sub- and super-viscosity solutions. Thus an analytic interpretation of the continuation region $\mathcal{E}$ will be justified in Proposition 7. Several properties of the solution V to the HJB equation (3.9) in this case will be discussed in Theorem 6.

Proposition 6. *The regions $\mathcal{S}$ (stopping) and $\mathcal{E}$ (exploration) are non-empty.*

Proof. i) The region $\mathcal{S} \neq \emptyset$. We show that there exists a continuous viscosity supersolution v, such that $v = G$ in $[0, \varepsilon] \cup [1-\varepsilon, 1]$ for sufficiently small $\varepsilon > 0$. To this end, for some $M > 0$ and $\varepsilon > 0$ to be determined, let

$$v(q) := \mu + M \left(\frac{q}{\varepsilon} - 1\right)_+^3, \quad 0 \leq q \leq 2\varepsilon,$$

$$v(q) := \mu + M + \frac{q - 2\varepsilon}{1 - 4\varepsilon}(h(1-2\varepsilon) + 2l\varepsilon - \mu), \quad 2\varepsilon \leq q \leq 1 - 2\varepsilon,$$

$$v(q) := qh + (1-q)l + M\left(\frac{1-q}{\varepsilon} - 1\right)_+^3, \quad 1 - 2\varepsilon \leq q \leq 1.$$

Note that, by construction, v is continuous and twice differentiable at $q = \varepsilon$ and $q = 1 - \varepsilon$. Furthermore,

$$v'(q) = 3\frac{M}{\varepsilon}\left(\frac{q}{\varepsilon} - 1\right)_+^2 \quad \text{and} \quad v''(q) = 6\frac{M}{\varepsilon^2}\left(\frac{q}{\varepsilon} - 1\right)_+, \quad 0 < q < 2\varepsilon,$$

$$v'(q) := \frac{1}{1 - 4\varepsilon}(h(1-2\varepsilon) + 2l\varepsilon - \mu) \quad \text{and } v''(q) = 0, \ 2\varepsilon < q < 1 - 2\varepsilon,$$

$$v'(q) := h - l - 3\frac{M}{\varepsilon}\left(\frac{1-q}{\varepsilon} - 1\right)_+^2 \quad \text{and} \quad 6\frac{M}{\varepsilon^2}\left(\frac{1-q}{\varepsilon} - 1\right)_+, \ 1 - 2\varepsilon < q < 1.$$

Then, for $q \in (2\varepsilon, 1 - 2\varepsilon)$,

$$\rho v(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 v''(q) + C_I(q) = \rho v(q) + C_I(q) > 0,$$

and, for $q \in [0, 2\varepsilon) \cup (1 - 2\varepsilon, 1]$,

$$\rho v(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 v''(q) + C_I(q) \geq \rho\mu - 6\frac{M}{2\varepsilon^2}\left(\frac{h-l}{\sigma}\right)^2 (2\varepsilon)^2 \geq \rho\mu - 12M\left(\frac{h-l}{\sigma}\right)^2.$$

Setting $M := \frac{\sigma^2 \rho\mu}{24(h-l)^2}$, then the right hand side of above is positive.

Next, note that $v'_-(2\varepsilon) = 3M\varepsilon^{-1}$ and $v'_+(1-2\varepsilon) = -3M\varepsilon^{-1} + h - l$. Therefore, by choosing $\varepsilon < \min\left(\frac{3M}{2(\mu-l)}, \frac{3M}{2(h-\mu)}, \frac{1}{8}\right)$, it holds that

$$v'_-(2\varepsilon) > \left(\mu + M + \frac{q-2\varepsilon}{1-4\varepsilon}(h(1-2\varepsilon) + 2l\varepsilon - \mu)\right)'_+\Bigg|_{q=2\varepsilon}$$

$$= \left(\mu + M + \frac{q-2\varepsilon}{1-4\varepsilon}(h(1-2\varepsilon) + 2l\varepsilon - \mu)\right)'_-\Bigg|_{q=1-2\varepsilon}$$

$$= \frac{h(1-2\varepsilon) + 2l\varepsilon - \mu}{1-4\varepsilon} > v'_+(1-2\varepsilon).$$

Thus, any $\mathcal{C}_2$ test function φ can only touch v from below at some q_0 in $[0,2\varepsilon)\cup(2\varepsilon,1-2\varepsilon)\cup(1-2\varepsilon,1]$, on which v is $\mathcal{C}^2$ and $v''(q_0) \geq \varphi''(q_0)$. Therefore, v is a supersolution with $v = G$ in $[0,\varepsilon] \cup [1-\varepsilon,1]$. As a consequence, we have, by uniqueness, that the value function $G \leq V \leq v$, and hence $V = G$, $q \in [0,\varepsilon] \cup [1-\varepsilon,1]$, and we conclude.

ii) The region $\mathcal{E} \neq \emptyset$. We show that exists a continuous viscosity subsolution u, such that $u(\hat{p}) > G(\hat{p})$ with $\hat{p} := \frac{\mu-l}{h-l} \in (0,1)$. To this end, for some $M > 0$ and $\varepsilon > 0$ to be chosen in the sequel, let

$$w(q) := \mu + M\left(-\frac{\varepsilon}{2} + \frac{1}{\varepsilon}(q-(\hat{p}-\varepsilon))^2\right), \quad \hat{p}-\varepsilon \leq q \leq \hat{p}+\varepsilon,$$

with

$$\varepsilon < \min\left(\frac{\hat{p}}{2}, \frac{1}{32(15M + \overline{C_I} + \rho\mu)}\left(\frac{h-l}{\sigma}\right)^2 \hat{p}^2(1-\hat{p})^2 M\right) \quad \text{and} \quad M < \frac{2}{7}(h-l),$$

with $\overline{C_I} = \max_{|q-\hat{p}|\leq\varepsilon} C_I(q)$. We claim that w is a viscosity subsolution in $[\hat{p}-\varepsilon, \hat{p}+\varepsilon]$. Indeed, $w''(q) = \frac{2M}{\varepsilon}$, and

$$\rho w(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 w''(q) + C_I(q)$$
$$\leq \rho(\mu + 15\varepsilon M) + \overline{C_I} - \frac{M}{2\varepsilon}\left(\frac{h-l}{\sigma}\right)^2 (\hat{p}-\varepsilon)^2(1-\hat{p}-\varepsilon)^2 < 0,$$

where the last inequality holds from the choice of ε. Next, let

$$u(q) := \begin{cases} \max\left(w(q), G(q)\right), & \hat{p} - \varepsilon \leq q \leq \hat{p} + \varepsilon, \\ G(q), & \text{otherwise}. \end{cases}$$

Then, u is continuous on $[0, 1]$ since

$$\begin{aligned} u(\hat{p} - \varepsilon) &= \mu - \frac{1}{2}M\varepsilon < \mu = G(\hat{p} - \varepsilon), \\ u(\hat{p} + \varepsilon) &= \mu + \frac{7}{2}M\varepsilon < \mu + (h - l)\varepsilon = G(\hat{p} + \varepsilon). \end{aligned}$$

Therefore, u is a continuous subsolution, and moreover, $u(\hat{p}) = w(\hat{p}) = \mu + \frac{1}{2}M\varepsilon > \mu = G(\hat{p})$. Therefore, by comparison, $V(\hat{p}) \geq u(\hat{p}) > G(\hat{p})$, and we conclude.

□

Proposition 7. *The value function* V *is* $\mathcal{C}^1\left(\partial\mathcal{E}\right)$ *and the unique* $\mathcal{C}^2\left(int\mathcal{E}\right)$ *solution to*

$$\rho V = \frac{1}{2}\left(\frac{h - l}{\sigma}\right)^2 q^2(1 - q)^2 V'' - C_I\left(q\right), \quad q \in int\mathcal{E}. \tag{3.20}$$

Proof. From **Proposition 6**, we have that, for some $\varepsilon > 0$, $\mathcal{E} \subset (\varepsilon, 1 - \varepsilon)$ and, thus, the above equation is uniformly elliptic in $\mathcal{E}$. Classical results (see e.g., Fleming and Rishel (2012); more recent references (Lian et al., 2020, Theorem 2.6) and (Tang et al., 2022, Lemma 5)) yield the existence and uniqueness of a smooth $\mathcal{C}^2$ solution of (3.20), say w, in any open set $\mathcal{O} \subset \mathcal{E}$, with boundary condition $w = V$. On the other hand, the value function V is the unique viscosity solution and, therefore, $w \equiv V$.

To show that $V \in \mathcal{C}^1(\partial\mathcal{E})$ we argue by contradiction. To this end, for $\hat{q} \in \partial\mathcal{E}$, we have $V(\hat{q}) = G(\hat{q})$, and $V(p) \geq G(p)$, $p \in [0, 1]$. In addition,

$$\frac{V(p) - V(\hat{q})}{p - \hat{q}} \leq \frac{G(p) - G(\hat{q})}{p - \hat{q}}, \; p < \hat{q} \quad \text{and} \quad \frac{V(p) - V(\hat{q})}{p - \hat{q}} \geq \frac{G(p) - G(\hat{q})}{p - \hat{q}}, \; p > \hat{q}.$$

Therefore, $V'_-(\hat{q}) \leq G'(\hat{q}) \leq V'_+(\hat{q})$. If V is not differentiable at $\hat{q}$, there exists some $d \in (V'_-(\hat{q}), V'_+(\hat{q}))$. Let,

$$\varphi_\varepsilon(p) := V(\hat{q}) + d(p - \hat{q}) + \frac{1}{2\varepsilon}(p - \hat{q})^2.$$

Then, V dominates φ_ε locally in a neighborhood of $\hat{q}$, i.e., $\hat{q}$ is a local minimum of $V - \varphi_\varepsilon$. From the viscosity supersolution property of V, we deduce that

$$\rho V(\hat{q}) - \frac{1}{2\varepsilon}\frac{(h-l)^2}{2\sigma^2}\hat{q}^2(1-\hat{q})^2 + C_I(\hat{q}) \geq 0.$$

Sending $\varepsilon \longrightarrow 0$ provides contradiction since $\hat{q} \in (0,1)$, and both V and C_I are Lipsitcz continuous on $[0,1]$. □

Theorem 6. *The following assertions hold:*

(I) *There exist cutoffs* $\underline{q}$ *and* $\overline{q}$*, with* $0 < \underline{q} < \overline{q} < 1$ *such that*

$$V(q) = G(q),\ \ q \leq \underline{q} \text{ and } q \geq \overline{q}, \quad \text{and } V(q) > G(q),\ \ q \in [\underline{q}, \overline{q}]. \tag{3.21}$$

(II) V *is convex and non-decreasing on* $[0,1]$.

Proof. i) Recall that $V(0) = \mu$. We claim that, if there exists $\underline{q} > 0$ such that the $V(\underline{q}) = \mu$, then it must be

$$V(q) = \mu, \quad q \in [0, \underline{q}].$$

We argue by contradiction, assuming there exists p such that $V(p) := \sup_{p\in[0,\underline{q}]} V(p) > \mu$. From the Dynamic Programming Principle, we have for $s \in \mathfrak{T}_Y$,

$$V(p) = \sup_{\tau\in\mathfrak{T}_Y} \mathbb{E}\left[-\int_0^{\tau\wedge s} e^{-\rho t}C_I(q_t^p)dt + e^{-\rho\tau}G(q_\tau^p)1_{\tau<s} + e^{-\rho s}V(q_s^p)1_{s\leq\tau}\right]. \tag{3.22}$$

Then, for $s := \inf\left(t \ \geq 0 |\ q_t^p = 0 \text{ or } q_t^p = \underline{q}\right)$,

$$V(p) = \sup_{\tau\in\mathfrak{T}_Y} \mathbb{E}\left[-\int_0^{\tau\wedge s} e^{-\rho t}C_I(q_t^p)dt + e^{-\rho\tau}G(q_\tau^p)1_{\tau<s} + e^{-\rho s}\mu 1_{s\leq\tau}\right],$$

which holds since $V(q_s^p)1_{s\leq\tau} = \mu 1_{s\leq\tau}$. Conditionally on the event $\{\tau < s\}$, we have (almost surely), that $0 \leq q_\tau^p \leq \underline{q} \leq \frac{\mu-l}{h-l}$ and, hence $G(q_\tau^p) = \mu$. This, however, yields a contradiction, since

$$\begin{aligned} V(p) &= \sup_{\tau\in\mathcal{T}_Y} \mathbb{E}\left[-\int_0^{\tau\wedge s} e^{-\rho t} C_I(q_t^p)dt + e^{-\rho\tau}\mu 1_{\tau<s} + e^{-\rho s}\mu 1_{s\leq\tau}\right] \\ &= \sup_{\tau\in\mathcal{T}_Y} \mathbb{E}\left[-\int_0^{\tau\wedge s} e^{-\rho t} C_I(q_t^p)dt + e^{-\rho\tau\wedge s}\mu\right] \leq \mu, \end{aligned}$$

where the last inequality holds since $C_I(\cdot) \geq 0$.

Similarly, we can show that if there exists $\overline{q} \in (0,1)$ such that the $V(\overline{q}) = \overline{q}h + (1-\overline{q})l$, then we must have

$$V(q) = qh + (1-q)l, \quad q \in [\overline{q}, 1].$$

Using the above results we deduce for the continuation region that $\mathcal{E} =(\underline{q}, \overline{q})$.

ii) By Proposition 7, we have that V satisfies

$$\rho V(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V''(q) + C_I(q) = 0 \quad \text{in} \quad \mathcal{E}.$$

Since $V(q) \geq G(q) > 0$ and $C_I(q) \geq 0$, $q \in [0,1]$, the above gives $V''(q) > 0$, $q \in \mathcal{E}$, and thus V is convex in $\mathcal{E}$. Furthermore, V is constant on $[0, \underline{q}]$ and linear on $[\overline{q}, 1]$. Therefore, to establish the convexity on $[0,1]$, it suffices to show that V is convex at both $\underline{q}$ and $\overline{q}$. To this end, for any (a, b, λ) such that $\underline{q} = a\lambda + b(1-\lambda)$ with $a < \underline{q} < b$ and $\lambda \in (0,1)$, we have

$$\lambda V(a) + (1-\lambda)V(b) \geq \mu = V(\underline{q}),$$

since $V(b) \geq \mu$. Similarly, for any (a, b, λ) such that $\overline{q} = a\lambda + b(1-\lambda)$ with $a < \overline{q} < b$ and $\lambda \in (0,1)$, we deduce

$$\lambda V(a) + (1-\lambda)V(b) = \lambda V(a) + (1-\lambda)\left(V(\overline{q}) + (h-l)(b-\overline{q})\right)$$

$$= \lambda V(a) + (1-\lambda)\left(V(\overline{q}) + \lambda(h-l)(b-a)\right) > V(\bar{q}),$$

where the last inequality holds since $V(a) > ha + l(1-a)$.

The monotonicity follows easily as $V'(q) \geq 0$, on $\mathcal{E}$, and V is non-decreasing on $\mathcal{S}$. □

Corollary 3. *Let $\underline{q}$ and $\overline{q}$ be the cutoff points in Theorem 6. Then, the value function V is the unique $\mathcal{C}^2\left(\left(\underline{q}, \overline{q}\right)\right)$ solution of*

$$\begin{cases} \rho V(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V''(q) + C_I(q) = 0, \\ V(\underline{q}) = \mu, V(\overline{q}) = h\overline{q} + l(1-\overline{q}), \\ V'(\underline{q}) = 0, V'(\overline{q}) = h - l, \end{cases} \tag{3.23}$$

in the class of Lipschitz continuous functions.

We introduce the notation

$$\mathcal{S}_1 := [0, \underline{q}] \quad \text{and} \quad \mathcal{S}_2 := [\overline{q}, 1]. \tag{3.24}$$

Discussion: **Theorem 6** implies that the DM will choose Product A if the initial belief $q \in \mathcal{S}_1$ and product B if $q \in \mathcal{S}_2$. If, on the other hand, $q \in \left(\underline{q}, \overline{q}\right)$ the DM starts learning about the unknown product B and makes a decision when the belief process hits either of the cut-off points, $\underline{q}$ or $\overline{q}$.

We will be calling $[0, \underline{q}]$ the safe choice region, $[\underline{q}, \overline{q}]$ the exploration region, and $[\overline{q}, 1]$ the new choice region. In general, calculating $\underline{q}$ and $\overline{q}$ in closed form is not possible, even for simple cases for the information cost C_I and the payoff function G. One of the main contributions herein is that, despite this lack of tractability, we are still able to study the behavior of the solution and the various regions for general information costs.

Remark: When $\sigma = 0$, the DM can immediately observe the true value of Θ as soon as they have access to the signal process Y, which degenerates to $dY_t = \Theta dt$. If the DM starts with belief $q(0_-) = q$ at time $t = 0_-$ and has access to Y at time $t = 0$, then the belief follows the càdlàg process

$$q_0 = \begin{cases} 1, & \text{if } \Theta = h, \\ 0, & \text{if } \Theta = l, \end{cases} \qquad \text{and } q_t = q_0, \ t \geq 0. \tag{3.25}$$

Given the possible discontinuity of the belief process at time $t = 0$, the corresponding value function is now defined as

$$V(q) = \sup_{\tau \in \mathcal{T}_Y} \mathbb{E}\left[-\int_0^\tau e^{-\rho t} C_I(q_t) dt + e^{-\rho\tau} G(q_\tau) \middle| q(0_-) = q \right]. \tag{3.26}$$

We argue that, in this case, the optimal stopping time $\tau^* = 0$. Indeed, from (3.25) it holds almost surely

$$\begin{aligned} e^{-\rho t} C_I(q_t) &= e^{-\rho t} C_I(q_0) \leq C_I(q_0), \ t \geq 0; \\ e^{-\rho t} G(q_t) &= e^{-\rho t} G(q_0) \leq G(q_0), \ t \geq 0. \end{aligned} \tag{3.27}$$

In turn, (3.26) yields

$$V(q) = \mathbb{E}\left[G(q_0) | q(0_-) = q \right] = qG(1) + (1-q)G(0) = qh + (1-q)\mu.$$

3.3.2 Sensitivity Analysis for General Information Costs

We analyze the effects of C_I, σ and ρ on the regions $\mathcal{S}_1, \mathcal{S}_2$ and $\mathcal{E}$ for arbitrary information cost functions and payoff functions of form (3.19). The solution approach depends on building appropriate sub- and super- viscosity solutions. We note that the results in Proposition 8 will be also used for the general case of reversible decisions, developed in Section 3.4.

The following result will be used repeatedly.

Lemma 4. *Let* $V_1, V_2 \in \mathcal{C}([0,1])$ *satisfying, respectively,*

$$V_i(q) = G(q) \text{ for } q \in [0, \underline{q}_i] \cup [\overline{q}_i, 1], \quad \text{and } V_i(q) > G(q), \; q \in (\underline{q}_i, \overline{q}_i), \; i = 1, 2.$$

If $V_1(q) \geq V_2(q)$, $q \in [0,1]$, *then it must be that*

$$\underline{q}_1 \leq \underline{q}_2, \overline{q}_2 \leq \overline{q}_1.$$

Proof. For every $q \in (\underline{q}_2, \overline{q}_2)$, it holds that $V_1(q) \geq V_2(q) > G(q)$, and, thus, $q \in (\underline{q}_1, \overline{q}_1)$. Therefore, $(\underline{q}_2, \overline{q}_2) \subseteq (\underline{q}_1, \overline{q}_1)$. □

Proposition 8. *The following assertions hold:*

i) If $\rho_1 \leq \rho_2$, *then* $\underline{q}_1 \leq \underline{q}_2$ *while* $\overline{q}_1 \geq \overline{q}_2$, *and, thus* $\mathcal{E}_1 \supseteq \mathcal{E}_2$.

ii) If $\sigma_1 \leq \sigma_2$, *then* $\underline{q}_1 \leq \underline{q}_2$ *while* $\overline{q}_1 \geq \overline{q}_2$, *and, thus,* $\mathcal{E}_1 \supseteq \mathcal{E}_2$.

iii) If $C_I^1(q) \geq C_I^2(q)$, $q \in [0,1]$, *then, then* $\underline{q}_1 \leq \underline{q}_2$ *while* $\overline{q}_1 \geq \overline{q}_2$ *and, thus,* $\mathcal{E}_1 \supseteq \mathcal{E}_2$.

For the reader's convenience, the proof is provided in Appendix C.1.

Discussion: When the discount rate ρ or the information cost C_I increases, the width, $\underline{q} - \overline{q}$, of the exploration region $\mathcal{E}$ decreases. Hence, the DM tends to spend less effort in learning and processing information about product B.

If the volatility σ is too large to extract useful information, the DM will spend less effort learning and prefers to choose a product faster. On the contrary, the DM will spend more effort in learning when σ is low, as the signal contains more precise information about product B.

In addition to the monotonicity properties in Proposition 8, we also study the limiting cases $\rho, C_I, \sigma \longrightarrow +\infty$. We recall the critical value

$$\hat{p} := \frac{\mu - l}{h - l}.$$

Proposition 9. *Let* $G(q) = \max(\mu, qh + (1-q)l)$. *Then, the following assertions hold:*

i) If $\rho \to +\infty$, *then* $\underline{q} \to \hat{p}$ *and* $\overline{q} \to \hat{p}$.

ii) If $C_I \to +\infty$, *then* $\underline{q} \to \hat{p}$ *and* $\overline{q} \to \hat{p}$.

iii) If $\sigma \to +\infty$, *then* $\underline{q} \to \hat{p}$ *and* $\overline{q} \to \hat{p}$.

The proof is provided in Appendix C.1.

Proposition 9 implies that $\hat{p}$ is a critical value that $\underline{q}$ and $\overline{q}$ converges to when $\rho, C_I, \sigma \to \infty$. The DM will choose either of them based on her initial belief without any exploration.

Proposition 10. *Let* $G(q) = \max(\mu, qh + (1-q)l)$. *If* $\mu_1 \leq \mu_2$, *then* $\underline{q}_1 \leq \underline{q}_2$ *and* $\overline{q}_1 \leq \overline{q}_2$.

The proof is provided in Appendix C.1.

Discussion: When the reward μ of product A is higher than the expected reward of B, the DM is more reluctant to leave the safe choice region $\mathcal{S}_1$ (since $\underline{q}$ increases) versus choosing B (since $\overline{q}$ increases).

Proposition 11. *Let* $G(q) = \max(\mu, qh + (1-q)l)$. *Then, the following assertions hold:*

i) When $l \uparrow \mu$, *then* $\overline{q} \downarrow 0$ *and* $\underline{q} \downarrow 0$.

ii) When $h \uparrow \infty$, *then* $\overline{q} \uparrow 1$ *and* $\underline{q} \downarrow 0$.

For the proof, see Appendix C.1. We note that the above results are by no means trivial as the variables l and h appear in both the "volatility" term $\frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2$ and the obstacle term in (3.9).

Monotonicity of the cutoff points with respect to parameters h and l does not, in general, hold. At the moment, we can only show that $\underline{q}$ decreases when h increases and that $\overline{q}$ decreases when l increases.

3.4 Learning with Reversible Decisions for Product B

This section introduces a framework that enables the DM to reverse her initial decision if she chooses the less-known Product B as the first decision. Specifically, if the DM explores the value of Θ at a cost in the first period and subsequently decides to make a conservative decision by selecting the well-known Product A as the initial choice, the DM will stop observing any information about Product B and stick with Product A. However, if the DM decides to try out Product B after the preliminary exploration, the DM will have a more accurate observation process $\widetilde{Y}_t$ for Product B when start using it, and there will be no additional cost associated with this new observation, since they are already using Product B. In this scenario, the DM will have one opportunity to reverse her initial decision and switch to Product A at a fixed cost ($R(\widetilde{q}) = c_S > 0$).

Mathematically, the value function for the DM is defined as:

$$V(q) = \sup_{\tau \in \mathfrak{T}_Y} \mathbb{E}\Bigg[-\int_0^{\tau} e^{-\rho t} C_I(q_t) dt \tag{3.28}$$

$$+e^{-\rho\tau} \max\Bigg(\mu, \underbrace{\sup_{\widetilde{\tau} \in \mathfrak{T}_{\widetilde{Y}}} \widetilde{\mathbb{E}}\Bigg[\int_{\tau}^{\tau+\widetilde{\tau}} e^{-\rho t}\Big(\rho(\widetilde{q}_t h + (1-\widetilde{q}_t)l)\Big) dt + e^{-\rho(\tau+\widetilde{\tau})}\big(\mu - c_S\big) \Bigg| \widetilde{q}_0 = q_\tau \Bigg]}_{(V_B)} \Bigg) \Bigg| q_0 = q \Bigg].$$

where

$$\widetilde{q}_t = \mathbb{P}\left[\Theta = h \middle| \mathcal{F}_t^{\widetilde{Y}}\right],$$

with $\widetilde{q}_0 = q_\tau$ and $\mathcal{F}_t^{\widetilde{Y}} = \sigma(\widetilde{Y}_s, s \leq t)$ is the filtration generated by $\widetilde{Y}$. The admissible control set for stopping time $\widetilde{\tau}$ is defined $\mathcal{T}_{\widetilde{Y}} = \{\widetilde{\tau} \geq 0 \mid \widetilde{\tau} \in \mathcal{F}^{\widetilde{Y}}\}$. In the running reward function of V_B, $\rho(\widetilde{q}_t h + (1-\widetilde{q}_t)l) = \rho\,\mathbb{E}[\Theta|\mathcal{F}_t^{\widetilde{Y}}]$ is the per-unit-time reward that the DM thinks she could collect under the current belief $\widetilde{q}_t$. To make the problem non-trivial, we assume $\mu - c_S > l$. See a demonstration of the decision timeline in Figure 3.2.

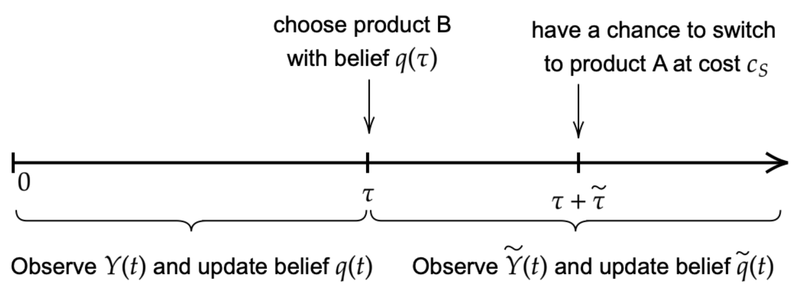


Figure 3.2: Timeline of the decisions when the initial choice is Product B.

Comparing to the formulation in Section 3.2.1, for product B we replace the per-unit-time reward $\mathbb{E}[\Theta|\mathcal{F}_\tau^Y] = q_\tau h + (1 - q_\tau)l$ by $\mathbb{E}[\Theta|\mathcal{F}_t^{\widetilde{Y}}]$ at time $t \geq \tau$. This is because, under the "irreversible decison" setup in Section 2, the DM will have no access to any additional signals after the stopping time τ. Hence her understanding of the per-unit-time reward is frozen at the status of q_τ. On the other hand for many practical applications, users' understanding of the less known product could be improved once they receive the product and their decisions may change if switching products is allowed at a certain cost. This perspective is captured in the formulation (3.28). Here $\widetilde{Y}_t$ is a new signal process that models the DM's understanding of the new product B after receiving it. This $\widetilde{Y}_t$ process may be very different from the initial signal process Y_t.

Here we examine two possibilities of the new signal process in the second period if the DM picks Product B as the initial choice – (1) $\widetilde{Y}$ follows a Poisson signal reviewing the true value of Θ at an arrival rate λ and (2) $\widetilde{Y}$ follows a Gaussian signal with a smaller variance. The Poisson signal corresponds to a scenario in which the DM will randomly discover the true value of the new product B. A higher arrival rate suggests that the DM will acquire this information relatively soon after receiving the product. On the other hand, a Gaussian signal with a smaller variance implies that the DM can collect information about Product B more efficiently, as she can start using it. This is different from the initial period, during which the DM must collect noisy information at a cost without access to any of the products.

In the next two subsections, we derive an explicit form of the function V_B

$$V_B(q_\tau) = \sup_{\widetilde{\tau}\in\mathcal{T}_{\widetilde{Y}}} \widetilde{\mathbb{E}}\left[\int_\tau^{\tau+\widetilde{\tau}} e^{-\rho t}\Big(\rho\left(\widetilde{q}_t h + (1-\widetilde{q}_t)l\right)\Big)dt + e^{-\rho(\tau+\widetilde{\tau})}\left(\mu - c_S\right)\middle|\widetilde{q}_0 = q_\tau\right], \quad (3.29)$$

under the above-mentioned two cases – a Poisson signal and a Gaussian signal with smaller variance. Then we use the explicit form of V_B to characterize the value function V defined in (3.28).

3.4.1 Case 1: Poisson Signal

If the DM starts with product B, then the decision-making problem after the first stopping time τ follows (3.29), subjecting to the following dynamics (3.30) with initial value $\widetilde{q}$ Keller and Rady (2010); Hörner and Skrzypacz (2017),

$$d\widetilde{q}_t = (1-\widetilde{q}_t)dJ_t^1(\lambda\widetilde{q}_t) + (0-\widetilde{q}_t)dJ_t^0(\lambda(1-\widetilde{q}_t)), \quad (3.30)$$

in which $\langle J_t^i(\cdot)\rangle$ are independent Poisson counting processes with intensity rate $(\cdot)$. Here λ is the rate under which the true value (precise information) of Θ arrives. Under the agent's current belief at time t, this Poisson process is splitted into two independent Poisson processes $\langle J_t^1(\cdot)\rangle$ (with information $\Theta = h$) and $\langle J_t^0(\cdot)\rangle$ (with information $\Theta = l$) at rates $\lambda\widetilde{q}_t$ and $\lambda(1-\widetilde{q}_t)$, respectively. Mathematically,

$$\begin{aligned} \lambda(1-\widetilde{q}_t)\mathrm{d}t &= \mathbb{P}\left(J^0_{t+\mathrm{d}t} - J^0_t = 1 \,\middle|\, \mathcal{F}^{\widetilde{Y}}_t\right), \\ \lambda\widetilde{q}_t\mathrm{d}t &= \mathbb{P}\left(J^1_{t+\mathrm{d}t} - J^1_t = 1 \,\middle|\, \mathcal{F}^{\widetilde{Y}}_t\right), \end{aligned}$$

with the initialization $J_0^0 = J_0^1 = 0$. Note that the belief process will jump immediately to one when the Poisson signal arrives with the true information that $\Theta = h$. Similarly, the belief process will jump immediately to zero when the Poisson signal arrives with the true information that $\Theta = l$.

The corresponding HJB equation follows:

$$\min\left(\rho U_B^\lambda - \rho\left(\widetilde{q}h + (1-\widetilde{q})l\right) - \lambda\left(\widetilde{q}h + (1-\widetilde{q})(\mu - c_S) - U_B^\lambda\right), \right. \\ \left. U_B^\lambda - \left(\mu - c_S\right)\right) = 0. \tag{3.31}$$

Note that the term $\lambda\left(\widetilde{q}h + (1-\widetilde{q})(\mu - c_S) - U_B^\lambda(\widetilde{q})\right)$ in (3.31) is a simplification from the following derivation:

$$\begin{aligned} & \lambda\widetilde{q}\left(V_B^\lambda(1) - U_B^\lambda(\widetilde{q})\right) + \lambda(1-\widetilde{q})\left(U_B^\lambda(0) - U_B^\lambda(\widetilde{q})\right) \\ = \; & \lambda\widetilde{q}\left(h - U_B^\lambda(\widetilde{q})\right) + \lambda(1-\widetilde{q})\left((\mu - c_S) - U_B^\lambda(\widetilde{q})\right) = \lambda(\widetilde{q}h + (1-\widetilde{q})(\mu - c_S) - U_B^\lambda(\widetilde{q})). \end{aligned}$$

Now we provide the explicit solution of the control problem (3.29) under dynamics (3.30).

Theorem 7. *The optimal boundary for problem* (3.29) *(under dynamics* (3.30)*) is*

$$q_B^\lambda \;:=\; \frac{(\mu - c_S - l)\rho}{\lambda(c_S - \mu + h) + \rho(h - l)}, \tag{3.32}$$

and the corresponding optimal value function U_B^λ takes the form:

$$U_B^\lambda(\widetilde{q}) = \begin{cases} \mu - c_S, & \widetilde{q} \le q_B^\lambda, \\ h\widetilde{q} + \frac{\rho l + \lambda(\mu - c_S)}{\rho + \lambda}(1 - \widetilde{q}), & \widetilde{q} > q_B^\lambda. \end{cases} \tag{3.33}$$

Note that the new "lower value" defined as

$$\widetilde{l} := \frac{\rho l + \lambda(\mu - c_S)}{\rho + \lambda} \tag{3.34}$$

can be viewed as a convex combination of $(\mu - c_S)$ and l.c It is easy to check that $\widetilde{l} > l$.

Proof. The first term of (3.31) can be re-arranged as

$$(\lambda + \rho)V_B^\lambda - (\rho + \lambda)h\widetilde{q} - (\rho + l)(1 - \widetilde{q}).$$

Hence we have V_B^λ taking the maximum of two linear pieces: $\frac{(\rho+\lambda)h\widetilde{q} + (\rho + l)(1-\widetilde{q})}{\lambda+\rho}$ and $\mu - c_S$. The intersection point q_B^λ satisfies:

$$\frac{(\rho + \lambda)q_B^\lambda h + (1 - q_B^\lambda)(\rho + l)}{\lambda + \rho} = \mu - c_S, \tag{3.35}$$

which leads to the solution in (3.32) and hence (3.33) holds. □

Corollary 4. *The following facts hold:*

(I) *We have*

$$q_B^\lambda < \hat{p} = \frac{\mu - l}{h - l}. \tag{3.36}$$

(II) $q_B^\lambda \to 0$ *when* $\lambda \to \infty$ *and* $q_B^\lambda \to \frac{\mu - c_S - l}{h - l}$ *when* $\lambda \to 0$.

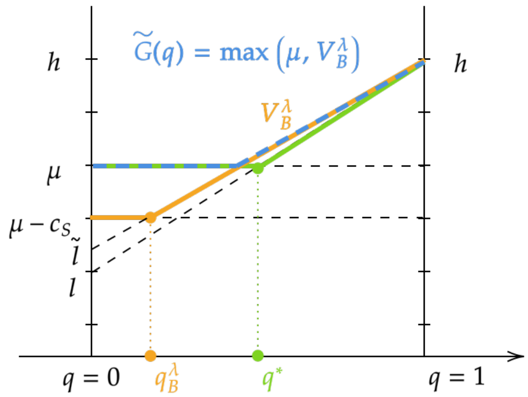


Figure 3.3: U_B^λ defined in (3.33) (in orange) v.s. $G(q)$ defined in (3.19) (in green).

This suggests that (3.28) reduces to

$$V(q) = \sup_{\tau \in \mathcal{T}_Y} \mathbb{E}\left[-\int_0^\tau e^{-\rho t} C_I(q_t) dt + e^{-\rho\tau} \widetilde{G}\left(q_\tau\right) \middle| q_0 = q \right].$$

with $\widetilde{G}(q) = \max\left(\mu,\, V_B^\lambda(q)\right)$. The corresponding HJB equation takes the form

$$\min\left(\rho V(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V''(q) + C_I(q), V(q) - \widetilde{G}(q)\right) = 0. \qquad (3.37)$$

Here $\widetilde{G}(q)$ takes the same structure as the $G(q)$ function by replacing l by $\widetilde{l}$. Therefore the general results in Section 3.3 all apply to this setting and the results in Section 3.2.1 can be applied with simple modifications. However, by comparing the results to Section (3.2.1), it sheds light on users change their behaviors when there is a chance to reverse versus no chance to reverse.

To further investigate some properties of the value function and compare to the results in Section 3.2.1, we consider a simple situation where the cost function is a constant $C_I(q) = c_I > 0$. In this case, (3.37) becomes

$$\min\left(\rho V(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V''(q) + c_I, V(q) - V_B^\lambda(q)\right) = 0. \qquad (3.38)$$

Theorem 8 (Optimal Stopping Policy and Optimal Value Function). *For the optimal control with terminal condition* V_B^λ*, there exists a unique pair* $(\underline{q}, \overline{q})$ *such that the value function* $V(x) \in \mathcal{C}^1([0,1])$ *satisfies*

$$V(x) = \begin{cases} \mu, & \text{for} \quad q \le \underline{q}, \\ -\frac{c_I}{\rho} + d_1 q^{\frac{1}{2}-\frac{k}{2}}(1-q)^{\frac{1}{2}+\frac{k}{2}} + d_2 (1-q)^{\frac{1}{2}-\frac{k}{2}} q^{\frac{1}{2}+\frac{k}{2}}, & \text{for} \quad \underline{q} \le q \le \overline{q}, \\ qh + (1-q)\widetilde{l}, & \text{for} \quad q \ge \overline{q}, \end{cases} \tag{3.39}$$

where $\widetilde{l}$ *is defined in* (3.34)*,* $k := \sqrt{1 + \frac{8\rho\sigma^2}{(h-l)^2}}$ *and*

$$d_1 = \frac{\left(\mu + \frac{c_I}{\rho} - c_S\right)\left(\frac{1}{2} + \frac{k}{2} - \underline{q}\right)}{(1-\underline{q})^{\frac{1}{2}+\frac{k}{2}} (\underline{q})^{\frac{1}{2}-\frac{k}{2}} k}, \quad d_2 = \frac{h\left(\frac{1}{2} + \frac{k}{2}\right)\overline{q} - \widetilde{l}\left(\frac{1}{2} - \frac{k}{2}\right)(1-\overline{q}) + \frac{c_I}{\rho}(\overline{q} - \frac{1}{2} + \frac{k}{2})}{(1-\overline{q})^{\frac{1}{2}-\frac{k}{2}} (\overline{q})^{\frac{1}{2}+\frac{k}{2}} k} \tag{3.40}$$

The optimal strategy is

$$\tau^* = \inf\left\{t \ge 0 : V(q_t) = V_B^\lambda(q_t)\right\} = \inf\left\{t \ge 0 : q_t \le \underline{q} \quad \text{or} \quad q_t \ge \overline{q}\right\}. \tag{3.41}$$

Proof. To discuss some properties of the solution to the above HJB equation, we first consider the solution to the second order ODE

$$\frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2 (1-q)^2 V'' = \rho V + c_I. \tag{3.42}$$

The general solution to (3.42) can be written as

$$V(q) = -\frac{c_I}{\rho} + d_1 v_1(q) + d_2 v_2(q), \tag{3.43}$$

where d_1, d_2 are two free parameters to be determined, and

$$v_1(q) = q^{\frac{1}{2}-\frac{k}{2}}(1-q)^{\frac{1}{2}+\frac{k}{2}}, \quad v_2(q) = q^{\frac{1}{2}+\frac{k}{2}}(1-q)^{\frac{1}{2}-\frac{k}{2}}. \tag{3.44}$$

Note that since $V(q) \ge \max(qh + (1-q)\widetilde{l}, \mu)$, V needs to be positive on $[0,1]$. This forces $d_1, d_2 > 0$. Hence the solution (3.43) satisfies $V(0^+) = V(1^-) = +\infty$. To seek a $\mathcal{C}^1([0,1])$ solution that satisfies the boundary condition, by Theorem 6 there exist some cut-offs $0 < \underline{q} < \overline{q} < 1$, such that $V(q) = \mu$ for $q \in [0, \underline{q}]$, $V(q) = qh + (1-q)\widetilde{l}$ for $q \in [\overline{q}, 1]$, and $V(q)$ satisfies the ODE (3.43) for $q \in [\underline{q}, \overline{q}]$. Therefore, by the smooth-fit principle, $d_1, d_2, \underline{q}, \overline{q}$ need to satisfy

$$\begin{cases} -\frac{c_I}{\rho} + d_1 v_1(\underline{q}) + d_2 v_2(\underline{q}) & = \mu - c_S, \\ d_1 v_1'(\underline{q}) + d_2 v_2'(\underline{q}) & = 0, \\ -\frac{c_I}{\rho} + d_1 v_1(\overline{q}) + d_2 v_2(\overline{q}) & = \overline{q}h + (1-\overline{q})\widetilde{l}, \\ d_1 v_1'(\overline{q}) + d_2 v_2'(\overline{q}) & = h - \widetilde{l}. \end{cases} \tag{3.45}$$

To simplify the notation, denote $m = \frac{1}{2}(1-k)$. Then the general solution to the second order ODE could be written as

$$V(q) = -\frac{c_I}{\rho} + d_1 q^m (1-q)^{-m+1} + d_2 (1-q)^m q^{-m+1}. \tag{3.46}$$

where d_1 and d_2 are two free parameters to be determined. By straightforward calculation,

$$\begin{aligned} V'(q) &= d_1 m q^{m-1}(1-q)^{-m+1} + d_1(m-1)q^m(1-q)^{-m} \\ &\quad + d_2(-m)(1-q)^{m-1}q^{-m+1} + d_2(-m+1)(1-q)^m q^{-m} \\ &= d_1 q^{m-1}(1-q)^{-m}[m(1-q) + (m-1)q] + d_2(1-q)^{m-1}q^{-m}[-mq + (-m+1)(1-q)] \\ &= d_1 q^{m-1}(1-q)^{-m}(m-q) + d_2(1-q)^{m-1}q^{-m}[-m+1-q]. \end{aligned}$$

For the left cut-off point $\underline{q}$, we have

$$d_1(\underline{q})^m(1-\underline{q})^{-m+1} + d_2(1-\underline{q})^m(\underline{q})^{-m+1} = \mu + \tfrac{c_I}{\rho} - c_S \tag{3.47}$$

$$d_1(\underline{q})^{m-1}(1-\underline{q})^{-m}(m-\underline{q}) + d_2(1-\underline{q})^{m-1}(\underline{q})^{-m}[-m+1-\underline{q}] = 0 \tag{3.48}$$

From (3.48), we have

$$d_1 = -d_2 \frac{(1-\underline{q})^{m-1}(\underline{q})^{-m}[-m+1-\underline{q}]}{(\underline{q})^{m-1}(1-\underline{q})^{-m}(m-\underline{q})} = -d_2 \frac{-m+1-\underline{q}}{m-\underline{q}} \frac{(1-\underline{q})^{2m-1}}{(\underline{q})^{2m-1}}. \tag{3.49}$$

Plugging (3.49) into (3.47), we have

$$-d_2\frac{-m+1-\underline{q}}{m-\underline{q}}\frac{(1-\underline{q})^{2m-1}}{(\underline{q})^{2m-1}}(\underline{q})^m(1-\underline{q})^{-m+1}+d_2(1-\underline{q})^m(\underline{q})^{-m+1}=\mu+\frac{c_I}{\rho}-c_S. \quad (3.50)$$

By direct computation,

$$-d_2\frac{-m+1-\underline{q}}{m-\underline{q}}(1-\underline{q})^m q^{-m+1}+d_2(1-\underline{q})^m(\underline{q})^{-m+1}=\mu+\frac{c_I}{\rho}. \quad (3.51)$$

Hence $d_2\left(1-\frac{-m+1-\underline{q}}{m-\underline{q}}\right)(1-\underline{q})^m(\underline{q})^{-m+1}=\mu+\frac{c_I}{\rho}-c_S$, and finally

$$d_2=\left(\mu+\frac{c_I}{\rho}-c_S\right)\frac{m-\underline{q}}{2m-1}(1-\underline{q})^{-m}(\underline{q})^{m-1}. \quad (3.52)$$

Plugging (3.52) into (3.49), we have

$$d_1=-\left(\mu+\frac{c_I}{\rho}-c_S\right)\frac{-m+1-\underline{q}}{2m-1}(1-\underline{q})^{m-1}(\underline{q})^{-m}. \quad (3.53)$$

Similarly, for the right cut-off point $\overline{q}$, we have

$$d_1(\overline{q})^m(1-\overline{q})^{-m+1}+d_2(1-\overline{q})^m(\overline{q})^{-m+1}=\overline{q}h+(1-\overline{q})\tilde{l}+\tfrac{c_I}{\rho} \quad (3.54)$$

$$d_1(\overline{q})^{m-1}(1-\overline{q})^{-m}(m-\overline{q})+d_2(1-\overline{q})^{m-1}(\overline{q})^{-m}[-m+1-\overline{q}]=h-\tilde{l} \quad (3.55)$$

Multiplying both sides of (3.55) by $(1-\overline{q})\overline{q}$, we have

$$d_1(\overline{q})^m(1-\overline{q})^{-m+1}(m-\overline{q})+d_2(1-\overline{q})^m(\overline{q})^{-m+1}[-m+1-\overline{q}]=(h-\tilde{l})(1-\overline{q})\overline{q}. \quad (3.56)$$

Multiplying both sides of (3.54) by $(m-\overline{q})$, we have

$$d_1(\overline{q})^m(1-\overline{q})^{-m+1}(m-\overline{q})+d_2(1-\overline{q})^m(\overline{q})^{-m+1}(m-\overline{q})=\left(\overline{q}h+(1-\overline{q})\tilde{l}+\frac{c_I}{\rho}\right)(m-\overline{q}) \quad (3.57)$$

Take the difference between (3.56) and (3.57), we have

$$\begin{aligned} d_2(1-\overline{q})^m(\overline{q})^{-m+1}[-2m+1] &= (h-l)(1-\overline{q})\overline{q}-\left(\overline{q}h+(1-\overline{q})\tilde{l}+\frac{c_I}{\rho}\right)(m-\overline{q}) \\ &= h\overline{q}(1-m)+\tilde{l}m(\overline{q}-1)+\frac{c_I}{\rho}(\overline{q}-m) \\ &= \overline{q}(h-m(h-\tilde{l})+\frac{c_I}{\rho})-m(\tilde{l}+\frac{c_I}{\rho}). \end{aligned}$$

Therefore

$$d_2 = \frac{h\overline{q}(1-m) + \widetilde{l}m(\overline{q}-1) + \frac{c_I}{\rho}(\overline{q}-m)}{(1-\overline{q})^m(\overline{q})^{-m+1}(1-2m)}. \tag{3.58}$$

Plugging (3.58) into (3.54), we have

$$\begin{aligned} d_1 &= \frac{\overline{q}h + (1-\overline{q})\widetilde{l} + \frac{c_I}{\rho} - \frac{h\overline{q}(1-m)+\widetilde{l}m(\overline{q}-1)+\frac{c_I}{\rho}(\overline{q}-m)}{(1-2m)}}{(\overline{q})^m(1-\overline{q})^{-m+1}} \\ &= \frac{\frac{c_I}{\rho}(1-m-\overline{q}) + (1-\overline{q})\widetilde{l}(1-m) - m\overline{q}h}{(1-2m)(\overline{q})^m(1-\overline{q})^{-m+1}} \end{aligned} \tag{3.59}$$

□

We investigate how the cost of switching c_S affects the exploration behavior of the DM before she makes her initial choices as follows.

Proposition 12 (Monotonicity with respect to c_S)**.** *If $c_S^1 \le c_S^2$, then $\underline{q}_1 \le \underline{q}_2$ and $\overline{q}_1 \le \overline{q}_2$*

Proposition 12 implies that, when the switching cost c_S is higher, the DM is more reluctant to leave the safe choice region (since $\underline{q}$ increases) and take the risk to choose the new Product (since $\overline{q}$ increases).

Proof of Proposition 12. Recall that in (3.38) , $\widetilde{G}(q)$ takes the same structure of $G(q)$ by replacing l by $\widetilde{l}$. Hence the monotonicity analysis with respect to c_S is equivalent to the monotonicity with respect to $\widetilde{l} = \frac{\rho l + \lambda(\mu - c_S)}{\rho+\lambda}$. For $\widetilde{l}_2 > \widetilde{l}_1$, define

$$\widetilde{G}_1(q) = \max\{\mu, qh + (1-q)\widetilde{l}_1\}, \qquad \widetilde{G}_2(q) = \max\{\mu, qh + (1-q)\widetilde{l}_2\}.$$

Then $\frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}\widetilde{G}_1 + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h \geq \widetilde{G}_2 \geq \widetilde{G}_1$. Let V_1, V_2 be the viscosity solutions to

$$\min\left\{\rho V_1(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V_1''(q) + c_I, V_1(q) - \widetilde{G}_1(q)\right\} = 0, \tag{3.60}$$

$$\min\left\{\rho V_2(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V_2''(q) + c_I, V_2(q) - \widetilde{G}_2(q)\right\} = 0. \tag{3.61}$$

We first show V_1 is a viscosity subsolution to (3.61). Take $x \in (0,1)$ and a test function $\varphi \in C^2(0,1)$ such that

$$(V_1 - \varphi)(x) = \max_{q\in(0,1)} (V_1 - \varphi)(q) = 0.$$

Since V_1 is the viscosity solution to (3.60), it is also a viscosity subsolution to (3.60), then the test function φ satisfies

$$\min\left\{\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 x^2(1-x)^2\varphi''(x) + c_I, \varphi(x) - \widetilde{G}_1(x)\right\} \leq 0.$$

Since $\widetilde{G}_1(x) \leq \widetilde{G}_2(x)$, we also have

$$\min\left\{\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 x^2(1-x)^2\varphi''(x) + c_I, \varphi(x) - \widetilde{G}_2(x)\right\} \leq 0,$$

which implies that V_1 is a viscosity subsolution to (3.61). By comparison principle, $V_1 \leq V_2$.

Next, we want to show $\frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}V_1 + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h$ is a viscosity supersolution to (3.61). We take $y \in (0,1)$ and a test function $\psi \in C^2(0,1)$, such that

$$\left(\frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}V_1 + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h - \psi\right)(y) = \min_{q\in(0,1)}\left(\frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}V_1 + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h - \psi\right)(q) = 0.$$

Then $\hat{\psi} := \frac{h-\widetilde{l}_1}{h-\widetilde{l}_2}\left(\psi - \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h\right)$ satsifies $\hat{\psi} \in C^2(0,1)$ and

$$(V_1 - \hat{\psi})(y) = \min_{q\in(0,1)} (V_1 - \hat{\psi})(q) = 0.$$

Since V_1 is the viscosity solution to (3.60), it is also a viscosity supersolution to (3.60), then the test function $\hat{\psi}$ satisfies

$$\min\left\{\rho\hat{\psi}(y) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 y^2(1-y)^2\hat{\psi}''(y) + c_I, \hat{\psi}(y) - \widetilde{G}_1(y)\right\} \geq 0.$$

Since $\hat{\psi} \leq \psi$, $\hat{\psi}'' = \frac{h-\widetilde{l}_1}{h-\widetilde{l}_2}\psi'' \geq \psi''$, and $\hat{\psi} - \widetilde{G}_1 = \frac{h-\widetilde{l}_1}{h-\widetilde{l}_2}\psi - \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_2}h - \widetilde{G}_1 \leq \frac{h-\widetilde{l}_1}{h-\widetilde{l}_2}(\psi - \widetilde{G}_2)$, we have

$$\min\left\{\rho\psi(y) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 y^2(1-y)^2\psi''(y) + c_I, \psi(y) - \widetilde{G}_2(y)\right\} \geq 0.$$

It shows that $\frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}V_1 + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h$ is a viscosity supersolution to (3.61). By comparison principle, $\frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}V_1 + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h \geq V_2$.

So far we have shown that

$$\frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}V_1(q) + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h \geq V_2(q) \geq V_1(q) \qquad \text{for } q \in [0,1].$$

By Theorem 6, there exist two pairs of cut-offs $\underline{q}_1, \overline{q}_1, \underline{q}_2, \overline{q}_2$ such that

$$\begin{cases} V_1(q) = \mu & q \in [0, \underline{q}_1] \\ V_1(q) > \widetilde{G}_1(q) & q \in (\underline{q}_1, \overline{q}_1) \\ V_1(q) = qh + (1-q)\widetilde{l}_1 & q \in [\overline{q}_1, 1] \end{cases}, \qquad \begin{cases} V_2(q) = \mu & q \in [0, \underline{q}_2] \\ V_2(q) > \widetilde{G}_2(q) & q \in (\underline{q}_2, \overline{q}_2) \\ V_2(q) = qh + (1-q)\widetilde{l}_2 & q \in [\overline{q}_2, 1] \end{cases}.$$

To compare $\underline{q}_1, \underline{q}_2$, notice that $V_2(q) \geq V_1(q) \geq \mu$. For any $q \in [0,1]$ such that $V_2(q) = \mu$, we also have $V_1(q) = \mu$, hence

$$[0, \underline{q}_2] = \{q \in [0,1] : V_2(q) = \mu\} \subseteq \{q \in [0,1] : V_1(q) = \mu\} = [0, \underline{q}_1],$$

which yields $\underline{q}_2 \leq \underline{q}_1$. To compare $\overline{q}_1, \overline{q}_2$, notice that

$$\frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}V_1 + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h \geq \frac{h-\widetilde{l}_2}{h-\widetilde{l}_1}(hq + \widetilde{l}_1(1-q)) + \frac{\widetilde{l}_2-\widetilde{l}_1}{h-\widetilde{l}_1}h = hq + \widetilde{l}_2(1-q).$$

For any $q \in [0,1]$ such that $V_1(q) = hq + \widetilde{l}_1(1-q)$, we also have $V_2(q) = hq + \widetilde{l}_2(1-q)$, hence

$$[\overline{q}_1, 1] = \{q \in [0,1] : V_1(q) = hq + \widetilde{l}_1(1-q)\} \subseteq \{q \in [0,1] : V_2(q) = hq + \widetilde{l}_2(1-q)\} = [\overline{q}_2, 1],$$

which yields $\overline{q}_2 \leq \overline{q}_1$.

In conclusion, for $\widetilde{l}_2 > \widetilde{l}_1$, i.e. $c_S^2 < c_S^1$, we have $\underline{q}_2 \leq \underline{q}_1$ and $\overline{q}_2 \leq \overline{q}_1$. □

Comparison to Irreversible Decisions. Recall that in (3.38) , $\widetilde{G}(q)$ takes the same structure as the $G(q)$ function by replacing l by $\widetilde{l}$. Hence the solution to the Poisson problem can be treated as an irreversible decision-making problem with a bigger "lower value" $\widetilde{l}$.

Also for the reversible decision-making problem with Poisson signals, $\widetilde{l}$ decreases when c_S increases, with a limiting value $\widetilde{l} = l$ achieved when $c_S = \mu - l$. By rewriting $\widetilde{l} = -\frac{\rho}{\rho+\lambda}(\mu - c_S - l) + (\mu - c_S)$, we know that $\widetilde{l}$ increases when λ increases.

Proposition 11-11 suggests that, asymptotically, the bandwidth of the exploration region $[\underline{q}, \overline{q}]$ for the reversible decision-making problem converges to zero when $\lambda \to \infty$ and $c_S \to 0$. This implies that when the Poisson signal arrives fast enough and when the cost of switching is low, the DM will spend less time acquiring costly information in the first period and make her initial choice comparatively sooner than the irreversible case.

3.4.2 Case 2: Gaussian Signal with Smaller Variance

If the DM starts with product B, the new observation process $\widetilde{Y}$ is modeled as a signal process

$$d\widetilde{Y}_t = \Theta\, dt + \widetilde{\sigma} dW_t,$$

in which $\widetilde{\sigma} \leq \sigma < 0$. Then the belief process follows the dynamics (3.62),

$$d\widetilde{q}_t = \frac{h-l}{\widetilde{\sigma}}\widetilde{q}_t\Big(1-\widetilde{q}_t\Big)dZ_t, \tag{3.62}$$

with initial value $\widetilde{q}$. Since when she chooses Product B as her initial choice, she has a better information source to know the Product B (e.g. her own using experience during the period).

The corresponding HJB equation follows:

$$\min\left(\rho V_B - \rho\Big(\widetilde{q}h + (1-\widetilde{q})l\Big) - \frac{1}{2}\left(\frac{h-l}{\widetilde{\sigma}}\right)^2 \widetilde{q}^{\ 2}(1-\widetilde{q})^2 V_B'', V_B - \Big(\mu - c_S\Big)\right) = 0. \tag{3.63}$$

Now we provide the explicit solution of the control problem under dynamics (3.62).

Theorem 9. *The optimal boundary for problem* (3.29) *(under dynamics* (3.62)*) is*

$$\widetilde{q}_B \quad := \quad \frac{\Big(l-\mu+c_S\Big)\Big(\frac{1}{2}-\frac{\widetilde{k}}{2}\Big)}{(h-l)\Big(\frac{1}{2}+\frac{\widetilde{k}}{2}\Big)+l-\mu+c_S}, \tag{3.64}$$

and the corresponding optimal value function V_B takes the form:

$$V_B(\widetilde{q}) = \begin{cases} \mu - c_S, & \widetilde{q} \leq \widetilde{q}_B, \\ \widetilde{q}h + (1-\widetilde{q})l + d_B\widetilde{q}^{\frac{1}{2}-\frac{\widetilde{k}}{2}}(1-\widetilde{q})^{\frac{1}{2}+\frac{\widetilde{k}}{2}}, & \widetilde{q} > \widetilde{q}_B, \end{cases} \tag{3.65}$$

with $d_B = \frac{(\mu-c_S)-\widetilde{q}_B h-(1-\widetilde{q}_B)l}{(\widetilde{q}_B)^{\frac{1}{2}-\frac{\widetilde{k}}{2}}(1-\widetilde{q}_B)^{\frac{1}{2}+\frac{\widetilde{k}}{2}}}$*, and* $\widetilde{k} := \sqrt{1+\frac{8\rho\widetilde{\sigma}^2}{(h-l)^2}} > 1$.

Proof. First we consider the solution to the second order ODE in (3.63):

$$\rho U_B - \rho\Big(\widetilde{q}h + (1-\widetilde{q})l\Big) - \frac{1}{2}\left(\frac{h-l}{\widetilde{\sigma}}\right)^2 \widetilde{q}^{\ 2}(1-\widetilde{q})^2 U_B'' = 0.$$

It is easy to see that the solution of the ODE can be written as

$$U_B(\widetilde{q}) = \widetilde{q}h + (1-\widetilde{q})l + c_1 v_1(\widetilde{q}) + c_2 v_2(\widetilde{q}), \tag{3.66}$$

with

$$v_1(\widetilde{q}) = \widetilde{q}^{\frac{1}{2}-\frac{\widetilde{k}}{2}}(1-\widetilde{q})^{\frac{1}{2}+\frac{\widetilde{k}}{2}}, \quad \text{and} \quad v_2(\widetilde{q}) = \widetilde{q}^{\frac{1}{2}+\frac{\widetilde{k}}{2}}(1-\widetilde{q})^{\frac{1}{2}-\frac{\widetilde{k}}{2}}.$$

Here c_1 and c_2 are two free parameters to be determined and $\widetilde{k} = \sqrt{1+\frac{8\rho\widetilde{\sigma}^2}{(h-l)^2}}$.

Similarly logic follows the proof of Theorem 7, we seek for $V_B \in \mathcal{C}^1([0,1])$ and $\widetilde{q}_B \in (0,1)$ such that $V_B(q) = \mu - c_S$ when $q \leq \widetilde{q}_B$ and $V_B(\widetilde{q}) = U_B(\widetilde{q})$ when $\widetilde{q} > \widetilde{q}_B$. Since V_B is bounded near $\widetilde{q} = 1$, we conclude that $c_2 = 0$. Otherwise $\lim_{\widetilde{q}\to 1} |V_B(\widetilde{q})| = \infty$ since $\lim_{\widetilde{q}\to 1} v_2(\widetilde{q}) = \infty$.

We next apply the smooth-fit principle to c_1 and $\widetilde{q}_B$. To get a $\mathcal{C}^1$ solution, set:

$$V_B(\widetilde{q}_B) = \widetilde{q}_B h + (1-\widetilde{q}_B)l + c_1(\widetilde{q}_B)^{\frac{1}{2}-\frac{\widetilde{k}}{2}}(1-\widetilde{q}_B)^{\frac{1}{2}+\frac{\widetilde{k}}{2}} = \mu - c_S, \tag{3.67}$$

$$V_B'(\widetilde{q}_B) = h - l + c_1(\widetilde{q}_B)^{-\frac{1}{2}-\frac{\widetilde{k}}{2}}(1-\widetilde{q}_B)^{-\frac{1}{2}+\frac{\widetilde{k}}{2}}\left(\frac{1}{2}-\frac{\widetilde{k}}{2}-\widetilde{q}_B\right) = 0, \tag{3.68}$$

which leads to the following solution:

$$\widetilde{q}_B = \frac{\Big(l-(\mu-c_S)\Big)\left(\frac{1}{2}-\frac{\widetilde{k}}{2}\right)}{(h-l)\left(\frac{1}{2}+\frac{\widetilde{k}}{2}\right)+l-(\mu-c_S)}, \tag{3.69}$$

$$d_B = \frac{(\mu-c_S)-\widetilde{q}_B h-(1-\widetilde{q}_B)l}{(\widetilde{q}_B)^{\frac{1}{2}-\frac{\widetilde{k}}{2}}(1-\widetilde{q}_B)^{\frac{1}{2}+\frac{\widetilde{k}}{2}}}. \tag{3.70}$$

□

We now provide some properties of the terminal condition.

Lemma 5. *When $c_S > 0$, there exists a unique $q_{AB} \in (0,1)$ such that $V_B(q_{AB}) = \mu$.*

Proof. Since $V_B(\widetilde{q}_B) = \mu - c_S < \mu$ and $V_B(1) = h > \mu$, by continuity of V_B there exists at least a $\widetilde{q} \in (\widetilde{q}_B, 1)$ such that $V_B(\widetilde{q}) = \mu$.

Moreover, by the formulation of V_B, we can show that for $\widetilde{q} \in (\widetilde{q}_B, 1)$,

$$V_B''(\widetilde{q}) = \frac{k^2 - 1}{4}((1 - \widetilde{q})\widetilde{q})^{\frac{-3-k}{2}}(d_B(1 - \widetilde{q})^k) > 0. \tag{3.71}$$

So $V_B(\widetilde{q})$ is a convex function for $\widetilde{q} \in (\widetilde{q}_B, 1)$.

Suppose there exist at least two intersections, such that $V_B(q_1) = V_B(q_2) = \mu$, where $\widetilde{q}_B < q_1 < q_2 < 1$. Then by the convexity of V_B, $V_B(q_1) \leq \frac{q_2 - q_1}{q_2 - \widetilde{q}_B} V_B(\widetilde{q}_B) + \frac{q_1 - \widetilde{q}_B}{q_2 - \widetilde{q}_B} V_B(q_2) < \mu$, which is a contradiction to $V_B(q_1) = \mu$. Therefore, the intersection is unique. We denote q_{AB} to be the unique intersection. See Figure 3.4 for a demonstration.

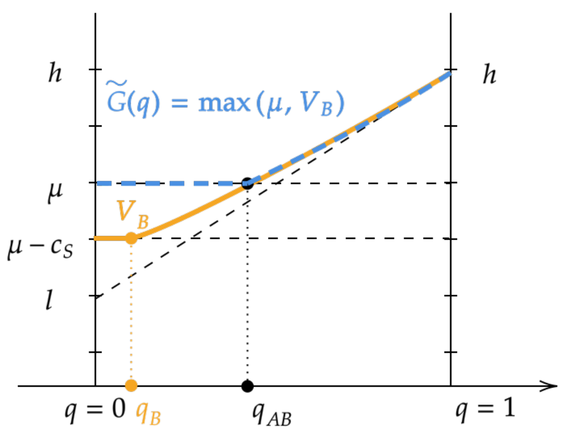


Figure 3.4: Demonstration of the terminal condition: $\widetilde{G}$ (in dotted blue), and V_B (in orange).

□

This suggests that (3.28) reduces to

$$V(q) = \sup_{\tau \in \mathcal{T}_Y} \mathbb{E}\left[-\int_0^\tau e^{-\rho t} C_I(q_t)dt + e^{-\rho\tau} \max(\mu, V_B(q_\tau)) \,\middle|\, q_0 = q \right].$$

To further investigate some properties of the value function and the optimal stopping strategy, we consider a simple situation where the cost function is a constant $C_I(q) = c_I > 0$. The corresponding HJB equation takes the form

$$\min\left(\rho V(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V''(q) + c_I, V(q) - \widetilde{G}(q)\right) = 0. \tag{3.72}$$

where

$$\widetilde{G}(q) = \max\left(\mu, V_B(q)\right) \tag{3.73}$$

The optimal stopping strategy and the corresponding optimal value function are provided as follows.

Theorem 10 (Optimal Stopping Policy and Optimal Value Function)**.** *For the optimal control with terminal condition* (3.73)*, there exists a unique pair* $(\underline{q}, \overline{q})$ *such that the value function* $V(q) \in \mathcal{C}^1$ *satisfies*

$$V(q) = \begin{cases} \mu - c_S, & \text{for} \quad q \le \underline{q}, \\ -\frac{c_I}{\rho} + d_1 q^{\frac{1}{2}-\frac{k}{2}}(1-q)^{\frac{1}{2}+\frac{k}{2}} + d_2(1-q)^{\frac{1}{2}-\frac{k}{2}} q^{\frac{1}{2}+\frac{k}{2}}, & \text{for} \quad \underline{q} \le q \le \overline{q}, \\ hq + l(1-q) + d_B q^{\frac{1}{2}-\frac{\widetilde{k}}{2}}(1-q)^{\frac{1}{2}+\frac{\widetilde{k}}{2}}, & \text{for} \quad q \ge \overline{q}, \end{cases} \tag{3.74}$$

where

$$d_1 = \frac{\left(\mu + \frac{c_I}{\rho} - c_S\right)\left(\frac{1}{2} + \frac{k}{2} - \underline{q}\right)}{(1-\underline{q})^{\frac{1}{2}+\frac{k}{2}}(\underline{q})^{\frac{1}{2}-\frac{k}{2}} k}, \quad d_2 = -\frac{\left(\mu + \frac{c_I}{\rho} - c_S\right)\left(\frac{1}{2} - \frac{k}{2} - \underline{q}\right)}{(1-\underline{q})^{\frac{1}{2}-\frac{k}{2}}(\underline{q})^{\frac{1}{2}+\frac{k}{2}} k} \tag{3.75}$$

d_B *is defined in Theorem 7, where*

$$d_B = \frac{\mu - l - c_S}{\frac{1}{2} + \frac{\widetilde{k}}{2}}\left(\frac{\left(\frac{1}{2} + \frac{\widetilde{k}}{2}\right)(h - \mu + c_S)}{-\left(\frac{1}{2} - \frac{\widetilde{k}}{2}\right)(\mu - l - c_S)}\right)^{\frac{1}{2}-\frac{\widetilde{k}}{2}} \tag{3.76}$$

The optimal strategy is

$$\tau^* = \inf\left\{t \ge 0 : V(q_t) = \widetilde{G}(q_t)\right\} = \inf\left\{t \ge 0 : q_t \le \underline{q} \quad \text{or} \quad q_t \ge \overline{q}\right\}. \tag{3.77}$$

Proof. To simplify the notation, denote $m = \frac{1}{2}(1-k)$. Then the general solution to the second order ODE could be written as

$$V(q) = -\frac{c_I}{\rho} + d_1 q^m (1-q)^{-m+1} + d_2 (1-q)^m q^{-m+1}. \tag{3.78}$$

where d_1 and d_2 are two free parameters to be determined. By straightforward calculation,

$$\begin{aligned} V'(q) &= d_1 m q^{m-1}(1-q)^{-m+1} + d_1(m-1)q^m(1-q)^{-m} \\ &\quad + d_2(-m)(1-q)^{m-1}q^{-m+1} + d_2(-m+1)(1-q)^m q^{-m} \\ &= d_1 q^{m-1}(1-q)^{-m}[m(1-q)+(m-1)q] + d_2(1-q)^{m-1}q^{-m}[-mq+(-m+1)(1-q)] \\ &= d_1 q^{m-1}(1-q)^{-m}(m-q) + d_2(1-q)^{m-1}q^{-m}[-m+1-q]. \end{aligned}$$

For the left cut-off point $\underline{q}$, we have

$$d_1(\underline{q})^m(1-\underline{q})^{-m+1} + d_2(1-\underline{q})^m(\underline{q})^{-m+1} = \mu + \tfrac{c_I}{\rho} - c_S \tag{3.79}$$

$$d_1(\underline{q})^{m-1}(1-\underline{q})^{-m}(m-\underline{q}) + d_2(1-\underline{q})^{m-1}(\underline{q})^{-m}[-m+1-\underline{q}] = 0 \tag{3.80}$$

From (3.80), we have

$$d_1 = -d_2\frac{(1-\underline{q})^{m-1}(\underline{q})^{-m}[-m+1-\underline{q}]}{(\underline{q})^{m-1}(1-\underline{q})^{-m}(m-\underline{q})} = -d_2\frac{-m+1-\underline{q}}{m-\underline{q}}\frac{(1-\underline{q})^{2m-1}}{(\underline{q})^{2m-1}}. \tag{3.81}$$

Plugging (3.81) into (3.79), we have

$$-d_2\frac{-m+1-\underline{q}}{m-\underline{q}}\frac{(1-\underline{q})^{2m-1}}{(\underline{q})^{2m-1}}(\underline{q})^m(1-\underline{q})^{-m+1} + d_2(1-\underline{q})^m(\underline{q})^{-m+1} = \mu + \frac{c_I}{\rho} - c_S. \tag{3.82}$$

By direct computation,

$$-d_2\frac{-m+1-\underline{q}}{m-\underline{q}}(1-\underline{q})^m q^{-m+1} + d_2(1-\underline{q})^m(\underline{q})^{-m+1} = \mu + \frac{c_I}{\rho} - c_S. \tag{3.83}$$

Hence $d_2\left(1-\frac{-m+1-\underline{q}}{m-\underline{q}}\right)(1-\underline{q})^m(\underline{q})^{-m+1}=\mu+\frac{c_I}{\rho}-c_S$, and finally

$$d_2=\left(\mu+\frac{c_I}{\rho}-c_S\right)\frac{m-\underline{q}}{2m-1}(1-\underline{q})^{-m}(\underline{q})^{m-1}. \tag{3.84}$$

Plugging (3.84) into (3.81), we have

$$d_1=-\left(\mu+\frac{c_I}{\rho}-c_S\right)\frac{-m+1-\underline{q}}{2m-1}(1-\underline{q})^{m-1}(\underline{q})^{-m}. \tag{3.85}$$

Similarly, for the right cut-off point $\overline{q}$, we have

$$d_1(\overline{q})^m(1-\overline{q})^{-m+1}+d_2(1-\overline{q})^m(\overline{q})^{-m+1}=\overline{q}h+(1-\overline{q})l+\frac{c_I}{\rho}+d_B(\overline{q})^{\widetilde{m}}(1-\overline{q})^{-\widetilde{m}+1} \tag{3.86}$$

$$d_1(\overline{q})^{m-1}(1-\overline{q})^{-m}(m-\overline{q})+d_2(1-\overline{q})^{m-1}(\overline{q})^{-m}[-m+1-\overline{q}]=h-l+d_B(\overline{q})^{\widetilde{m}-1}(1-\overline{q})^{-\widetilde{m}}(\widetilde{m}-\overline{q}) \tag{3.87}$$

where we denote $\widetilde{m}=\frac{1}{2}(1-\widetilde{k})$. □

Comparison to Irreversible Decisions. Analytically, as $\widetilde{\sigma}\to 0$, we can show from Theorem 9 that $q_B\to 0$ and $V_B(\widetilde{q})=\widetilde{q}h+(1-\widetilde{q})(\mu-c_S)$. Thus, in this scenario, the $\widetilde{G}(q)$ function is equivalent to the $G(q)$ function in the irreversible decision-making case with the lower value of Product B replaced by $\mu-c_S$. This implies that if the DM could observe the true value of Product B once she chooses it, she would immediately finalize her decision whether to keep Product B if $\Theta=h$, or switch to Product A at a cost of c_S if $\Theta=l$.

Also note that when $\widetilde{\sigma}=0$, the same monotonicity result with respect to changes in c_S in Proposition 12 holds, i.e., as c_S increases, both $\underline{q}$ and $\overline{q}$ increase. This indicates that when the cost of revising the decision is higher, the DM is more likely to choose the well-known Product A.

We further conduct a series of numerical experiments when $0<\widetilde{\sigma}<\sigma$, as illustrated in Figure 3.5, to delve deeper into the impact of the cost of reverse (c_S) on the behavior of DM. Here we re-denote the cutoffs of the irreversible counterpart as $\underline{q}^*$ and $\overline{q}^*$. As c_S increases from zero to $\mu - l$ (with the assumption that $\mu - l = h - \mu$ without loss of generality), we observe that

- Both $\underline{q}$ and $\overline{q}$ increase monotonically, moving towards to $(\underline{q}^*, \overline{q}^*)$. This indicates that when the cost of reverse becomes more costly, the behavior of the DM tends to align more closely with "irreversible decisions."

- The width of the exploration range, $\overline{q} - \underline{q}$, also expands from zero to an upper limit below $\overline{q}^* - \underline{q}^*$. It shows that when the cost of reverse is zero, the DM tends to make her initial choice without any exploration. This situation also presents a significant advantage for selecting Product B initially (since both $\underline{q}$ and $\overline{q}$ have a small value), as the DM can always switch to Product A at no additional cost later on. As the cost of reverse increase, the DM engages in more exploration before reaching her initial decision.

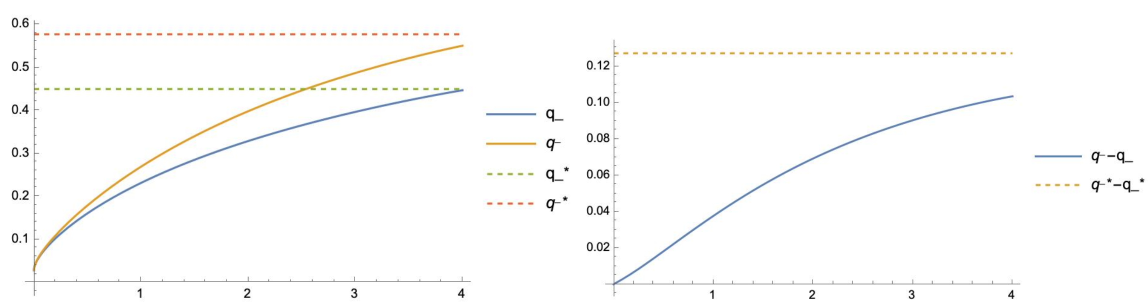


Figure 3.5: Behavior of $\underline{q}$, $\overline{q}$, and $\overline{q} - \underline{q}$. In this experiment, we take $\rho = 1$, $l = 1$, $h = 9$, $\mu = 5$, $c_I = 1$, $\sigma = 5$, $\widetilde{\sigma} = 1$.

Chapter 4: Conclusions and Future Research Directions

This dissertation contains two main parts dedicated to quantitative and modeling aspects of optimal decision-making in stochastic environments under model ambiguity and costly information acquisition, respectively.

The first part provides an offline data-driven distributionally robust stochastic optimization and its application to decision making with contextual information. In Chapter 1, 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. The proposed Shapley extension provides a novel family of policies for adjustable robust optimization with probably zero optimality gap and can be easily extended to contextual decision-making problems with general convex costs, beyond the newsvendor one. The work can be generalized in two main directions: i) investigate how the dynamic inventory management problem can be solved with a demand-feature data set and ii) extend the proposed Shapley extension for other adjustable robust optimization problems and more general classes of stochastic optimization problems. In Chapter 2, we propose a new distributionally robust decision-rule optimization for decision-making with side information based on causal transport distance. We study its computational properties by providing tractable formulations for both the inner worst-case loss problem and the outer optimization over the decision rules problem. We derive a tractable dual reformulation for evaluating the worst-case expected cost and show that the worst-case distribution has a similar conditional information structure as the nominal distribution. We also identify tractable cases to find the optimal decision rules over an affine class or the entire nonparametric class and apply our work in conditional regression, incumbent pricing, and portfolio selection. These results open new research directions for distributionally robust optimization and adjustable robust optimization. For future work, it would be interesting to further investigate the application of causal transport distance and to investigate the performance guarantees of the proposed framework.

The second part examines the interplay between costly information acquisition and optimal decision making. The main question is how to choose between two products, a known product and a new one, when the return of the latter is not perfectly known to the consumer. The analysis is done in a Bayesian framework using filtering techniques. However, the use of the filter is costly, which gives rise to a Stefan problem for optimal stopping. The novelty of the work is on allowing revision of the initial decision at a return fee, a rather realistic feature in today's retail platforms. Allowing for revision of the first decision (return the product initially bought) creates a more complex sequential optimal stopping problem. The second novelty of the model is the differential information structure after the initial buying choice. This refined information gives rise to a "nested" optimal stopping problem with various interesting features. There are various directions for future research, among others, related to the number of products, improved learning, and market design. Specifically, a consumer may have the option to choose among more than two products, which will give rise to a new multi-dimensional combined filtering and optimal stopping problem. Improved learning is related to the second period of the problem during which the characteristics of the initially bought products are, on the one hand, revealed more accurately but, on the other, may themselves change dynamically. Finally, revisions of initial decisions are directly related to the level of exchange/return fees. How to determine this fee from the retailer's perspective is a rather interesting question especially because it will force us to consider a continuum of consumers, which might give rise to new mean field games.

Appendix A: Appendices to *"Data-driven Decision Making and Distributionally Robust Stochastic Optimization"*

A.1 Proofs for Section 1.5.1

The following Lemma 6 is a direct consequence of the strong duality result in Wasserstein distributionally robust optimization (e.g., Gao and Kleywegt (2022); Esfahani and Kuhn (2018); Blanchet and Murthy (2019)). To ease the notation, in the sequel we denote $\Psi_f(x,z) := \Psi(f(x),z)$ and $\Psi_{\widehat{f}}(\widehat{x},z) = \Psi(\widehat{f}(\widehat{x}),z)$.

Lemma 6. *For each $f \in \mathcal{F}$, the inner primal problem*

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

is equal to the following inner dual problem,

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

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

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

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

In Figure A.1, we plot the graph of the function A_{12} when $K = 2$ (left) and A_{12}, A_{23}, A_{13} when $K = 3$ (right), in the case $\mathcal{X} = \mathbb{R}^2$. Same as the setting of Figure 1.3, when

$K = 2$ the Shapley policy is A_{12}, and when $K = 3$ the Shapley policy is the middle one among A_{12}, A_{13} and A_{23}, which is rendered with a mesh in this figure.

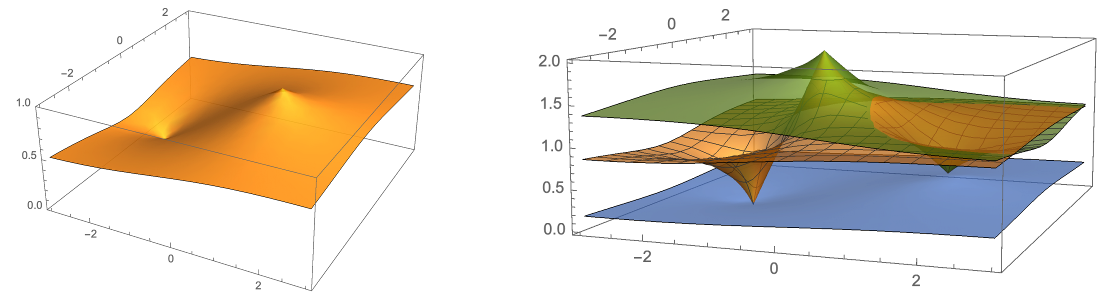


Figure A.1: Graph of the Shapley policy $y = f(x)$ when $K = 2, 3$, $x \in \mathbb{R}^2$

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

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

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

Next we show the boundedness and the Lipschitzness. It suffices to show them for each A_{jk}, because both bounds are compatible with min max operations. Because $A_{jk}(x)$ is just an interpolation between y_j and y_k, clearly we have $\min\{y_j, y_k\} \le A_{jk}(x) \le \max\{y_j, y_k\}$. As for the Lipschitz bound of A_{jk}, when $j = k$, $A_{kk} \equiv y_k$ are constant functions, so they always satisfy the Lipschitz bound. When $j \ne k$, fix $x, x' \in \mathcal{X}$, and we denote

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

Then

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

By triangular inequality,

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

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

It remains to prove (1.9), which is to show that $y = A^+(x), A^-(x)$ satisfy the following condition for every k:

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

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

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

Using definition of M_k in (Mk),

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

Plug in into the above inequality,

$$\Phi(A_{jk}(x)) - \|x-\widehat{x}_k\| \leq M_k + \frac{\|x-\widehat{x}_k\|}{\|x-\widehat{x}_j\| + \|x-\widehat{x}_k\|}\|\widehat{x}_j - \widehat{x}_k\| - \|x-\widehat{x}_k\|$$
$$= M_k + \|x-\widehat{x}_k\|\left(\frac{\|\widehat{x}_j - \widehat{x}_k\|}{\|x-\widehat{x}_j\| + \|x-\widehat{x}_k\|} - 1\right) \leq M_k.$$

In the last step we used the triangle inequality $\|\widehat{x}_j - \widehat{x}_k\| \leq \|x-\widehat{x}_j\| + \|x-\widehat{x}_k\|$. 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 in a closed interval $[A_{j_1k}(x), A_{j_2k}(x)]$ convex function Φ can only attain maximum at the endpoints due to the maximum principle, we conclude that

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

As a result, since both $y = A_{j_1k}(x)$ and $y = A_{j_2k}(x)$ satisfy (Mk), we conclude that $y = A^+(x), A^-(x)$ also satisfy (Mk), which implies (1.9).

The existence of the saddle point is proved in Lemma 7 below. It is purely algebraic and it makes use of the Shapley's theorem. □

Lemma 7. *With the same notation as Lemma 1,* $A^+ = A^-$.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$(d_i + d_k)(y_i - y_\ell) > (d_i + d_\ell)(y_i - y_k),$$

$$(d_j + d_\ell)(y_j - y_k) > (d_j + d_k)(y_j - y_\ell),$$

$$(d_i + d_\ell)(y_j - y_\ell) > (d_j + d_\ell)(y_i - y_\ell).$$

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

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

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

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

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

To show $v_D \leq v_{\hat{D}}$, we let f be the Shapley extension of a given $\widehat{f} \in \widehat{\mathcal{F}}$ as defined in Lemma 1, and split into two cases: $\lambda > 0$ and $\lambda = 0$. If $\lambda > 0$, by Lemma 1 we know that f satisfies (1.9). In particular, If we choose $\Phi(y) = \frac{1}{\lambda}\Psi(y, z) - |z - \widehat{Z}|$, then

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

Consequently,

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

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

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

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

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

The left dominates v_D, so taking the inf over $\widehat{f} \in \widehat{\mathcal{F}}$ on the right gives $v_D \le v_{\hat{D}}$, which completes the proof of $v_D = v_{\hat{D}}$. Note that the above also proves that if $\widehat{f}^*$ is a minimizer of $v_{\hat{D}}$, then the Shapley extension of $\widehat{f}^*$ is a minimizer of v_D, also a minimizer of v_P by Lemma 6. □

A.2 Proofs for Section 1.6.1

Lemma 8. *For each fixed $\lambda \ge b \vee h$ and $\widehat{f} \in \widehat{\mathcal{F}}$, we can always find a $\frac{\lambda}{b\vee h}$-Lipschitz policy $\widehat{g} \in \widehat{\mathcal{F}}$ satisfying $\underline{z}_k \le \widehat{g}(\widehat{x}_k) \le \overline{z}_k$, where*

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

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

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

Proof. Proof of Lemma 8. Denote $y_k = \widehat{f}(\widehat{x}_k)$. First, we show that if $\widehat{f}$ is "not Lipschitz enough" at some point, in the sense that its local Lipschitz constant at a point is too large, then we can reduce it by modifying $\widehat{f}$. Suppose there exists j_0, k_0 such that

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

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

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

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

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

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

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

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

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

This completes the proof of the claim.

Now we make use of the above claim recursively.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Proof. Proof of Proposition 1.

We start by proving that

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

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

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

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

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

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

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

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

Then

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

where we replace $\widehat{\mathcal{F}}$ by $\bar{\hat{\mathcal{F}}} \cap \widehat{\mathcal{F}}_\lambda$ in (A.2) using Lemma 8. For each $\widehat{f} \in \widehat{\mathcal{F}}_\lambda$, because $\|\Psi(\widehat{f}(\cdot), \widehat{z})\|_{\mathrm{Lip}} \leq \|\widehat{f}\|_{\mathrm{Lip}}(b \vee h) \leq \lambda$, the max over x is attained at $x = \widehat{X}$. Therefore,

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

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

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

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

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

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

A.3 Additional Results for Section 1.6.2 and Proofs for Section 1.6.3

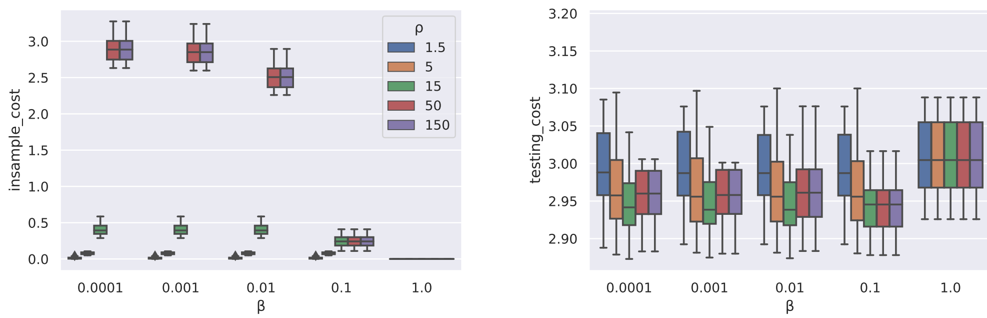


Figure A.2: Impact of hyperparameters ρ and β on the in-sample and out-of-sample performance

In Figure A.2, we illustrate the effect of different choices of hyperparameters on the model performance. We plot the in-sample costs and the out-of-sample costs in the case when $d = 5000$, $n = 100$, and $h = b = 1$, with various hyperparameters. The in-sample cost increases in the radius ρ and decreases in the distance weight β. As β becomes large ($\beta = 1$), in-sample cost reduces to zero and the policy becomes empirical risk minimization. This can be seen from the regularization point of view ($\widehat{\mathsf{R}}$): when we are indifferent in Lipschitz norms that are below a very high threshold—higher than the Lipschitz norm of the critical conditional quantile—the optimal policy is simply the unconditional quantile function. The out-of-sample performance seems less sensitive in β as long as it is sufficiently small (less than 0.1), but it has a clear preference in suitable tuning value of ρ. For large or small ρ, the policies are either too conservative or too restrictive, which lead to undesirable out-of-sample cost.

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

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

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

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

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

A.4 Proofs for Section 1.6.4

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

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

Proof. Proof of Proposition 4.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

To see the second part, if condition (I) is not satisfied, then we claim that the optimizer $\widehat{f}^*$ must have Lipschitz constant $\|\widehat{f}^*\|_{\text{Lip}} = L \geq 1$. Indeed, if $L < 1$, we can always adjust the value of $\widehat{f}^*$ to reduce costs in the second term of $v_{\hat{R}}$ without paying more cost in the first term. So, the only possibility that $\|\widehat{f}^*\|_{\text{Lip}} < 1$ is that the second term is already optimized, that is $\underline{q} \leq \widehat{f}^* \leq \overline{q}$. However, this would imply (I) holds, which is a contradiction.

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

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

Proposition 13 (Worst-case Distribution)**.** *Let $\widehat{f}^*$ be an in-sample optimal robust policy. Then there exists a worst-case distribution $\mathbb{P}^*$ of $(\widehat{\mathsf{P}})$ such that the following holds.*

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

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

In Figure A.3, we have the identical setting as in Figure 1.5. Above the graph of f^* is a blue shadow region representing $\{(x, z) : z \geq f^*(x)\}$, and $\mathbb{P}^*$ moves the probability mass in this region when backorder costs more than holding. In the left figure $\|\widehat{f}^*\|_{\text{Lip}} < 1$, so it is more cost-efficient to move along the direction of z. In the right figure $\|\widehat{f}^*\|_{\text{Lip}} > 1$, and it is more cost-efficient to move along the direction of x; since the worst-case distribution is for the in-sample problem, it perturbs x from one empirical value to another.

Proof. Proof of Proposition 13. To ease the notation we use f to represent $\widehat{f}^*$ in this proof.

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

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

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

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

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

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

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

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

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

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

This implies that for every k,

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

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

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

We will show that

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Finally, take expectation in X, we have

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

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

Remark 4. *Similarly, when $b < h$, the transport map should move $X \in \widehat{\mathcal{X}}_<$ in $\{Z \le \widehat{f}^*(X)\}$ when $L > 1$, and should decrease Z in $\{Z \le \widehat{f}^*(X)\}$ when $L \le 1$. Considering the demand must be nonnegative, in the case $L \le 1$ we should use the transport map*

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

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

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

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

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

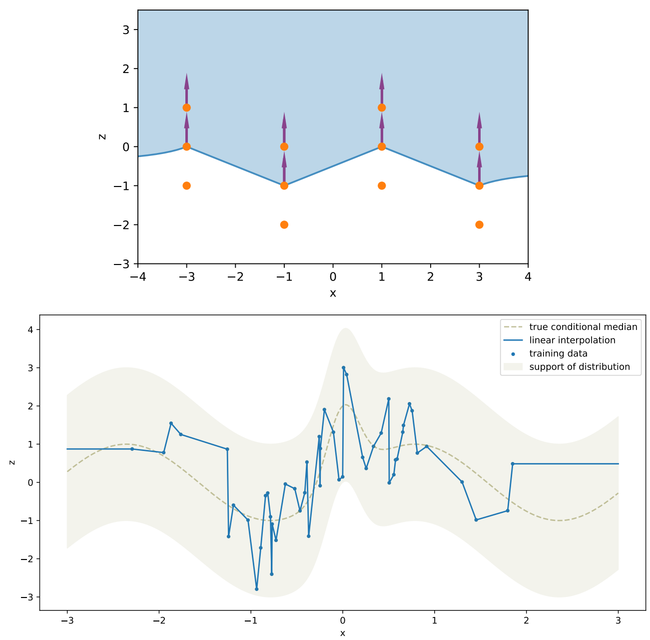


Figure A.3: Transport map of worst-case distribution (purple arrows). When the optimal policy is 1-Lipschitz, worst-case distribution moves in z (left). Otherwise, worst-case distribution moves in x (right).

Appendix B: Appendices to *"Distributionally Robust Stochastic Optimization with Causal Transport Distance"*

B.1 Causal Transport Distance

Lemma 9 (Equivalent Definition)**.** *Let $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$ be a transport plan. Then the following are equivalent.*

(I) $\gamma \in \Gamma_c(\widehat{\mathbb{P}}, \mathbb{P})$.

(II) *For $\widehat{\mathbb{P}}$-almost every $(\widehat{X}, \widehat{Z}) \in \mathcal{X} \times \mathcal{Z}$,*

$$\gamma_{X|(\widehat{X},\widehat{Z})} = \gamma_{X|\widehat{X}}.$$

(III) *Let $\mathrm{Proj}_X : \mathcal{X} \times \mathcal{Z} \to \mathcal{X}$ be the projection into X coordinate. For $\widehat{\mathbb{P}}$-almost every $(\widehat{x}, \widehat{z}_1), (\widehat{x}, \widehat{z}_2) \in \mathcal{X} \times \mathcal{Z}$,*

$$(\mathrm{Proj}_X)_{\#}\gamma(\mathrm{d}x|\widehat{x}, \widehat{z}_1) = (\mathrm{Proj}_X)_{\#}\gamma(\mathrm{d}x|\widehat{x}, \widehat{z}_2).$$

(IV) *For $\widehat{\mathbb{P}}_{\widehat{X}}$-almost every $\widehat{X}$ and $\mathbb{P}_X$-almost every X,*

$$\gamma_{\widehat{Z}|(\widehat{X},X)} = \gamma_{\widehat{Z}|\widehat{X}} = \widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}.$$

(V) *Let $\mathrm{Proj}_{\widehat{Z}} : \mathcal{Z} \times \mathcal{Z} \to \mathcal{Z}$ be the projection into $\widehat{Z}$ coordinate: $\mathrm{Proj}_{\widehat{Z}}(\widehat{z}, z) = \widehat{z}$. For $\widehat{\mathbb{P}}_{\widehat{X}}$-almost every $\widehat{x} \in \mathcal{X}$ and $\mathbb{P}_X$-almost every $x_1, x_2 \in \mathcal{X}$,*

$$(\mathrm{Proj}_{\widehat{Z}})_{\#}\gamma(\mathrm{d}\widehat{z}|\widehat{x}, x_1) = (\mathrm{Proj}_{\widehat{Z}})_{\#}\gamma(\mathrm{d}\widehat{z}|\widehat{x}, x_2).$$

Moreover, $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$ plus any one from the above is equivalent to $\gamma \in \mathcal{P}((\mathcal{X} \times \mathcal{Z}) \times (\mathcal{X} \times \mathcal{Z}))$, satisfying

(VI) *γ has a decomposition into successive regular kernels*

$$\gamma(\mathrm{d}\widehat{x}\,\mathrm{d}\widehat{z}\,\mathrm{d}x\,\mathrm{d}z) = \gamma_1(\mathrm{d}\widehat{x}\,\mathrm{d}x)\gamma_2(\mathrm{d}\widehat{z}\,\mathrm{d}z|\widehat{x},x)$$

satisfying

$$\begin{aligned} \gamma_1 &\in \Gamma(\widehat{\mathbb{P}}_{\widehat{X}}, \mathbb{P}_X), \\ (\mathrm{Proj}_{\widehat{Z}})_{\#}\gamma_2(\mathrm{d}\widehat{z}|\widehat{x},x) &= \widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}(\mathrm{d}\widehat{z}|\widehat{x}) \qquad \textit{for } \gamma_1\textit{-almost every } (\widehat{x},x), \\ (\mathrm{Proj}_{(X,Z)})_{\#}\gamma_{Z|X}(\mathrm{d}z|x) &= \mathbb{P}_{Z|X}(\mathrm{d}z|x) \qquad \textit{for } \mathbb{P}_X\textit{-almost every } x. \end{aligned}$$

That is,

$$\gamma_1 \in \Gamma(\widehat{\mathbb{P}}_{\widehat{X}}, \mathbb{P}_X), \qquad \gamma_2 \in \Gamma(\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}, \mathbb{Q}^{(\widehat{X})}) \textit{ where } \mathbb{E}_{\widehat{X}\sim(\gamma_1)_{\widehat{X}|X}}[\mathbb{Q}^{(\widehat{X})}|X] = \mathbb{P}_{Z|X}.$$

Proof. Proof. The equivalence of (I), (II) and (IV) follows from the definition. It is also easy to check from the definition that (II) is equivalent to (III), and (IV) is equivalent to (V).

Suppose (VI) holds, then projecting γ onto $(X, \widehat{X}, \widehat{Z})$ coordinate, we have

$$(\mathrm{Proj}_{(X,\widehat{X},\widehat{Z})})_{\#}\gamma(\mathrm{d}x\,\mathrm{d}\widehat{x}\,\mathrm{d}\widehat{z}) = \gamma_1(\mathrm{d}\widehat{x}\,\mathrm{d}x)\cdot(\mathrm{Proj}_{\widehat{Z}})_{\#}\gamma_2(\mathrm{d}\widehat{z}|\widehat{x},x) = \gamma_1(\mathrm{d}\widehat{x}\,\mathrm{d}x)\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}(\mathrm{d}\widehat{z}|\widehat{x}).$$

Projecting onto $(\widehat{X}, \widehat{Z})$ yields

$$(\mathrm{Proj}_{(\widehat{X},\widehat{Z})})_{\#}\gamma(\mathrm{d}\widehat{x}\,\mathrm{d}\widehat{z}) = (\mathrm{Proj}_{\widehat{X}})_{\#}\gamma_1(\mathrm{d}\widehat{x})\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}(\mathrm{d}\widehat{z}|\widehat{x}) = \widehat{\mathbb{P}}_{\widehat{X}}(\mathrm{d}\widehat{x})\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}(\mathrm{d}\widehat{z}|\widehat{x}) = \widehat{\mathbb{P}}(\widehat{x},\widehat{z}).$$

As for the other marginal,

$$(\mathrm{Proj}_{(X,Z)})_{\#}\gamma(\mathrm{d}x\,\mathrm{d}z) = (\mathrm{Proj}_X)_{\#}\gamma_1(\mathrm{d}x)\cdot(\mathrm{Proj}_{(X,Z)})_{\#}\gamma_{Z|X}(\mathrm{d}z|x) = \mathbb{P}_X(\mathrm{d}x)\mathbb{P}_{Z|X}(\mathrm{d}z|x) = \mathbb{P}(\mathrm{d}x\,\mathrm{d}z).$$

So indeed we have $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$.

Example 13. *Here we show a few examples of causal transport. A transport plan induced by a causal transport map* $\boldsymbol{T} : \mathcal{X}\times\mathcal{Z} \to \mathcal{X}\times\mathcal{Z}$ *is causal. Recall that* $\boldsymbol{T}$ *is causal if it is in the form* $\boldsymbol{T}(x, z) = (T_1(x), T_2(x, z))$*, where* $T_1 : \mathcal{X} \to \mathcal{Z}$ *and* $T_2 : \mathcal{X} \times \mathcal{Z} \to \mathcal{Z}$ *are measurable. To see why this is true, let* $\gamma = (\mathrm{Id}\times T)_\# \widehat{\mathbb{P}}$*, then take any two points* $(\widehat{x}, \widehat{z})$*, we have*

$$(\mathrm{Proj}_X)_\# \gamma(\mathrm{d}x|\widehat{x}, \widehat{z}) = \delta_{T_1(\widehat{x})}$$

which is independent of the choice of $\widehat{z}$*.*

Proof. Proof of Lemma 2. Since $\gamma^{(q)}$ are transport plans starting from $\widehat{\mathbb{P}}$,

$$\gamma^{(q)}_{(\widehat{X},\widehat{Z})} = \widehat{\mathbb{P}}, \qquad \gamma^{(q)}_{\widehat{X}} = \widehat{\mathbb{P}}_{\widehat{X}}, \qquad \forall q \in [0, 1].$$

Together with

$$\gamma^{(q)}_{(X,\widehat{X},\widehat{Z})} = (1 - q)\gamma^{(0)}_{(X,\widehat{X},\widehat{Z})} + q\gamma^{(1)}_{(X,\widehat{X},\widehat{Z})}, \qquad \gamma^{(q)}_{(X,\widehat{X})} = (1 - q)\gamma^{(0)}_{(X,\widehat{X})} + q\gamma^{(1)}_{(X,\widehat{X})},$$

we know that

$$\gamma^{(q)}_{X|(\widehat{X},\widehat{Z})} = (1 - q)\gamma^{(0)}_{X|(\widehat{X},\widehat{Z})} + q\gamma^{(1)}_{X|(\widehat{X},\widehat{Z})}, \qquad \gamma^{(q)}_{X|\widehat{X}} = (1 - q)\gamma^{(0)}_{X|\widehat{X}} + q\gamma^{(1)}_{X|\widehat{X}}.$$

Because $\gamma^{(0)}$ and $\gamma^{(1)}$ are causal, by equivalent definition (2), for $\widehat{\mathbb{P}}$-almost every $(\widehat{X}, \widehat{Z}) \in \mathcal{X} \times \mathcal{Z}$,

$$\gamma^{(0)}_{X|(\widehat{X},\widehat{Z})} = \gamma^{(0)}_{X|\widehat{X}}, \qquad \gamma^{(1)}_{X|(\widehat{X},\widehat{Z})} = \gamma^{(1)}_{X|\widehat{X}}.$$

Therefore

$$\gamma^{(q)}_{X|(\widehat{X},\widehat{Z})} = \gamma^{(q)}_{X|\widehat{X}},$$

so $\gamma^{(q)}$ is also causal.

Proof. Proof. With probability one, each $\widehat{x}$ in the support of $\widehat{\mathbb{P}}$ corresponds to only one $\widehat{z}$, so that

$$\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}=\widehat{x}_k} = \delta_{\widehat{z}_k}.$$

Now let $\gamma \in \Gamma(\widehat{\mathbb{P}}, \mathbb{P})$. Because

$$\mathbb{E}_{X|\widehat{X}}[\gamma_{\widehat{Z}|(\widehat{X},X)}] = \gamma_{\widehat{Z}|\widehat{X}} = \delta_{\widehat{Z}},$$

the only choice is $\gamma_{\widehat{Z}|(\widehat{X},X)} = \delta_{\widehat{X}}$, for $(\gamma_1)_{X|\widehat{X}}$-a.e. X. Therefore γ is causal.

B.2 Sup of Convex Functions

Lemma 10 (Dual Objective Function)**.** *The dual objective function h has the following properties. Let $c_I = \{h < \infty\}$. Then*

(I) *There exists $\kappa \geq 0$, such that either $c_I = (\kappa, \infty)$ or $c_I = [\kappa, \infty)$.*

(II) *h is convex and continuous in c_I.*

(III) *$h(\lambda) \to \infty$ as $\lambda \to \infty$.*

(IV) *h has a minimizer $\lambda^* \in [\kappa, \infty)$.*

Proof. Proof.

(I) $h(\lambda) - \lambda\rho^p$ is monotonously decreasing in λ, therefore we can find κ such that h is infinite for smaller λ, and finite for greater λ.

(II) h is a combination of supremums and expectations of convex functions, therefore h is convex. Since $h < \infty$ in c_I, h is continuous in c_I with only a possible exception at $\kappa \in c_I$. Notice that

$$\begin{aligned}
\liminf_{\lambda \downarrow \kappa} F_{(x)}(\lambda; \widehat{x}) &= \liminf_{\lambda \downarrow \kappa} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{x}}} \left[\sup_{z \in \mathcal{Z}} \left\{ G_{(z)}(\lambda; x, \widehat{Z}) \right\} \mid \widehat{X} = \widehat{x} \right] - \kappa \|x - \widehat{x}\|^p \\
&\geq \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{x}}} \left[\liminf_{\lambda \downarrow \kappa} \sup_{z \in \mathcal{Z}} \left\{ G_{(z)}(\lambda; x, \widehat{Z}) \right\} \mid \widehat{X} = \widehat{x} \right] - \kappa \|x - \widehat{x}\|^p \\
&\geq \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{x}}} \left[\sup_{z \in \mathcal{Z}} \left\{ \liminf_{\lambda \downarrow \kappa} G_{(z)}(\lambda; x, \widehat{Z}) \right\} \mid \widehat{X} = \widehat{x} \right] - \kappa \|x - \widehat{x}\|^p \\
&\geq \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{x}}} \left[\sup_{z \in \mathcal{Z}} \left\{ G_{(z)}(\kappa; x, \widehat{Z}) \right\} \mid \widehat{X} = \widehat{x} \right] - \kappa \|x - \widehat{x}\|^p = F_{(x)}(\kappa; \widehat{x}).
\end{aligned}$$

Similarly

$$\begin{aligned}
\liminf_{\lambda \downarrow \kappa} h(\lambda) &= \kappa \rho^p + \liminf_{\lambda \downarrow \kappa} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\sup_{x \in \mathcal{X}} \left\{ F_{(x)}(\lambda; \widehat{X}) \right\} \right] \\
&\geq \kappa \rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\sup_{x \in \mathcal{X}} \left\{ F_{(x)}(\kappa; \widehat{X}) \right\} \right] = h(\kappa).
\end{aligned}$$

Therefore h is continuous in c_I.

(III) This is simply because we can pick $x = \widehat{X}$, $z = \widehat{Z}$ so

$$\begin{aligned}
h(\lambda) &\geq \lambda \rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{x}}} \left[\Psi(f(\widehat{X}), \widehat{Z}) - \lambda \|\widehat{Z} - \widehat{Z}\|^p \mid \widehat{X} \right] - \lambda \|\widehat{X} - \widehat{X}\|^p \right] \\
&= \lambda \rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{x}}} \left[\Psi(f(\widehat{X}), \widehat{Z}) \mid \widehat{X} \right] \right] = \lambda \rho^p + \mathbb{E}_{\widehat{\mathbb{P}}} \left[\Psi(f(\widehat{X}), \widehat{Z}) \right] \to +\infty
\end{aligned}$$

as $\lambda \to +\infty$.

(IV) It follows from (1)-(3).

Lemma 11 (Exchange sup and derivative for Convex Functions)**.** *Let Λ be an index set. Let $\{F_\alpha\}_{\alpha \in \Lambda}$ be a family of real-valued convex functions defined on an interval c_I. Suppose its sup is pointwise bounded, $\Phi(\lambda) = \sup_{\alpha \in \Lambda} F_\alpha(\lambda) < \infty$. Denote $f_\alpha(\lambda) = F'_\alpha(\lambda)$, and $\phi(\lambda) = \Phi'(\lambda)$. For any function f we denote f^* [resp. f_*] to be the upper* *[resp. lower] semicontinous envelope of f. For every $\varepsilon > 0$, define the ε-argmax set Ω_ε and $\overline{D}, \underline{D}$ by*

$$\Omega_\varepsilon(\lambda) := \{\alpha \in \Lambda : F_\alpha(\lambda) \geq \Phi(\lambda) - \varepsilon\},$$

$$\overline{D}_\varepsilon(\lambda) := \sup_{\alpha \in \Omega_\varepsilon(\lambda)} f_\alpha^*(\lambda), \qquad \overline{D}(\lambda) = \lim_{\varepsilon \to 0} \overline{D}_\varepsilon(\lambda),$$

$$\underline{D}_\varepsilon(\lambda) := \inf_{\alpha \in \Omega_\varepsilon(\lambda)} f_{\alpha *}(\lambda), \qquad \underline{D}(\lambda) = \lim_{\varepsilon \to 0} \underline{D}_\varepsilon(\lambda).$$

Then

(I) *For every $\lambda \in c_I$, $\underline{D}(\lambda) \leq \overline{D}(\lambda)$.*

(II) *For every $\lambda, \mu \in c_I$ with $\lambda < \mu$, $\overline{D}(\lambda) \leq \phi^*(\lambda) \leq \phi_*(\mu) \leq \underline{D}(\mu)$.*

(III) *Fix $\lambda \in c_I$, $\delta > 0$, $\epsilon > 0$. If $\lambda_1 \in c_I$ such that $\lambda_1 < \lambda$ is sufficiently close to λ, then we can find $\alpha \in \Lambda$ such that*

$$f_\alpha^*(\lambda_1) \leq \phi_*(\lambda) + \delta, \qquad F_\alpha(\lambda_2) \geq \Phi(\lambda) - \epsilon.$$

If $\lambda_2 \in c_I$ such that $\lambda_2 > \lambda$ is sufficiently close to λ, we can find $\beta \in \Lambda$ such that

$$f_{\beta_*}(\lambda_2) \geq \phi^*(\lambda) - \delta, \qquad F_\beta(\lambda_2) \geq \Phi(\lambda) - \epsilon.$$

Proof. Proof. Φ is the sup of a family of convex functions, so Φ is convex. Since Φ and F_α are convex and finite in c_I, they have locally Lipschitz, monotonously increasing derivatives ϕ and f_α. Monotonicity implies f_α^* and ϕ^* [resp. $f_{\alpha *}$ and ϕ_*] are right [resp. left] continuous, and thus convexity implies for $\lambda < \mu$,

$$f_\alpha{}^*(\lambda) \leq \frac{F_\alpha(\mu) - F_\alpha(\lambda)}{\mu - \lambda} \leq f_{\alpha *}(\mu), \qquad \phi^*(\lambda) \leq \frac{\Phi(\mu) - \Phi(\lambda)}{\mu - \lambda} \leq \phi_*(\mu). \tag{B.1}$$

(I) ε-argmax set Ω_ε is never empty by definition. Therefore, $\underline{D}_\varepsilon(\lambda) \leq \overline{D}_\varepsilon(\lambda)$ holds for all ε. As $\varepsilon \to 0$, $\Omega_\varepsilon(\lambda)$ shrinks, so $\overline{D}_\varepsilon(\lambda) \downarrow \overline{D}(\lambda)$, $\underline{D}_\varepsilon(\lambda) \uparrow \underline{D}(\lambda)$, we have $\underline{D}(\lambda) \leq \overline{D}(\lambda)$.

(II) Fix any $\varepsilon > 0$, and $\lambda < \mu$. For any $\alpha \in \Omega_\varepsilon(\lambda)$, $\beta \in \Omega_\varepsilon(\mu)$, using (B.1) we have

$$F_\alpha(\mu) - \varepsilon \le \Phi(\mu) - \varepsilon \le F_\beta(\mu) \le F_\beta(\lambda) + (\mu - \lambda) f_{\beta_*}(\mu) \le \Phi(\lambda) + (\mu - \lambda) f_{\beta_*}(\mu),$$

$$F_\beta(\lambda) - \varepsilon \le \Phi(\lambda) - \varepsilon \le F_\alpha(\lambda) \le F_\alpha(\mu) - (\mu - \lambda) {f_\alpha}^*(\lambda) \le \Phi(\mu) - (\mu - \lambda) {f_\alpha}^*(\lambda).$$

By these two inequalities, we conclude

$$-\varepsilon + (\mu - \lambda) f_\alpha^*(\lambda) \le \Phi(\mu) - \Phi(\lambda) \le \varepsilon + (\mu - \lambda) f_{\beta_*}(\mu),$$
$$\Rightarrow -\frac{\varepsilon}{\mu - \lambda} + f_\alpha^*(\lambda) \le \frac{\Phi(\mu) - \Phi(\lambda)}{\mu - \lambda} \le \frac{\varepsilon}{\mu - \lambda} + f_{\beta_*}(\mu).$$

By taking the sup over $\alpha \in \Omega_\varepsilon(\lambda)$, taking the inf over $\beta \in \Omega_\varepsilon(\mu)$, we have

$$-\frac{\varepsilon}{\mu - \lambda} + \overline{D}_\varepsilon(\lambda) \le \frac{\Phi(\mu) - \Phi(\lambda)}{\mu - \lambda} \le \frac{\varepsilon}{\mu - \lambda} + \underline{D}_\varepsilon(\mu).$$

Let $\varepsilon \to 0$,

$$\overline{D}(\lambda) \le \frac{\Phi(\mu) - \Phi(\lambda)}{\mu - \lambda} \le \underline{D}(\mu). \tag{B.2}$$

We now combine (B.1) with (B.2) to show that $\phi^*(\lambda) \le \underline{D}(\mu)$, $\overline{D}(\lambda) \le \phi_*(\mu)$. To finish the proof of (2), we use the monotonicity $\phi^*(\lambda) \le \phi_*(\mu)$, and

$$\phi^*(\lambda) = \lim_{\mu \downarrow \lambda} \phi(\mu) \ge \lim_{\mu \downarrow \lambda} \phi_*(\mu) \ge \overline{D}(\lambda), \qquad \phi_*(\mu) = \lim_{\lambda \uparrow \mu} \phi(\lambda) \le \lim_{\lambda \uparrow \mu} \phi^*(\lambda) \le \underline{D}(\mu).$$

(III) Since Φ is continuous in the interior of c_I, we can let λ_1 and λ_2 be close enough to λ such that

$$\Phi(\lambda_1), \Phi(\lambda_2) \ge \Phi(\lambda) - \frac{\epsilon}{2}.$$

Let $\varepsilon < \frac{\epsilon}{2}$ be small enough such that $\overline{D}_\varepsilon(\lambda_1) < \overline{D}(\lambda_1) + \delta$, $\underline{D}_\varepsilon(\lambda_2) > \underline{D}(\lambda_2) - \delta$. Pick any $\alpha \in \Omega_\varepsilon(\lambda_1), \beta \in \Omega_\varepsilon(\lambda_2)$, then

$$f_\alpha^*(\lambda_1) \le \overline{D}_\varepsilon(\lambda_1) < \overline{D}(\lambda_1) + \delta \le \phi_*(\lambda) + \delta,$$

$$f_{\beta_*}(\lambda_2) \ge \underline{D}_\varepsilon(\lambda_2) > \underline{D}(\lambda_2) - \delta \ge \phi^*(\lambda) - \delta.$$

Moreover, by the definition of $\Omega_\varepsilon(\lambda)$,

$$F_\alpha(\lambda_1) \geq \Phi(\lambda_1) - \varepsilon \geq \Phi(\lambda_1) - \frac{\epsilon}{2} \geq \Phi(\lambda) - \epsilon,$$
$$F_\beta(\lambda_2) \geq \Phi(\lambda_2) - \varepsilon \geq \Phi(\lambda_2) - \frac{\epsilon}{2} \geq \Phi(\lambda) - \epsilon.$$

Lemma 12. *With the same notations as the previous lemma, let* Λ *be a Euclidean space. Suppose for each* $\lambda \in \mathrm{Int}(c_I)$*,* $F_\alpha(\lambda)$ *is upper semicontinuous in* α*, and* $|f_\alpha(\lambda)| \to \infty$ *as* $|\alpha| \to \infty$*. Then*

(I) $\Omega_0(\lambda)$ *is nonempty.*

(II) *There exists* $\alpha, \beta \in \Omega_0(\lambda)$*, such that*

$$f_{\alpha *}(\lambda) = \phi_*(\lambda), \qquad f^*_\beta(\lambda) = \phi^*(\lambda), \qquad F_\alpha(\lambda) = F_\beta(\lambda) = \Phi(\lambda).$$

(III) $\overline{D}_0(\lambda) = \overline{D}(\lambda) = \phi^*(\lambda)$*, and* $\underline{D}_0(\lambda) = \underline{D}(\lambda) = \phi_*(\lambda)$*.*

Proof. Proof. Let $\lambda_0 \in \mathrm{Int}(c_I)$. Then we can find $\kappa < \lambda_0 < \mu$ all inside $\mathrm{Int}(c_I)$. For some small δ, $\kappa' = \kappa - \delta$ and $\mu' = \mu + \delta$ are also inside $\mathrm{Int}(c_I)$.

(I) By Lemma 11 (2), $\phi_*(\lambda) \leq \underline{D}(\lambda) \leq \overline{D}(\lambda) \leq \phi^*(\lambda)$, and since λ is in the interior of c_I, Φ is locally Lipschitz, $\underline{D}(\lambda), \overline{D}(\lambda)$ are finite. Thus for some small ε, $\underline{D}_\varepsilon(\lambda)$ and $\overline{D}_\varepsilon(\lambda)$ are finite. This implies that Ω_ε is bounded, otherwise $|f_\alpha(\lambda)| \to \infty$ as $\alpha \to \infty$. Because F_α is upper semicontinuous, Ω_ε is also closed, so it is compact, thus

$$\Phi(\lambda) = \sup_{\alpha \in \Lambda} F_\alpha(\lambda) = \sup_{\alpha \in \Omega_\varepsilon(\lambda)} F_\alpha(\lambda)$$

is attainable, i.e.,

$$\Omega_0(\lambda) = \arg\max_{\alpha \in \Lambda} F_\alpha(\lambda)$$

is nonempty.

(II) For every λ, since $\Omega_0(\lambda) \subset \Omega_\varepsilon(\lambda)$ for any ε, we know that $\overline{D}_\varepsilon(\lambda) \geq \overline{D}_0(\lambda)$, $\underline{D}_\varepsilon(\lambda) \leq \underline{D}_0(\lambda)$. Let $\varepsilon \to 0$ we have $\overline{D}(\lambda) \geq \overline{D}_0(\lambda)$, $\underline{D}(\lambda) \leq \underline{D}_0(\lambda)$. So for every $\alpha \in \Omega_0(\lambda)$,

$$\phi_*(\lambda) \leq \underline{D}(\lambda) \leq \underline{D}_0(\lambda) \leq f_{\alpha*}(\lambda) \leq f_\alpha^*(\lambda) \leq \overline{D}_0(\lambda) \leq \overline{D}(\lambda) \leq \phi^*(\lambda). \tag{B.3}$$

Let $\lambda_n \uparrow \lambda_0$ be an increasing sequence inside $[\kappa, \mu]$. For each λ_n, $\Omega_0(\lambda_n)$ is nonempty, so we can find α_n such that

$$F_{\alpha_n}(\lambda_n) = \Phi(\lambda_n), \qquad \phi_*(\lambda_n) \leq f_{\alpha_n*}(\lambda_n) \leq f_{\alpha_n}^*(\lambda_n) \leq \phi^*(\lambda_n).$$

First, we claim that F_{α_n} are uniformly bounded in $[\kappa, \mu]$. The upper bound $F_{\alpha_n} \leq \Phi$ is clear. As for the lower bound, we first use the convexity of Φ, for all $\lambda \in [\kappa, \mu]$,

$$\Phi(\lambda) \geq \Phi(\kappa) + \phi^*(\kappa)(\lambda - \kappa), \qquad \Phi(\lambda) \geq \Phi(\mu) - \phi_*(\mu)(\mu - \lambda).$$

then we use the convexity of F_{α_n}, for $\lambda \in [\lambda_n, \mu]$,

$$\begin{aligned} F_{\alpha_n}(\lambda) &\geq F_{\alpha_n}(\lambda_n) + f_{\alpha_n}^*(\lambda_n)(\lambda - \lambda_n) \\ &\geq \Phi(\lambda_n) + \phi_*(\lambda_n)(\lambda - \lambda_n) \\ &\geq \Phi(\kappa) + \phi^*(\kappa)(\lambda_n - \kappa) + \phi^*(\kappa)(\lambda - \lambda_n) \\ &= \Phi(\kappa) + \phi^*(\kappa)(\lambda - \kappa). \end{aligned}$$

For $\lambda \in [\kappa, \lambda_n]$,

$$\begin{aligned} F_{\alpha_n}(\lambda) &\geq F_{\alpha_n}(\lambda_n) - f_{\alpha_n*}(\lambda_n)(\lambda_n - \lambda) \\ &\geq \Phi(\lambda_n) - \phi^*(\lambda_n)(\lambda_n - \lambda) \\ &\geq \Phi(\mu) - \phi_*(\mu)(\mu - \lambda_n) - \phi_*(\mu)(\lambda_n - \lambda) \\ &= \Phi(\mu) - \phi_*(\mu)(\mu - \lambda). \end{aligned} \tag{B.4}$$

Therefore, for all $\lambda \in [\kappa, \mu]$,

$$F_{\alpha_n}(\lambda) \geq \min\left\{\Phi(\kappa) + \phi^*(\kappa)(\lambda - \kappa), \Phi(\mu) - \phi_*(\mu)(\mu - \lambda)\right\}.$$

Next, we claim that F_{α_n} are equicontinuous in $[\kappa, \mu]$. Since

$$F_{\alpha_n}(\kappa) \geq \min\left\{\Phi(\kappa), \Phi(\mu) - \phi_*(\mu)(\mu - \kappa)\right\} = \Phi(\mu) - \phi_*(\mu)(\mu - \kappa),$$

by convexity of F_{α_n} we have

$$f_{\alpha_n *}(\kappa) \geq \frac{F_{\alpha_n}(\kappa) - F_{\alpha_n}(\kappa')}{\kappa - \kappa'} \geq \frac{\Phi(\mu) - \phi_*(\mu)(\mu - \kappa) - \Phi(\kappa')}{\delta}.$$

Similarly we have

$$f^*_{\alpha_n}(\mu) \leq \frac{F_{\alpha_n}(\mu') - F_{\alpha_n}(\mu)}{\mu' - \mu} \leq \frac{\Phi(\mu') - \Phi(\kappa) - \phi^*(\kappa)(\mu - \kappa)}{\delta}.$$

f_{α_n} are increasing between κ and μ, so they are uniformly bounded, thus F_{α_n} are uniformly Lipschitz.

Since f_{α_n} are uniformly bounded, we know that $\{\alpha_n\}_{n\in\mathbb{N}}$ is bounded by the assumption of the lemma. Up to a subsequence we may assume $\alpha_n \to \alpha$. Since F_{α_n} are uniformly bounded and equicontinuous in $[\kappa, \mu]$, by Arzelà-Ascoli Lemma it admits a subsequence uniformly converging to some F_∞, and since F_α is upper semicontinuous in α, we know that $F_\alpha \geq \lim_{n\to\infty} F_{\alpha_n} = F_\infty$. Therefore, up to a subsequence,

$$\Phi(\lambda_0) \geq F_\alpha(\lambda_0) \geq F_\infty(\lambda_0) = \lim_{n\to\infty} F_{\alpha_n}(\lambda_n) = \lim_{n\to\infty} \Phi(\lambda_n) = \Phi(\lambda_0).$$

Thus $\alpha \in \Omega_0(\lambda_0)$. Moreover, by taking $n \to \infty$ in (B.4), for any $\lambda \in [\kappa, \lambda_0)$ we have

$$\begin{aligned}\Phi(\lambda) \geq F_\alpha(\lambda) \geq F_\infty(\lambda) &= \lim_{n\to\infty} F_{\alpha_n}(\lambda_n) - f_{\alpha_n *}(\lambda_n)(\lambda_n - \lambda) \\ &\geq \lim_{n\to\infty} F_{\alpha_n}(\lambda_n) - \phi_*(\lambda_n)(\lambda_n - \lambda) = \Phi(\lambda_0) - \phi_*(\lambda_0)(\lambda_0 - \lambda),\end{aligned}$$

and they all equal at $\lambda = \lambda_0$. So the left derivative at λ_0

$$\phi_*(\lambda_0) \geq f_{\alpha*}(\lambda_0) \geq \phi_*(\lambda_0)$$

are equal. This shows that $f_{\alpha*}(\lambda_0) = \phi_*(\lambda_0)$. The proof for the β part is exactly symmetric to the α, so we omit here.

(III) This is the consequence of part (2) and (B.3).

B.3 Proofs for Section 2.2.1

Proof. Proof of Theorem 2.

What remains to be proved is the strong duality $v_{\mathsf{P}}^f \geq v_{\mathsf{D}}^f$. For each $x \in \mathcal{X}$, $\widehat{z} \in \mathcal{Z}$ we denote

$$G_{(z)}(\lambda; x, \widehat{z}) := \Psi(f(x), z) - \lambda \|z - \widehat{z}\|^p.$$

It is a linearly decreasing function of λ. Thus, the supremum over z

$$\Upsilon(\lambda; x, \widehat{z}) := \sup_{z \in \mathcal{Z}} \left\{ G_{(z)}(\lambda; x, \widehat{z}) \right\} \tag{B.5}$$

is a decreasing convex function of λ. Because the expectation of decreasing convex functions are decreasing and convex, we have for each $\widehat{x} \in \mathcal{X}$,

$$F_{(x)}(\lambda; \widehat{x}) := \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}} \left[\Upsilon(\lambda; x, \widehat{Z}) \mid \widehat{X} = \widehat{x} \right] - \lambda \|x - \widehat{x}\|^p$$

is a family of decreasing convex functions of λ. Their supremum

$$\Phi(\lambda; \widehat{x}) := \sup_{x \in \mathcal{X}} \left\{ F_{(x)}(\lambda; \widehat{x}) \right\} \tag{B.6}$$

is again convex and decreasing. Finally, the dual objective function

$$h(\lambda) = \lambda \rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\Phi(\lambda; \widehat{X}) \right]$$

is also convex. By Lemma 10, there exists $\kappa \in [0, \infty]$ such that h is finite in (κ, ∞) and infinite in $[0, \kappa)$. Moreover, in the case $\kappa < \infty$, h attains its global minimum at $\lambda^* \geq \kappa$. Therefore we can separate into the following cases.

Case 1: $\kappa = \infty$

This means $h(\lambda) = \infty$ for any $\lambda \geq 0$, therefore $v_{\mathsf{D}}^f = \infty$.

Now fix $\lambda > 0$, then

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\Phi(\lambda; \widehat{X})\right] = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\sup_{x\in\mathcal{X}} F_{(x)}(\lambda; \widehat{X})\right] = \infty.$$

We may assume

$$\mathbb{E}_{\widehat{\mathbb{P}}}[\Psi(f(\widehat{X}), \widehat{Z})] < \infty,$$

otherwise $v_{\mathsf{P}}^f = \infty$ because $\widehat{\mathbb{P}}$ is feasible, and the strong duality holds automatically. For each $\widehat{X}$ we can find an $X \in \mathcal{X}$, denoted by $X = T_1(\widehat{X})$, such that

$$\begin{aligned}
\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[F_{(X)}(\lambda; \widehat{X})\right] &\geq \mathbb{E}_{\widehat{\mathbb{P}}}[\Psi(f(\widehat{X}), \widehat{Z})] + 2\lambda\rho^p, \\
\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\lambda; X, \widehat{Z}) \mid \widehat{X}\right] - \lambda\|X - \widehat{X}\|^p\right] &\geq \mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}), \widehat{Z})\right] + 2\lambda\rho^p, \\
2\lambda\rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\lambda\|X - \widehat{X}\|^p\right] &\leq \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\lambda; X, \widehat{Z}) - \Psi(f(\widehat{X}), \widehat{Z}) \mid \widehat{X}\right]\right] \\
&= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}} G_{(z)}(\lambda; x, \widehat{Z}) - \Psi(f(\widehat{X}), \widehat{Z}) \mid \widehat{X}\right]\right]
\end{aligned}$$

For each $(\widehat{X}, \widehat{Z})$ pair, we can find $Z \in \mathcal{Z}$, denoted by $T_2(\widehat{X}, \widehat{Z})$, such that

$$\begin{aligned}
\lambda\rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\lambda\|X - \widehat{X}\|^p\right] &\leq \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[G_{(Z)}(\lambda; X, \widehat{Z}) - \Psi(f(\widehat{X}), \widehat{Z}) \mid \widehat{X}\right]\right] \\
&= \mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(X), Z) - \Psi(f(\widehat{X}), \widehat{Z}) - \lambda\|Z - \widehat{Z}\|^p\right]
\end{aligned}$$

Denote $\gamma_1 = ((T_1, T_2) \otimes \mathrm{id}_{\mathcal{X}\times\mathcal{Z}})_\sharp \widehat{\mathbb{P}}$, with $\sharp$ means push-forward of measure. Then $((X, Z), (\widehat{X}, \widehat{Z})) \sim \gamma_1$, and denote the distance between $(\widehat{X}, \widehat{Z})$ and (X, Z) by

$$D = \mathbb{E}_{\gamma_1}\left[\|X - \widehat{X}\|^p + \|Z - \widehat{Z}\|^p\right],$$

then

$$\mathbb{E}_{\gamma_1}\left[\Psi(f(X),Z)-\Psi(f(\widehat{X}),\widehat{Z})\right]\geq\lambda\rho^p+\lambda D.$$

Let $\gamma_0=(\mathrm{id}_{\mathcal{X}\times\mathcal{Z}}\otimes\mathrm{id}_{\mathcal{X}\times\mathcal{Z}})_\sharp\widehat{\mathbb{P}}$ denote the joint distribution induced by identity transport map. Let $\gamma_\theta=\theta\gamma_1+(1-\theta)\gamma_0$ be the transport plan which perturbs γ_0 by moving $\theta:=\min\{1,\frac{\rho^p}{D}\}$ portion of mass from $(\widehat{X},\widehat{Z})$ to (X,Z). By the convexity lemma 2 this transport plan is causal. Denote $\mathbb{P}_\theta=(\gamma_\theta)_{(X,Z)}$ to be the marginal of γ_θ. Then

$$\mathcal{C}_p(\widehat{\mathbb{P}},\mathbb{P})^p\leq\mathbb{E}_{\gamma_\theta}\left[\|X-\widehat{X}\|^p+\|Z-\widehat{Z}\|^p\right]=\theta D\leq\rho^p,$$

So $\mathbb{P}_\theta$ is primal feasible, and

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_\theta}[\Psi(f(X),Z)]-\mathbb{E}_{\widehat{\mathbb{P}}}[\Psi(f(\widehat{X}),\widehat{Z})]&=\mathbb{E}_{\gamma_\theta}\left[\Psi(f(X),Z)-\Psi(f(\widehat{X}),\widehat{Z})\right]\\&=\theta\mathbb{E}_{\gamma_1}\left[\Psi(f(x),Z)-\Psi(f(\widehat{x}),\widehat{Z})\right]\\&\geq\theta\left(\lambda\rho^p+\lambda D\right)\\&\geq\lambda\rho^p.\end{aligned}$$

Therefore

$$v_{\mathsf{P}}^f\geq\mathbb{E}_{\mathbb{P}_\theta}[\Psi(f(X),Z)]\geq\mathbb{E}_{\widehat{\mathbb{P}}}[\Psi(f(\widehat{X}),\widehat{Z})]+\lambda\rho^p,$$

and since λ can be arbitrarily large, we have

$$v_{\mathsf{P}}^f=\infty=v_{\mathsf{D}}^f.$$

Case 2: $\kappa<\infty,\lambda^*>\kappa$

Fix some small $\delta > 0$, $\varepsilon > 0$. Applying Lemma 11 on (B.6), for $\widehat{x} \in \mathcal{X}$ we can find $\overline{x}, \underline{x} \in \mathcal{X}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda^+} F_{(\underline{x})}(\lambda_1; \widehat{x}) \leq \frac{\mathrm{d}}{\mathrm{d}\lambda^-} \Phi(\lambda^*; \widehat{x}) + \delta, \qquad \frac{\mathrm{d}}{\mathrm{d}\lambda^-} F_{(\overline{x})}(\lambda_2; \widehat{x}) \geq \frac{\mathrm{d}}{\mathrm{d}\lambda^+} \Phi(\lambda^*; \widehat{x}) - \delta,$$
$$F_{(\underline{x})}(\lambda_1, \widehat{x}) \geq \Phi(\lambda^*, \widehat{x}) - \varepsilon, \qquad F_{(\overline{x})}(\lambda_2, \widehat{x}) \geq \Phi(\lambda^*, \widehat{x}) - \varepsilon$$

for $\kappa < \lambda_1 < \lambda^* < \lambda_2$ and λ_1, λ_2 sufficiently close to λ^*. Fix $x \in \mathcal{X}$. Apply Lemma 11 on (B.5), for $\widehat{z} \in \mathcal{Z}$ we can find $\overline{z}, \underline{z} \in \mathcal{Z}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda^+} G_{(\underline{z})}(\lambda_3; x, \widehat{z}) \leq \frac{\mathrm{d}}{\mathrm{d}\lambda^-} \Upsilon(\lambda_1; x, \widehat{z}) + \delta, \qquad \frac{\mathrm{d}}{\mathrm{d}\lambda^-} G_{(\overline{z})}(\lambda_4; x, \widehat{z}) \geq \frac{\mathrm{d}}{\mathrm{d}\lambda^+} \Upsilon(\lambda_2; x, \widehat{z}) - \delta,$$
$$G_{(\underline{z})}(\lambda_3; x, \widehat{z}) \geq \Upsilon(\lambda_1, x, \widehat{z}) - \varepsilon, \qquad G_{(\overline{z})}(\lambda_4; x, \widehat{z}) \geq \Upsilon(\lambda_2, x, \widehat{z}) - \varepsilon$$

for $\kappa < \lambda_3 < \lambda_1 < \lambda^* < \lambda_2 < \lambda_4$ and λ_3, λ_4 sufficiently close to λ_1, λ_2. Now suppose $\widehat{\mathbb{P}}$ is supported over a finite set of $\{(\widehat{x}_k, \widehat{z}_{ki})\}_{ki}$, we know that for $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ sufficiently close to λ^* we can find $\overline{x}_k, \underline{x}_k, \overline{z}_{ki}, \underline{z}_{ki}$ such that the above are satisfied simultaneously. We denote the transport map by $\overline{x}_k = \overline{T}_1(\widehat{x}_i)$, $\overline{z}_{ki} = \overline{T}_2(\widehat{x}_k, \widehat{z}_{ki})$, and $\overline{T}(\widehat{x}_k, \widehat{z}_{ki}) = (\overline{x}_k, \overline{z}_{ki})$. We define $\underline{T}$ similarly, so we can construct $(\overline{X}, \overline{Z}) = \overline{T}(\widehat{X}, \widehat{Z})$, $(\underline{X}, \underline{Z}) = \underline{T}(\widehat{X}, \widehat{Z})$. We denote the law of $((\overline{X}, \overline{Z}), (\underline{X}, \underline{Z}))$ by $\overline{\gamma} = (\overline{T} \otimes \mathrm{id}_{\mathcal{X}\times\mathcal{Z}})_\sharp \widehat{\mathbb{P}}$, and the law of $(\overline{X}, \overline{Z})$ is $\overline{\mathbb{P}} = \overline{\gamma}_{(X,Z)}$ the marginal. Similarly we define $\underline{\gamma}$ and $\underline{\mathbb{P}}$. We also define $\widehat{g} = (\mathrm{id}_{\mathcal{X}\times\mathcal{Z}} \otimes \mathrm{id}_{\mathcal{X}\times\mathcal{Z}})_\sharp \widehat{\mathbb{P}}$ to be the identity transport plan. For convenience, denote the law of $(\overline{X}, \widehat{X})$ to be $\overline{\gamma}_1 = \overline{\gamma}_{(X,\widehat{X})}$, and the law of $(\underline{X}, \widehat{X})$ to be $\underline{\gamma}_1 = \underline{\gamma}_{(X,\widehat{X})}$. Similarly define $\overline{\gamma}_2 = \overline{\gamma}_{(Z,\widehat{Z})|(X,\widehat{X})}$ and $\underline{\gamma}_2 = \underline{\gamma}_{(Z,\widehat{Z})|(X,\widehat{X})}$ to be the conditional law of $(\overline{Z}, \widehat{Z})$ and $(\underline{Z}, \widehat{Z})$ given $(\overline{X}, \widehat{X})$ and $(\underline{X}, \widehat{X})$, respectively.

We know that $h(\lambda)$ attains its minimum v_{D}^f at some $\lambda^* \in c_I$, so $h'(\lambda^*+) \geq 0$ and $h'(\lambda^*-) \leq 0$ (if $\lambda^* > \kappa$), so

$$\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^-}\right|_{\lambda=\lambda^*} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\Phi(\lambda, \widehat{X})\right] \leq -\rho^p \leq \left.\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\right|_{\lambda=\lambda^*} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\Phi(\lambda, \widehat{X})\right]$$

where

$$
\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^-}\right|_{\lambda=\lambda^*} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\Phi(\lambda,\widehat{X})\right] &= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^-}\right|_{\lambda=\lambda^*}\Phi(\lambda,\widehat{X})\right] \\
&\geq \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\right|_{\lambda=\lambda_1} F_{(\underline{X})}(\lambda;\widehat{X})\right] - \delta \\
&= \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\right|_{\lambda=\lambda_1} \left\{\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\lambda;\underline{X},\widehat{Z}) \mid (\underline{X},\widehat{X})\right] - \lambda\|\underline{X}-\widehat{X}\|^p\right\}\right] - \delta \\
&= \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\right|_{\lambda=\lambda_1}\Upsilon(\lambda;\underline{X},\widehat{Z}) \mid (\underline{X},\widehat{X})\right] - \|\underline{X}-\widehat{X}\|^p\right] - \delta \\
&\geq \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{(\underline{Z},\widehat{Z})\sim\underline{\gamma}_2}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\right|_{\lambda=\lambda_3} G_{(\underline{Z})}(\lambda;\underline{X},\widehat{Z}) \mid (\underline{X},\widehat{X})\right] - \|\underline{X}-\widehat{X}\|^p\right] - 2\delta \\
&\geq \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{(\underline{Z},\widehat{Z})\sim\underline{\gamma}_2}\left[-\|\underline{Z}-\widehat{Z}\|^p \mid (\underline{X},\widehat{X})\right] - \|\underline{X}-\widehat{X}\|^p\right] - 2\delta \\
&= -\mathbb{E}_{((\underline{X},\underline{Z}),(\widehat{X},\widehat{Z}))\sim\underline{\gamma}}\left[\|\underline{X}-\widehat{X}\|^p + \|\underline{Z}-\widehat{Z}\|^p\right] - 2\delta,
\end{aligned}
$$

$$
\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\right|_{\lambda=\lambda^*} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\Phi(\lambda,\widehat{X})\right] &= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\right|_{\lambda=\lambda^*}\Phi(\lambda,\widehat{X})\right] \\
&\leq \mathbb{E}_{(\overline{X},\widehat{X})\sim\overline{\gamma}_1}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^-}\right|_{\lambda=\lambda_2} F_{(\overline{X})}(\lambda;\widehat{X})\right] + \delta \\
&= \mathbb{E}_{(\overline{X},\widehat{X})\sim\overline{\gamma}_1}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^-}\right|_{\lambda=\lambda_2} \left\{\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\lambda;\overline{X},\widehat{Z}) \mid (\overline{X},\widehat{X})\right] - \lambda\|\overline{X}-\widehat{X}\|^p\right\}\right] + \delta \\
&= \mathbb{E}_{(\overline{X},\widehat{X})\sim\overline{\gamma}_1}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^-}\right|_{\lambda=\lambda_2}\Upsilon(\lambda;\overline{X},\widehat{Z}) \mid (\overline{X},\widehat{X})\right] - \|\overline{X}-\widehat{X}\|^p\right] + \delta \\
&\leq \mathbb{E}_{(\overline{X},\widehat{X})\sim\overline{\gamma}_1}\left[\mathbb{E}_{(\overline{Z},\widehat{Z})\sim\overline{\gamma}_2}\left[\left.\frac{\mathrm{d}}{\mathrm{d}\lambda^-}\right|_{\lambda=\lambda_4} G_{(\overline{Z})}(\lambda;\overline{X},\widehat{Z}) \mid (\overline{X},\widehat{X})\right] - \|\overline{X}-\widehat{X}\|^p\right] + 2\delta \\
&\leq \mathbb{E}_{(\overline{X},\widehat{X})\sim\overline{\gamma}_1}\left[\mathbb{E}_{(\overline{Z},\widehat{Z})\sim\overline{\gamma}_2}\left[-\|\overline{Z}-\widehat{Z}\|^p \mid (\overline{X},\widehat{X})\right] - \|\overline{X}-\widehat{X}\|^p\right] + 2\delta \\
&= -\mathbb{E}_{((\overline{X},\overline{Z}),(\widehat{X},\widehat{Z}))\sim\overline{\gamma}}\left[\|\underline{X}-\widehat{X}\|^p + \|\underline{Z}-\widehat{Z}\|^p\right] + 2\delta,
\end{aligned}
$$

Therefore,

$$\overline{d} := \mathbb{E}_{((\overline{X},\overline{Z}),(\widehat{X},\widehat{Z}))\sim\overline{\gamma}} \left[\|\underline{X} - \widehat{X}\|^p + \|\underline{Z} - \widehat{Z}\|^p\right] \leq \rho^p + 2\delta,$$
$$\underline{d} := \mathbb{E}_{((\underline{X},\underline{Z}),(\widehat{X},\widehat{Z}))\sim\underline{\gamma}} \left[\|\underline{X} - \widehat{X}\|^p + \|\underline{Z} - \widehat{Z}\|^p\right] \geq \rho^p - 2\delta.$$

Based on these, we construct a feasible primal solution. There exists $q_\delta^\varepsilon \in [0, 1]$ depending on $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, such that

$$\begin{aligned}\rho^p &= (1 - q_\delta^\varepsilon)\left(\overline{d} - 2\delta\right) + q_\delta^\varepsilon\left(\underline{d} + 2\delta\right),\\ \rho^p + 2(1 - 2q_\delta^\varepsilon)\delta &= (1 - q_\delta^\varepsilon)\overline{d} + q_\delta^\varepsilon\underline{d}.\end{aligned}$$

Let $q^\delta := \frac{\rho^p}{\rho^p + 2(1-2q_\delta^\varepsilon)_+\delta} \leq 1$. Define a transport plan $\gamma_\delta^\varepsilon$ by

$$\gamma_\delta^\varepsilon := q^\delta\left[(1 - q_\delta^\varepsilon)\overline{\gamma} + q_\delta^\varepsilon\underline{\gamma}\right] + (1 - q^\delta)\widehat{g}.$$

Its marginal distribution $\mathbb{P}_\delta^\varepsilon = (\gamma_\delta^\varepsilon)_{(X,Z)}$ is given by

$$\mathbb{P}_\delta^\varepsilon = q^\delta\left[(1 - q_\delta^\varepsilon)\overline{\mathbb{P}} + q_\delta^\varepsilon\underline{\mathbb{P}}\right] + (1 - q^\delta)\widehat{\mathbb{P}}.$$

Then $\mathbb{P}_\delta^\varepsilon$ is primal feasible, because

$$\begin{aligned}\mathcal{C}_p(\mathbb{P}_\delta^\varepsilon, \widehat{\mathbb{P}})^p &\leq \mathbb{E}_{((X,Z),(\widehat{X},\widehat{Z}))\sim\gamma_\delta^\varepsilon}\left[\|X - \widehat{X}\|^p + \|Z - \widehat{Z}\|^p\right]\\ &\leq q^\delta\left[(1 - q_\delta^\varepsilon)\overline{d} + q_\delta^\varepsilon\underline{d}\right] \leq \rho^p.\end{aligned}$$

In the mean time,

$$
\begin{aligned}
v_{\mathsf{D}}^{f} - \lambda^* \rho^p &= h(\lambda^*) - \lambda^* \rho^p \\
&= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\Phi(\lambda^*, \widehat{X}) \right] \\
&\le \mathbb{E}_{(\underline{X}, \widehat{X}) \sim \underline{\gamma}_1} \left[F_{(\underline{X})}(\lambda_1; \widehat{X}) \right] + \varepsilon \\
&= \mathbb{E}_{(\underline{X}, \widehat{X}) \sim \underline{\gamma}_1} \left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}} \left[\Upsilon(\lambda_1; \underline{X}, \widehat{Z}) \mid \widehat{X} \right] - \lambda_1 \| \underline{X} - \widehat{X} \|^p \right] + \varepsilon \\
&\le \mathbb{E}_{(\underline{X}, \widehat{X}) \sim \underline{\gamma}_1} \left[\mathbb{E}_{(\underline{Z}, \widehat{Z}) \sim \underline{\gamma}_2} \left[G_{(\underline{Z})}(\lambda_3; \underline{X}, \widehat{Z}) \mid (\underline{X}, \widehat{X}) \right] - \lambda_1 \| \underline{X} - \widehat{X} \|^p \right] + 2\varepsilon \\
&\le \mathbb{E}_{(\underline{X}, \widehat{X}) \sim \underline{\gamma}_1} \left[\mathbb{E}_{(\underline{Z}, \widehat{Z}) \sim \underline{\gamma}_2} \left[\Psi(f(\underline{X}), \underline{Z}) - \lambda_3 \| \underline{Z} - \widehat{Z} \|^p \mid (\underline{X}, \widehat{X}) \right] - \lambda_1 \| \underline{X} - \widehat{X} \|^p \right] + 2\varepsilon \\
&\le \mathbb{E}_{((\underline{X}, \underline{Z}), (\widehat{X}, \widehat{Z})) \sim \underline{\gamma}} \left[\Psi(f(\underline{X}), \underline{Z}) - \lambda_3 \| \underline{Z} - \widehat{Z} \|^p - \lambda_1 \| \underline{X} - \widehat{X} \|^p \right] + 2\varepsilon \\
&\le \mathbb{E}_{\underline{\mathbb{P}}} \left[\Psi(f(\underline{X}), \underline{Z}) \right] - \lambda_3 \underline{d} + 2\varepsilon.
\end{aligned}
$$

Similarly

$$
\begin{aligned}
v_{\mathsf{D}}^{f} - \lambda^* \rho^p &= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\Phi(\lambda^*, \widehat{X}) \right] \\
&\le \mathbb{E}_{(\overline{X}, \widehat{X}) \sim \overline{\gamma}_1} \left[F_{(\overline{X})}(\lambda_2; \widehat{X}) \right] + \varepsilon \\
&= \mathbb{E}_{(\overline{X}, \widehat{X}) \sim \overline{\gamma}_1} \left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}} \left[\Upsilon(\lambda_2; \overline{X}, \widehat{Z}) \mid (\overline{X}, \widehat{X}) \right] - \lambda_2 \| \overline{X} - \widehat{X} \|^p \right] + \varepsilon \\
&\le \mathbb{E}_{(\overline{X}, \widehat{X}) \sim \overline{\gamma}_1} \left[\mathbb{E}_{(\overline{Z}, \widehat{Z}) \sim \overline{\gamma}_2} \left[G_{(\overline{Z})}(\lambda_4; \overline{X}, \widehat{Z}) \mid (\overline{X}, \widehat{X}) \right] - \lambda_2 \| \overline{X} - \widehat{X} \|^p \right] + 2\varepsilon \\
&\le \mathbb{E}_{((\overline{X}, \overline{Z}), (\widehat{X}, \widehat{Z})) \sim \overline{\gamma}} \left[\Psi(f(\overline{X}), \overline{Z}) - \lambda_4 \| \overline{Z} - \widehat{Z} \|^p - \lambda_2 \| \overline{X} - \widehat{X} \|^p \right] + 2\varepsilon \\
&\le \mathbb{E}_{\overline{\mathbb{P}}} \left[\Psi(f(\overline{X}), \overline{Z}) \right] - \lambda_2 \overline{d} + 2\varepsilon.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
v_{\mathsf{P}}^{f} &\geq \mathbb{E}_{(X,Z)\sim\mathbb{P}_{\delta}^{\varepsilon}}\left[\Psi(f(X),Z)\right] \\
&= q^{\delta}\left((1-q_{\delta}^{\varepsilon})\mathbb{E}_{\overline{\mathbb{P}}}\left[\Psi(f(\overline{X}),\overline{Z})\right] + q_{\delta}^{\varepsilon}\mathbb{E}_{\underline{\mathbb{P}}}\left[\Psi(f(\underline{X}),\underline{Z})\right]\right) + (1-q^{\delta})\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right] \\
&\geq q^{\delta}\left((1-q_{\delta}^{\varepsilon})\left(v_{\mathsf{D}}^{f} - \lambda^{*}\rho^{p} + \lambda_{2}\overline{d} - 2\varepsilon\right) + q_{\delta}^{\varepsilon}\left(v_{\mathsf{D}}^{f} - \lambda^{*}\rho^{p} + \lambda_{3}\underline{d} - 2\varepsilon\right)\right) + (1-q^{\delta})\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right] \\
&\geq q^{\delta}\left(v_{\mathsf{D}}^{f} - \lambda^{*}\rho^{p} + \lambda_{3}((1-q_{\delta}^{\varepsilon})\overline{d} + q_{\delta}^{\varepsilon}\underline{d}) - 2\varepsilon\right) + (1-q^{\delta})\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right] \\
&\geq q^{\delta}\left(v_{\mathsf{D}}^{f} - \lambda^{*}\rho^{p} + \lambda_{3}(\rho^{p} + 2(1-2q_{\delta}^{\varepsilon})\delta) - 2\varepsilon\right) + (1-q^{\delta})\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right] \\
&= q^{\delta}\left(v_{\mathsf{D}}^{f} - (\lambda^{*}-\lambda_{3})\rho^{p} + 2\lambda_{3}(1-2q_{\delta}^{\varepsilon})\delta - 2\varepsilon\right) + (1-q^{\delta})\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right].
\end{aligned}
$$

As $\delta \to 0$, $q^{\delta} \to 1$. Thus take the limit as $\lambda_3 \to \lambda^*$ and $\delta \to 0$, it follows that

$$v_{\mathsf{P}}^{f} \geq v_{\mathsf{D}}^{f} - 2\varepsilon.$$

Since ε can be taken arbitrarily small, $v_{\mathsf{P}}^{f} \geq v_{\mathsf{D}}^{f}$.

Case 3: $\lambda^* = \kappa < \infty$

In this case, we can still choose $\overline{x}, \overline{z}$, and we still have

$$F_{(\overline{x})}(\lambda_2, \widehat{x}) > \Phi(\lambda^*, \widehat{x}) - \varepsilon, \qquad G_{(\overline{z})}(\lambda_4; x, \widehat{z}) > \Upsilon(\lambda_2, x, \widehat{z}) - \varepsilon.$$

and

$$\overline{d} = \mathbb{E}_{\overline{\gamma}}\left[\|\overline{X} - \widehat{X}\|^p + \|\overline{Z} - \widehat{Z}\|^p\right] \leq \rho^p + 2\delta.$$

We separate the cases $\kappa = 0$ and $\kappa > 0$.

Case 3.1: $\lambda^* = \kappa = 0$

Let $q^{\delta} := \frac{\rho^p}{\rho^p + 2\delta} \leq 1$. Define $\gamma_{\delta}^{\varepsilon} := q^{\delta}\overline{\gamma} + (1 - q^{\delta})\widehat{g}$, then its marginal is a distribution $\mathbb{P}_{\delta}^{\varepsilon}$ given by

$$\mathbb{P}_{\delta}^{\varepsilon} := q^{\delta}\overline{\mathbb{P}} + (1 - q^{\delta})\widehat{\mathbb{P}}.$$

Then it is primal feasible, because

$$\mathcal{C}_p(\mathbb{P}_\delta^\varepsilon, \widehat{\mathbb{P}})^p \le \mathbb{E}_{\gamma_\delta^\varepsilon}\left[\|\overline{X} - \widehat{X}\|^p + \|\overline{Z} - \widehat{Z}\|^p\right] \le q^\delta \overline{d} \le \rho^p,$$

thus

$$\begin{aligned} v_{\mathsf{P}}^f &\ge \mathbb{E}_{(X,Z)\sim\mathbb{P}_\delta^\varepsilon}[\Psi(f(X),Z)] \\ &= q^\delta \mathbb{E}_{(\overline{X},\overline{Z})\sim\overline{\mathbb{P}}}\left[\Psi(f(\overline{X}),\overline{Z})\right] + (1-q^\delta)\mathbb{E}_{\widehat{\mathbb{P}}}[\Psi(f(\widehat{X}),\widehat{Z})] \\ &\ge q^\delta\left(v_{\mathsf{D}}^f - \lambda^*\rho^p + \lambda_2\overline{d} - 2\varepsilon\right) + (1-q^\delta)\mathbb{E}_{\widehat{\mathbb{P}}}[\Psi(f(\widehat{X}),\widehat{Z})] \\ &\ge q^\delta\left(v_{\mathsf{D}}^f - 2\varepsilon\right) + (1-q^\delta)\mathbb{E}_{\widehat{\mathbb{P}}}[\Psi(f(\widehat{X}),\widehat{Z})] \end{aligned}$$

using $\lambda^* = 0$. Let $\delta \to 0$, $q^\delta \to 1$, we have $v_{\mathsf{P}}^f \ge v_{\mathsf{D}}^f - 2\varepsilon$, and by taking $\varepsilon \to 0$ we have $v_{\mathsf{P}}^f \ge v_{\mathsf{D}}^f$.

Case 3.2: $\lambda^* = \kappa > 0$

Fix any $0 < \kappa' < \kappa$. We have

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\Phi(\kappa';\widehat{X}) - \Phi(\kappa;\widehat{X})\right] = h(\kappa') - h(\kappa) = \infty. \tag{B.7}$$

We denote

$$\mathcal{X}^*(\lambda;\widehat{x}) := \left\{x \in \mathcal{X} : F_{(x)}(\lambda;\widehat{x}) \ge F_{(\widehat{x})}(\lambda;\widehat{x})\right\}.$$

Then $\mathcal{X}^*(\lambda;\widehat{x})$ is nonempty because $\widehat{x} \in \mathcal{X}^*(\lambda;\widehat{x})$. Since

$$\Phi(\kappa';\widehat{x}) = \sup_{x\in\mathcal{X}} F_{(x)}(\kappa';\widehat{x}) = \sup_{x\in\mathcal{X}^*(\kappa';\widehat{x})} F_{(x)}(\kappa';\widehat{x}),$$

we can rewrite (B.7) as

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\sup_{x\in\mathcal{X}^*(\kappa';\widehat{X})} F_{(x)}(\kappa';\widehat{X}) - \Phi(\kappa;\widehat{X})\right] = \infty$$

Thus for any fixed $R > 0$, we can pick $\underline{X} = \underline{T}_1(\widehat{X}) \in \mathcal{X}^*(\kappa'; \widehat{X})$, which induces $\underline{\gamma}_1$, such that

$$\begin{aligned}
R &< \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[F_{(\underline{X})}(\kappa'; \widehat{X}) - \Phi(\kappa; \widehat{X})\right] \\
&\leq \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[F_{(\underline{X})}(\kappa'; \widehat{X}) - F_{(\underline{X})}(\kappa; \widehat{X})\right] \\
&= \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\kappa'; \underline{X}, \widehat{Z}) - \Upsilon(\kappa; \underline{X}, \widehat{Z}) \mid (\underline{X}, \widehat{X})\right] + (\kappa - \kappa')\|\underline{X} - \widehat{X}\|^p\right]. \quad \text{(B.8)}
\end{aligned}$$

Moreover, because $\underline{X} \in \mathcal{X}^*(\kappa'; \widehat{X})$, we have

$$\begin{aligned}
F_{(\widehat{X})}(\kappa'; \widehat{X}) &\leq F_{(\underline{X})}(\kappa'; \widehat{X}), \\
\kappa'\|\underline{X} - \widehat{X}\|^p &\leq \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\kappa'; \underline{X}, \widehat{Z}) - \Upsilon(\kappa'; \widehat{X}, \widehat{Z}) \mid \widehat{X}\right], \\
\mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\kappa'\|\underline{X} - \widehat{X}\|^p\right] &\leq \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\kappa'; \underline{X}, \widehat{Z}) - \Upsilon(\kappa'; \widehat{X}, \widehat{Z}) \mid (\underline{X}, \widehat{X})\right]\right]. \quad \text{(B.9)}
\end{aligned}$$

We denote

$$\mathcal{Z}^*(\lambda; x, \widehat{z}) := \left\{z \in \mathcal{Z} : G_{(z)}(\lambda; x, \widehat{z}) \geq G_{(\widehat{z})}(\lambda; x, \widehat{z})\right\}.$$

Then $\mathcal{Z}^*(\lambda; x, \widehat{z})$ is nonempty because $\widehat{z} \in \mathcal{Z}^*(\lambda; x, \widehat{z})$. Since

$$\Upsilon(\kappa'; x, \widehat{z}) = \sup_{z\in\mathcal{Z}} G_{(z)}(\kappa'; x, \widehat{z}) = \sup_{z\in\mathcal{Z}^*(\kappa'; x,\widehat{z})} G_{(z)}(\kappa'; x, \widehat{z}),$$

we can rewrite (B.8) and (B.9) as

$$R < \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}^*(\kappa';\underline{X},\widehat{Z})} G_{(z)}(\kappa'; \underline{X}, \widehat{Z}) - \Upsilon(\kappa; \underline{X}, \widehat{Z}) \mid (\underline{X}, \widehat{X})\right] + (\kappa - \kappa')\|\underline{X} - \widehat{X}\|^p\right],$$

$$\mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\kappa'\|\underline{X} - \widehat{X}\|^p\right] \leq \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}^*(\kappa';\underline{X},\widehat{Z})} G_{(z)}(\kappa'; \underline{X}, \widehat{Z}) - \Upsilon(\kappa'; \widehat{X}, \widehat{Z}) \mid (\underline{X}, \widehat{X})\right]\right].$$

Thus we can pick $\underline{Z} = \underline{T}_2(\widehat{X}, \widehat{Z}) \in \mathcal{Z}^*(\kappa'; \underline{X}, \widehat{Z})$, which induces γ_2, such that

$$\begin{aligned}
R - \varepsilon &< \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{(\underline{Z},\widehat{Z})\sim\underline{\gamma}_2}\left[G_{(\underline{Z})}(\kappa'; \underline{X}, \widehat{Z}) - \Upsilon(\kappa; \underline{X}, \widehat{Z}) \mid (\underline{X}, \widehat{X})\right] + (\kappa - \kappa')\|\underline{X} - \widehat{X}\|^p\right] \\
&\leq \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{(\underline{Z},\widehat{Z})\sim\underline{\gamma}_2}\left[G_{(\underline{Z})}(\kappa'; \underline{X}, \widehat{Z}) - G_{(\underline{Z})}(\kappa; \underline{X}, \widehat{Z}) \mid (\underline{X}, \widehat{X})\right] + (\kappa - \kappa')\|\underline{X} - \widehat{X}\|^p\right] \\
&= \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{(\underline{Z},\widehat{Z})\sim\underline{\gamma}_2}\left[(\kappa - \kappa')\|\underline{Z} - \widehat{Z}\|^p \mid (\underline{X}, \widehat{X})\right] + (\kappa - \kappa')\|\underline{X} - \widehat{X}\|^p\right] \\
&= (\kappa - \kappa')\underline{d},
\end{aligned}$$

and simultaneously ensure

$$\begin{aligned}
\mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\kappa'\|\underline{X}-\widehat{X}\|^p\right]-\delta &\le \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{(\underline{Z},\widehat{Z})\sim\underline{\gamma}_2}\left[G_{(\underline{Z})}(\kappa';\underline{X},\widehat{Z})-\Upsilon(\kappa';\widehat{X},\widehat{Z})\mid(\underline{X},\widehat{X})\right]\right]\\
&\le \mathbb{E}_{(\underline{X},\widehat{X})\sim\underline{\gamma}_1}\left[\mathbb{E}_{(\underline{Z},\widehat{Z})\sim\underline{\gamma}_2}\left[G_{(\underline{Z})}(\kappa';\underline{X},\widehat{Z})-G_{(\widehat{Z})}(\kappa';\widehat{X},\widehat{Z})\mid(\underline{X},\widehat{X})\right]\right]\\
&= \mathbb{E}_{((\underline{X},\underline{Z}),(\widehat{X},\widehat{Z}))\sim\underline{\gamma}}\left[\Psi(f(\underline{X}),\underline{Z})-\kappa'\|\underline{Z}-\widehat{Z}\|^p-\Psi(f(\widehat{X}),\widehat{Z})\right],\\
\kappa'\underline{d} &\le \mathbb{E}_{((\underline{X},\underline{Z}),(\widehat{X},\widehat{Z}))\sim\underline{\gamma}}\left[\Psi(f(\underline{X}),\underline{Z})-\Psi(f(\widehat{X}),\widehat{Z})\right]\\
&\le \mathbb{E}_{\underline{\mathbb{P}}}\left[\Psi(f(\underline{X}),\underline{Z})\right]-\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right].
\end{aligned}$$

In conclusion, we have

$$\frac{R-\varepsilon}{\kappa-\kappa'}<\underline{d}\le\frac{\mathbb{E}_{\underline{\mathbb{P}}}\left[\Psi(f(\underline{X}),\underline{Z})\right]-\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right]}{\kappa'}.$$

We can choose $R=\varepsilon+(\kappa-\kappa')N\rho^p$ for some $N>>1$ to be specified later. Because

$$\overline{d}-2\delta\le\rho^p\le\frac{\underline{d}}{N}\le\underline{d}+2\delta,$$

there exists $q_\delta^\varepsilon\in[0,1]$ depending on $\lambda_2,\lambda_4,\kappa'$, such that

$$\begin{aligned}
\rho^p &= (1-q_\delta^\varepsilon)\left[\overline{d}-2\delta\right]+q_\delta^\varepsilon\left[\underline{d}+2\delta\right],\\
&= (1-q_\delta^\varepsilon)\overline{d}+q_\delta^\varepsilon\underline{d}-2(1-2q_\delta^\varepsilon)\delta,\\
\rho^p+2(1-2q_\delta^\varepsilon)\delta &= (1-q_\delta^\varepsilon)\overline{d}+q_\delta^\varepsilon\underline{d}.
\end{aligned}$$

Let $q^\delta:=\frac{\rho^p}{\rho^p+2(1-2q_\delta^\varepsilon)_+\delta}\le 1$. Define a distribution $\mathbb{P}_\delta^\varepsilon$ by

$$\mathbb{P}_\delta^\varepsilon:=q^\delta\left[(1-q_\delta^\varepsilon)\overline{\mathbb{P}}+q_\delta^\varepsilon\underline{\mathbb{P}}\right]+\left(1-q^\delta\right)\widehat{\mathbb{P}}.$$

Then $\mathbb{P}_\delta^\varepsilon$ is primal feasible, because

$$\begin{aligned}
\mathcal{C}_p(\mathbb{P}_\delta^\varepsilon,\widehat{\mathbb{P}})^p &\le q^\delta(1-q_\delta^\varepsilon)\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\|\overline{Z}-\widehat{Z}\|^p\mid\widehat{X}\right]+\|\overline{X}-\widehat{X}\|^p\right]\\
&\quad+q^\delta q_\delta^\varepsilon\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\|\underline{Z}-\widehat{Z}\|^p\mid\widehat{X}\right]+\|\underline{X}-\widehat{X}\|^p\right]\\
&\le q^\delta\left[(1-q_\delta^\varepsilon)\overline{d}+q_\delta^\varepsilon\underline{d}\right]\le\rho^p.
\end{aligned}$$

Therefore

$$
\begin{aligned}
v_{\mathsf{P}}^{f} &\geq \mathbb{E}_{(X,Z)\sim\mathbb{P}_{\delta}^{\varepsilon}}\left[\Psi(f(X),Z)\right]\\
&= \mathbb{E}_{\overline{\mathbb{P}}}\left[q^{\delta}(1-q_{\delta}^{\varepsilon})\Psi(f(\overline{X}),\overline{Z})\right] + \mathbb{E}_{\underline{\mathbb{P}}}\left[q^{\delta}q_{\delta}^{\varepsilon}\Psi(f(\underline{X}),\underline{Z})\right] + \mathbb{E}_{\widehat{\mathbb{P}}}\left[(1-q^{\delta})\Psi(f(\widehat{X}),\widehat{Z})\right]\\
&\geq q^{\delta}(1-q_{\delta}^{\varepsilon})\left(v_{\mathsf{D}}^{f} - \kappa\rho^{p} + \lambda_2\overline{d} - 2\varepsilon\right) + q^{\delta}q_{\delta}^{\varepsilon}\kappa'\underline{d}\\
&\qquad + \left(1 - q^{\delta} + q^{\delta}q_{\delta}^{\varepsilon}\right)\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right]\\
&\geq q^{\delta}\kappa'\left((1-q_{\delta}^{\varepsilon})\overline{d} + q_{\delta}^{\varepsilon}\underline{d}\right) + q^{\delta}(1-q_{\delta}^{\varepsilon})(v_{\mathsf{D}}^{f} - \kappa\rho^{p} - 2\varepsilon)\\
&\qquad + \left(1 - q^{\delta} + q^{\delta}q_{\delta}^{\varepsilon}\right)\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right]\\
&\geq q^{\delta}\kappa'(\rho^{p} + 2(1-2q_{\delta}^{\varepsilon})\delta) + q^{\delta}(1-q_{\delta}^{\varepsilon})(v_{\mathsf{D}}^{f} - \kappa\rho^{p} - 2\varepsilon) + \left(1 - q^{\delta} + q^{\delta}q_{\delta}^{\varepsilon}\right)\mathbb{E}_{\widehat{\mathbb{P}}}\left[\Psi(f(\widehat{X}),\widehat{Z})\right].
\end{aligned}
$$

As $\delta \to 0$, we have $q^{\delta} \to 1$. Moreover, because

$$\rho^{p} + 2\delta \geq (1-q_{\delta}^{\varepsilon})\overline{d} + q_{\delta}^{\varepsilon}\underline{d} \geq q_{\delta}^{\varepsilon}\underline{d} \geq q_{\delta}^{\varepsilon}N\rho^{p},$$

we know that $q_{\delta}^{\varepsilon} \leq \frac{1+2\delta\rho^{-p}}{N} \to 0$ as $N \to \infty$ and $\delta \to 0$. Therefore by taking these limit, we have

$$v_{\mathsf{P}}^{f} \geq \kappa'\rho^{p} + v_{\mathsf{D}}^{f} - \kappa\rho^{p} - 2\varepsilon = v_{\mathsf{D}}^{f} - 2\varepsilon - (\kappa - \kappa')\rho^{p}.$$

Since this is true for any $\kappa' < \kappa$ and $\varepsilon > 0$, we may take $\kappa' \to \kappa$ and $\varepsilon \to 0$ so $v_{\mathsf{P}}^{f} \geq v_{\mathsf{D}}^{f}$.

Proof. Proof of Theorem 3. Since $\Psi(f(\cdot),\cdot)$ is upper semicontinuous, we know that for each fixed $x \in \mathcal{X}$, $\widehat{z} \in \mathcal{Z}$, $\lambda > \kappa$, $G_{(z)}(\lambda; x, \widehat{z}) = \Psi(f(x), z) - \lambda|z - \widehat{z}|^{p}$ is upper semicontinuous in z. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}G_{(z)}(\lambda; x, \widehat{z}) = -|z - \widehat{z}|^{p} \to -\infty \qquad \text{as } |z| \to \infty,$$

By Lemma 12 (2), we can find $\overline{z}$, $\underline{z}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda^{+}}\Upsilon(\lambda; x, \widehat{z}) = -|\overline{z} - \widehat{z}|^{p}, \qquad \frac{\mathrm{d}}{\mathrm{d}\lambda^{-}}\Upsilon(\lambda; x, \widehat{z}) = -|\underline{z} - \widehat{z}|^{p}, \qquad \Upsilon(\lambda; x, \widehat{z}) = G_{(\overline{z})}(\lambda; x, \widehat{z}) = G_{(\underline{z})}(\lambda; x, \widehat{z}).$$

Now we claim that for each fixed $\widehat{z} \in \mathcal{Z}$, $\lambda > \kappa$, $\Upsilon(\lambda; x, \widehat{z})$ is upper semicontinuous in x. We prove by contradiction. Assume otherwise, then we can find $x_k \to x$, such that

$$\Upsilon(\lambda; x_k, \widehat{z}) > \Upsilon(\lambda; x, \widehat{z}) + \varepsilon$$

for all k. We can find $\underline{z}_k$ such that

$$\Upsilon(\lambda; x_k, \widehat{z}) = G_{(\underline{z}_k)}(\lambda; x_k, \widehat{z}), \qquad \frac{\mathrm{d}}{\mathrm{d}\lambda^-}\Upsilon(\lambda; x, \widehat{z}) = -|\underline{z}_k - \widehat{z}|^p.$$

If $\underline{z}_k$ is bounded, then up to a subsequence it converges to $\underline{z}_\infty$, and since G is upper semicontinuous,

$$\limsup_{k\to\infty} \Upsilon(\lambda; x_k, \widehat{z}) = \limsup_{k\to\infty} G_{(\underline{z}_k)}(\lambda; x_k, \widehat{z}) \le G_{(\underline{z}_\infty)}(\lambda; x, \widehat{z}) \le \Upsilon(\lambda; x, \widehat{z})$$

which is a contradiction. If $\underline{z}_k$ is unbounded, then up to a subsequence, for $\lambda' \in (\kappa, \lambda)$,

$$\begin{aligned} \Upsilon(\lambda'; x_k, \widehat{z}) &\ge \Upsilon(\lambda; x_k, \widehat{z}) - (\lambda - \lambda')\frac{\mathrm{d}}{\mathrm{d}\lambda^-}\Upsilon(\lambda; x_k, \widehat{z}) \\ &\ge \Upsilon(\lambda; x, \widehat{z}) + \varepsilon + (\lambda - \lambda')|\underline{z}_k - \widehat{z}|^p \to \infty \end{aligned}$$

as $k \to \infty$. Therefore

$$\begin{aligned} \lim_{k\to\infty} F_{(x_k)}(\lambda', \widehat{x}) &= \lim_{k\to\infty} \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\lambda; x_k, \widehat{Z}) \mid \widehat{X} = \widehat{x}\right] - \lambda'\|x_k - \widehat{x}\|^p \\ &= \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\lim_{k\to\infty} \Upsilon(\lambda; x_k, \widehat{Z}) \mid \widehat{X} = \widehat{x}\right] - \lambda'\|x - \widehat{x}\|^p = \infty. \end{aligned}$$

This contradicts with $\Phi(\lambda', \widehat{x}) < \infty$.

We can thus construct $\overline{Z}, \underline{Z}$ which depends on λ, $\widehat{Z}$ and x. Now we have

$$F_{(x)}(\lambda; \widehat{x}) = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Upsilon(\lambda; x, \widehat{Z})|\widehat{X} = \widehat{x}\right] - \lambda|x - \widehat{x}|^p.$$

It is upper semicontinuous in x, because each $\Upsilon(\lambda; x, \widehat{z})$ is upper semicontinuous in x, and the finite sum of upper semicontinuous functions is upper semicontinuous. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda^+}F_{(x)}(\lambda;\widehat{x}) = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\Upsilon(\lambda; x, \widehat{Z})|\widehat{X} = \widehat{x}\right] - |x - \widehat{x}|^p = -\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[|\overline{Z} - \widehat{Z}|^p|\widehat{X} = \widehat{x}\right] - |x - \widehat{x}|^p \to -\infty$$

as $x \to \infty$. By Lemma 12 (2) we can find $\overline{x}$ and $\underline{x}$ such that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\Phi(\lambda;\widehat{x}) = -\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[|\overline{Z} - \widehat{Z}|^p|\widehat{X} = \widehat{x}\right] - |\overline{x} - \widehat{x}|^p, \qquad \frac{\mathrm{d}}{\mathrm{d}\lambda^-}\Phi(\lambda;\widehat{x}) = -\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[|\underline{Z} - \widehat{Z}|^p|\widehat{X} = \widehat{x}\right] - |\underline{x} - \widehat{x}|^p,$$
$$\Phi(\lambda;\widehat{x}) = F_{(\underline{x})}(\lambda;\widehat{x}) = F_{(\overline{x})}(\lambda;\widehat{x}).$$

By constructing these for every $\widehat{x}$ in the support of $\widehat{\mathbb{P}}_{\widehat{X}}$, we have $\overline{X}$, $\underline{X}$, $\overline{Z}$, $\underline{Z}$ such that $((\overline{X}, \overline{Z}), (\widehat{X}, \widehat{Z})) \sim \overline{\gamma}$, $((\underline{X}, \underline{Z}), (\widehat{X}, \widehat{Z})) \sim \underline{\gamma}$, where

$$\overline{\gamma} = \sum_{k=1}^{K}\sum_{i=1}^{n_k}\hat{p}_{ki}\delta_{((\overline{x}_k,\overline{z}_{ki}),(\widehat{x}_k,\widehat{z}_{ki}))}, \qquad \underline{\gamma} = \sum_{k=1}^{K}\sum_{i=1}^{n_k}\hat{p}_{ki}\delta_{((\underline{x}_k,\underline{z}_{ki}),(\widehat{x}_k,\widehat{z}_{ki}))}.$$

We use notations $\overline{\gamma}_1, \underline{\gamma}_1, \overline{\gamma}_2, \underline{\gamma}_2$ similar as in the proof of theorem 1.

Now we have both

$$\begin{aligned}
h(\lambda) &= \lambda\rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\Phi(\lambda;\widehat{X})\right] \\
&= \lambda\rho^p + \mathbb{E}_{\overline{\gamma}_1}\left[F_{(\overline{X})}(\lambda;\widehat{X})\right] \\
&= \lambda\rho^p + \mathbb{E}_{\overline{\gamma}_1}\left[\mathbb{E}_{\overline{\gamma}_2}\left[\Upsilon(\lambda;\overline{X},\widehat{Z})|(\overline{X},\widehat{X})\right] - \lambda|\overline{X} - \widehat{X}|^p\right] \\
&= \lambda\rho^p + \mathbb{E}_{\overline{\gamma}_1}\left[\mathbb{E}_{\overline{\gamma}_2}\left[G_{(\overline{Z})}(\lambda;\overline{X},\widehat{Z})|(\overline{X},\widehat{X})\right] - \lambda|\overline{X} - \widehat{X}|^p\right] \\
&= \lambda\rho^p + \mathbb{E}_{\overline{\gamma}_1}\left[\mathbb{E}_{\overline{\gamma}_2}\left[\Psi(f(\overline{X}),\overline{Z}) - \lambda|\overline{Z} - \widehat{Z}|^p|(\overline{X},\widehat{X})\right] - \lambda|\overline{X} - \widehat{X}|^p\right] \\
&= \lambda\left(\rho^p - \overline{d}\right) + \mathbb{E}_{\overline{\mathbb{P}}}\left[\Psi(f(\overline{X}),\overline{Z})\right], \\
h(\lambda) &= \lambda\left(\rho^p - \underline{d}\right) + \mathbb{E}_{\underline{\mathbb{P}}}\left[\Psi(f(\underline{X}),\underline{Z})\right],
\end{aligned}$$

and

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\lambda^+}h(\lambda) &= \rho^p + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\frac{\mathrm{d}}{\mathrm{d}\lambda^+}\Phi(\lambda;\widehat{X})\right] \\
&= \rho^p + \mathbb{E}_{\overline{\gamma}_1}\left[-\mathbb{E}_{\overline{\gamma}_2}\left[|\overline{Z}-\widehat{Z}|^p|(\overline{X},\widehat{X})\right] - |\overline{X}-\widehat{X}|^p\right] \\
&= \rho^p - \overline{d}, \\
\frac{\mathrm{d}}{\mathrm{d}\lambda^-}h(\lambda) &= \rho^p - \underline{d}.
\end{aligned}
$$

At $\lambda = \lambda^*$, h is minimized, so $\frac{\mathrm{d}}{\mathrm{d}\lambda^-}h(\lambda^*) \leq 0 \leq \frac{\mathrm{d}}{\mathrm{d}\lambda^+}h(\lambda^*)$. Therefore there exists $q^* \in [0,1]$, such that

$$
q^*\left(\rho^p - \overline{d}\right) + (1-q^*)\left(\rho^p - \underline{d}\right) = 0.
$$

Then if we denote $\gamma^* = q^*\overline{\gamma} + (1-q^*)\underline{\gamma}$, then

$$
\mathbb{E}_{((X,Z),(\widehat{X},\widehat{Z}))\sim\gamma^*}\left[|X-\widehat{X}|^p + |Z-\widehat{Z}|^p\right] = q^*\overline{d} + (1-q^*)\underline{d} = \rho^p.
$$

Therefore, $\mathbb{P}^* = \gamma^*_{(X,Z)} = q^*\overline{\mathbb{P}} + (1-q^*)\underline{\mathbb{P}}$ is feasible, and

$$
\mathbb{E}_{\mathbb{P}^*}[\Psi(f(X),Z)] = q^*\mathbb{E}_{\overline{\mathbb{P}}}[\Psi(f(\overline{X}),\overline{Z})] + (1-q^*)\mathbb{E}_{\overline{\mathbb{P}}}[\Psi(f(\overline{X}),\overline{Z})] = h(\lambda^*) = v_{\mathsf{D}}^f = v_{\mathsf{P}}^f
$$

it is optimal.

Note that this optimal solution is

$$
\mathbb{P}^* = \sum_{k=1}^{K}\sum_{i=1}^{n_k}\hat{p}_{ki}\left(q^*\delta_{(\overline{x}_k,\overline{z}_{ki})} + (1-q^*)\delta_{(\underline{x}_k,\underline{z}_{ki})}\right).
$$

Now we first consider the following linear optimization problem,

$$
\begin{cases}
\sup_{q_i}\mathbb{E}_{(X,Z)\sim\mathbb{P}}[\Psi(f(X),Z)] \\
\qquad \text{where } \mathbb{P} = \sum_{k=1}^{K}\sum_{i=1}^{n_k}\hat{p}_{ki}\left(q_i\delta_{(\overline{x}_k,\overline{z}_{ki})} + (1-q_i)\delta_{(\underline{x}_k,\underline{z}_{ki})}\right), \\
\text{s.t. } \mathbb{E}_{((X,Z),(\widehat{X},\widehat{Z}))\sim\gamma}\left[|X-\widehat{X}|^p + |Z-\widehat{Z}|^p\right] \leq \rho^p, \qquad 0 \leq q_i \leq 1, \\
\qquad \text{where } \gamma = \sum_{k=1}^{K}\sum_{i=1}^{n_k}\hat{p}_{ki}\left(q_i\delta_{((\overline{x}_k,\overline{z}_{ki}),(\widehat{x}_k,\widehat{z}_{ki}))} + (1-q_i)\delta_{((\underline{x}_k,\underline{z}_{ki}),(\widehat{x}_k,\widehat{z}_{ki}))}\right).
\end{cases}
$$

The feasible domain is not empty because $q_k = q^*$ gives a feasible solution $\mathbb{P}^*$. The constraints and the target function are all linear functions of q_k, so the inf can be attained at the vertices of the feasible domain, thus we can find k_0 such that $q_k = 1$ or 0 whenever $k \neq k_0$. So we have found another optimal solution

$$\mathbb{P} = \sum_{k \neq k_0} \sum_{i=1}^{n_k} \hat{p}_{ki} \delta_{(x_k^*, z_{ki}^*)} + \sum_{i=1}^{n_{k_0}} \hat{p}_{i_0 j} \left(q \delta_{(\overline{x}_{k_0}, \overline{z}_{k_0 i})} + (1-q) \delta_{(\underline{x}_{k_0}, \underline{z}_{k_0 i})} \right).$$

where $(x_k^*, z_{ki}^*) = (\overline{x}_k, \overline{z}_{ki})$ or $(\underline{x}_k, \underline{z}_{ki})$ depending only on k. Note that the marginal $\mathbb{P}_X$ is supported over at most $I + 1$ points.

B.4 Proofs for Section 2.3

Proof. Proof of Corollary 1. Since $\Psi(\cdot, z)$ is affine for each z, Ψ can be written as

$$\Psi(w, z) = \ell^z(w), \qquad \ell^z(w) = \beta^{z\top} w + b^z.$$

Here ℓ^z is an affine function with gradient $\beta^z \in \mathcal{D}^*$ and intercept $b^z \in \mathbb{R}$. Then

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}} \left[\Psi(w, \widehat{Z}) | \widehat{X} = \widehat{x}_k \right] = \frac{1}{\sum_{i=1}^{n_k} \hat{p}_{ki}} \sum_{i=1}^{n_k} \hat{p}_{ki} \Psi(w, \widehat{z}_{ki}) = \frac{1}{\sum_{i=1}^{n_k} \hat{p}_{ki}} \sum_{i=1}^{n_k} \hat{p}_{ki} \ell^{\widehat{z}_{ki}}(w)$$

Denote

$$\beta_k := \frac{1}{\sum_{i=1}^{n_k} \hat{p}_{ki}} \sum_{i=1}^{n_k} \hat{p}_{ki} \beta^{\widehat{z}_{ki}}, \qquad b_k := \frac{1}{\sum_{i=1}^{n_k} \hat{p}_{ki}} \sum_{i=1}^{n_k} \hat{p}_{ki} b^{\widehat{z}_{ki}},$$

and

$$\ell_k(w) := \frac{1}{\sum_{i=1}^{n_k} \hat{p}_{ki}} \sum_{i=1}^{n_k} \hat{p}_{ki} \ell^{\widehat{z}_{ki}}(w) = \beta_k^\top w + b_k, \tag{B.10}$$

which is an affine function of w. Therefore, $\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}} \left[\Psi(w, \widehat{Z}) | \widehat{X} = \widehat{x}_k \right] = \ell_k(w)$ is affine. We have

$$\sup_{x \in \mathcal{X}} \left\{ \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}} \left[\Psi(f(x), \widehat{Z}) | \widehat{X} = \widehat{x}_k \right] - \lambda \|x - \widehat{x}_k\|^p \right\} = \sup_{x \in \mathcal{X}} \left\{ \ell_k(f(x)) - \lambda \|x - \widehat{x}_k\|^p \right\}.$$

Suppose $f : \mathcal{X} \to \mathcal{D}$ is an affine decision rule, then $f(x) = B^\top x + \delta$, and

$$\ell_k(f(x)) - \ell_k(f(\widehat{x}_k)) = \beta_k^\top (f(x) - f(\widehat{x}_k)) = \beta_k^\top B^\top (x - \widehat{x}_k).$$

Thus the supremum over x can be computed explicitly as

$$\begin{aligned}\sup_{x\in\mathcal{X}} \{\ell_k(f(x)) - \lambda\|x - \widehat{x}_k\|^p\} &= \ell_k(f(\widehat{x}_k)) + \sup_{x\in\mathcal{X}} \left\{(B\beta_k)^\top (x - \widehat{x}_k) - \lambda\|x - \widehat{x}_k\|^p\right\} \\ &= \ell_k(f(\widehat{x}_k)) + \sup_{t\geq 0} \{|B\beta_k|\, t - \lambda t^p\}.\end{aligned}$$

Using notation introduced in (2.3),

$$\sup_{x\in\mathcal{X}} \{\ell_k(f(x)) - \lambda\|x - \widehat{x}_k\|^p\} = \ell_k(f(\widehat{x}_k)) + R_p\left(\lambda, |B\beta_k|\right),$$

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\sup_{x\in\mathcal{X}} \left\{\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\Psi(f(x), \widehat{Z})|\widehat{X}\right] - \lambda\|x - \widehat{X}\|^p\right\}\right] = \sum_{k=1}^{K} \left(\sum_{i=1}^{n_k} \hat{p}_{ki}\right) \left[\ell_k(f(\widehat{x}_k)) + R_p\left(\lambda, |B\beta_k|\right)\right].$$

Note that R_p is a convex function in λ and B, $\ell_k(f(\widehat{x}_k)) = \ell_k(B^\top\widehat{x}_k + \delta)$ is affine in B and δ, so the right hand side of the last expression is convex in λ and B as well. Hence (2.1) is a convex program:

$$\inf_{\lambda\geq 0, (B,\delta)\in\Theta} \left\{\lambda\rho^p + \sum_{k=1}^{K} \left(\sum_{i=1}^{n_i} \hat{p}_{ki}\right) \left[\ell_k(B^\top\widehat{x}_k + \delta)) + R_p\left(\lambda, |B\beta_k|\right)\right]\right\},$$

where ℓ_k is an affine function defined by (B.10) and R_p is a convex function defined by (2.3).

Proof. Proof of Corollary 2

We start with sup over z:

$$\sup_{z\in\mathcal{Z}} \left\{\Psi(w, z) - \lambda\|z - \widehat{z}_{ki}\|^2\right\} = \Psi(w, \widehat{z}_{ki}) + \sup_{z\in\mathcal{Z}} \left\{(A^\top w + \alpha)^\top (z - \widehat{z}_{ki}) - \lambda\|z - \widehat{z}_{ki}\|^2\right\}.$$

Note that the decision is

$$w = f(x) = f(\widehat{x}_k) + B^\top (x - \widehat{x}_k) := w_k + B^\top (x - \widehat{x}_k). \tag{B.11}$$

We introduce an auxillary variable y_k satisfying

$$y_k \geq (A^\top w_k + A^\top B^\top (x - \widehat{x}_k) + \alpha)^\top (z - \widehat{z}_{ki}) - \lambda \|z - \widehat{z}_{ki}\|^2.$$

This is equivalent to

$$\begin{pmatrix}(x - \widehat{x}_k)^\top & (z - \widehat{z}_{ki})^\top & 1\end{pmatrix} \begin{pmatrix} O & -\frac{1}{2}BA & O \\ -\frac{1}{2}(BA)^\top & \lambda \operatorname{Id} & -\frac{1}{2}(A^\top w_k + \alpha) \\ O & -\frac{1}{2}(A^\top w_k + \alpha)^\top & y_k \end{pmatrix} \begin{pmatrix} x - \widehat{x}_k \\ z - \widehat{z}_{ki} \\ 1 \end{pmatrix} \geq 0, \qquad \forall z \in \mathcal{Z}.$$

By the linearity of $\mathbf{\Psi}$ in z,

$$\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}} \left[\sup_{z \in \mathcal{Z}} \left\{ \mathbf{\Psi}(w, z) - \lambda \|z - \widehat{Z}\|^2 \right\} \mid \widehat{X} = \widehat{x}_k \right] = \mathbf{\Psi}(w, \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}[\widehat{Z} \mid \widehat{X} = \widehat{x}_k]) + y_k = \mathbf{\Psi}(w, \bar{z}_k) + y_k,$$

where we define $\bar{z}_k = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}[\widehat{Z} \mid \widehat{X} = \widehat{x}_k]$. Using the notations in (B.11), we can write

$$\mathbf{\Psi}(w, \bar{z}_k) = \mathbf{\Psi}(w_k + B^\top (x - \widehat{x}_k), \bar{z}_k) = \mathbf{\Psi}(w_k, \bar{z}_k) + (\beta + A\bar{z}_k)^\top B^\top (x - \widehat{x}_k) = \mathbf{\Psi}_k + (\beta + A\bar{z}_k)^\top B^\top (x - \widehat{x}_k),$$

where we denote $\mathbf{\Psi}_k = \mathbf{\Psi}(w_k, \bar{z}_k)$. Now we introduce another auxillary variable Y_k satisfying

$$Y_k + y_k \geq \mathbf{\Psi}_k + (\beta + A\bar{z}_k)^\top B^\top (x - \widehat{x}_k) + y_k - \lambda \|x - \widehat{x}_k\|^2.$$

That is, we need

$$\begin{pmatrix}(x - \widehat{x}_k)^\top & 1\end{pmatrix} \begin{pmatrix} \lambda \operatorname{Id} & -\frac{1}{2}B(\beta + A\bar{z}_k) \\ -\frac{1}{2}(\beta + A\bar{z}_k)^\top B^\top & Y_k - \mathbf{\Psi}_k \end{pmatrix} \begin{pmatrix} x - \widehat{x}_k \\ 1 \end{pmatrix} \geq 0, \qquad \forall x \in \mathcal{X}.$$

Thus we have transformed (D) into a conic programming problem:

$$\begin{aligned} \inf_{\substack{(B,\delta)\in\Theta \\ \lambda \geq 0, y_k \geq 0, Y_k \geq 0}} \quad & \lambda \rho^2 + \sum_{k=1}^{K} \hat{p}_k (y_k + Y_k) \\ \text{subject to} \quad & \begin{pmatrix} \lambda \operatorname{Id} & -\frac{1}{2}B(\beta + A\bar{z}_k) \\ -\frac{1}{2}(\beta + A\bar{z}_k)^\top B^\top & Y_k \end{pmatrix} \succeq 0 \\ & \begin{pmatrix} O & -\frac{1}{2}BA & O \\ -\frac{1}{2}(BA)^\top & \lambda \operatorname{Id} & -\frac{1}{2}(A^\top w_k + \alpha) \\ O & -\frac{1}{2}(A^\top w_k + \alpha)^\top & y_k - \mathbf{\Psi}_k \end{pmatrix} \succeq 0 \\ & w_k = B^\top \widehat{x}_k + \delta, \qquad \mathbf{\Psi}_k = \mathbf{\Psi}(w_k, \bar{z}_k) = (A\bar{z}_k + \beta)^\top (B^\top \widehat{x}_k + \delta) + \alpha^\top \bar{z}_k + b. \end{aligned}$$

Proof. Proof of Theorem 4. First, we show that $\cap_k I_k(x)$ is nonempty. To begin with, each $I_k(x)$ is nonempty, because the definition of ϕ_k implies

$$\varphi_k(y_k) \leq \phi_k \leq \lambda^* \|x - x_k\| + \phi_k,$$

so $y_k \in I_k(x)$. Note that each $I_k(x)$ is an interval, since it is the sub-level set of a convex function φ_k. To prove they have a nonempty intersection, it suffices to show they pairwise intersect. For instance, we show here that $I_1(x)$ and $I_2(x)$ intersect, by contradiction. Suppose I_1 and I_2 are disjoint. Since $y_1 \in I_1(x)$, $y_2 \in I_2(x)$, we know that I_1 and I_2 are disjoint if and only if we can find y_3 in between y_1 and y_2 outside both intervals. This implies that

$$\begin{aligned}
&\varphi_1(y_3) > \lambda^* \|x - x_3\| + \phi_1 \geq \lambda^* \|x - x_1\| + \varphi_1(y_1),\\
&\varphi_1(y_3) > \lambda^* \|x - x_3\| + \phi_1 \geq \lambda^* \|x - x_1\| + \varphi_1(y_2) - \lambda^* \|x_1 - x_2\|,\\
&\varphi_2(y_3) > \lambda^* \|x - x_3\| + \phi_2 \geq \lambda^* \|x - x_2\| + \varphi_2(y_2),\\
&\varphi_2(y_3) > \lambda^* \|x - x_3\| + \phi_2 \geq \lambda^* \|x - x_2\| + \varphi_2(y_1) - \lambda^* \|x_1 - x_2\|.
\end{aligned}$$

Since y_3 is between y_1 and y_2, we can find $\alpha, \beta \in [0, 1]$ with $\alpha + \beta = 1$ and $y_3 = \alpha y_1 + \beta y_2$. By multiplying the first/fourth inequality with α and the second/third inequality with β then taking the sum, we have

$$\begin{aligned}
(\varphi_1 + \varphi_2)(y_3) &> \lambda^* (\|x - x_1\| + \|x - x_2\|) + \alpha(\varphi_1 + \varphi_2)(y_1) + \beta(\varphi_1 + \varphi_2)(y_2) - \lambda^* \|x_1 - x_2\|\\
&\geq \alpha(\varphi_1 + \varphi_2)(y_1) + \beta(\varphi_1 + \varphi_2)(y_2),
\end{aligned}$$

using the triangle inequality. However, this contradicts with the convexity of $\varphi_1 + \varphi_2$.

Next, we prove that any decision rule in the intersection $\cap_k I_k$ is optimal. For every $f \in \mathcal{F}$, let $\widehat{w} = f|_{\widehat{\mathcal{X}}} \in \hat{\mathcal{F}}$ be the restriction of f on the set $\widehat{X}$, then

$$\begin{aligned}
&\inf_{\lambda \geq 0} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\sup_{x \in \mathcal{X}} \left\{ \varphi(f(x); \lambda, \widehat{X}) - \lambda \|x - \widehat{X}\| \right\} \right] \right\} \\
&\geq \inf_{\lambda \geq 0} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\max_{x \in \widehat{\mathcal{X}}} \left\{ \varphi(f(x); \lambda, \widehat{X}) - \lambda \|x - \widehat{X}\| \right\} \right] \right\} \\
&= \inf_{\lambda \geq 0} \left\{ \lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\max_{1 \leq k \leq K} \left\{ \varphi(\widehat{w}(x_k); \lambda, \widehat{X}) - \lambda \|x_k - \widehat{X}\| \right\} \right] \right\} \geq v_{\hat{\mathsf{D}}}.
\end{aligned} \tag{B.12}$$

By taking the infimum over $f \in \mathcal{F}$ we would have $v_{\mathsf{D}} \geq v_{\hat{\mathsf{D}}}$. On the other hand, for the minimizer λ^* and $\widehat{w}^* \in \hat{\mathcal{F}}$ of (2.6), let $f \in \mathcal{F}$ be an extension in $\cap_k \boldsymbol{I}_k(x)$, then for every x we have

$$\varphi_k(f(x)) - \lambda^* \|x - \widehat{x}\| \leq \max_{1 \leq k \leq K} \left\{ \varphi(\widehat{w}(\widehat{x}_k); \lambda^*, \widehat{x}) - \lambda^* \|x_k - \widehat{x}\| \right\}.$$

Therefore,

$$\begin{aligned}
&\lambda^* \rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\max_{1 \leq k \leq K} \left\{ \varphi(\widehat{w}(x_k); \lambda^*, \widehat{X}) - \lambda^* \|x_k - \widehat{X}\| \right\} \right] \\
&\geq \lambda^* \rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}} \left[\sup_{x \in \mathcal{X}} \left\{ \varphi(f(x); \lambda^*, \widehat{X}) - \lambda^* \|x - \widehat{X}\| \right\} \right] \geq v_{\mathsf{D}}.
\end{aligned}$$

Thus $v_{\mathsf{D}} = v_{\hat{\mathsf{D}}}$.

Finally, we show the necessity of the interval condition. Suppose $f^* \in \mathcal{F}$ is an optimal policy to the problem (2.4) with optimal dual value λ^*. By (B.12), λ^* and the restriction $\hat{f}^* = f|_{\widehat{\mathcal{X}}} \in \hat{\mathcal{F}}$ are also an optimal dual value and an optimal policy to the problem (2.6). To show that $f^*(x) \in \cap_k \boldsymbol{I}_k(x)$, we prove by contradiction. Suppose for some $x \in \mathcal{X}$ and some $k \in [K]$, $f^*(x) \notin \boldsymbol{I}_k(x)$. This means

$$\varphi(f^*(x); \lambda^*, \widehat{x}_k) = \varphi_k(f^*(x)) > \lambda^* \|x - \widehat{x}_k\| + \phi_k = \lambda^* \|x - \widehat{x}_k\| + \max_j \left\{ \varphi_k(y_j) - \lambda^* \|\widehat{x}_k, \widehat{x}_j - \| \right\}.$$

That is, there exists $k \in [K]$ such that for all $j \in [K]$,

$$\varphi(f^*(x); \lambda^*, \widehat{x}_k) - \lambda^* \|x - \widehat{x}_k\| > \varphi(f^*(\widehat{x}_j); \lambda^*, \widehat{x}_k) - \lambda^* \|\widehat{x}_k - \widehat{x}_j\|.$$

Then

$$\begin{aligned} v_{\mathsf{D}} &= \lambda^*\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\sup_{x\in\mathcal{X}}\left\{\varphi(f^*(x);\lambda^*,\widehat{X}) - \lambda^*\|x-\widehat{X}\|\right\}\right] \\ &> \lambda^*\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\max_{j\in[K]}\left\{\varphi(f^*(\widehat{x}_j);\lambda^*,\widehat{X}) - \lambda^*\|\widehat{x}_j-\widehat{X}\|\right\}\right] \geq v_{\hat{\mathsf{D}}}, \end{aligned}$$

which contradicts with $v_{\mathsf{D}} = v_{\hat{\mathsf{D}}}$. Therefore, we must have $f^*(x) \in \cap_k \boldsymbol{I}_k(x)$ for all $x \in \mathcal{X}$, which completes the proof of the theorem. □

B.5 Proofs for Examples in Section 2.3

Proof. Proof of Example 7. Since f is real-valued and $\boldsymbol{\Psi}$ is convex in w, we use Theorem 4, so it has the following reformulation

$$\inf_{\substack{\widehat{w}:\widehat{\mathcal{X}}\to\mathbb{R}\\ \lambda\geq 0}}\left\{\lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\max_{1\leq k\leq K}\left\{\varphi(\hat{f}(\widehat{x}_k);\lambda,\widehat{X}) - \lambda\|\widehat{x}_k-\widehat{X}\|\right\}\right]\right\}$$

with

$$\varphi(w;\lambda;\widehat{x}) = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}}\left\{|w-z| - \lambda\|z-\widehat{Z}\|\right\}|\widehat{X}=\widehat{x}\right].$$

For any $\lambda < 1$, the supremum over z is infinite, hence $\varphi(w;\lambda,\hat{x}) = \infty$. For $\lambda \geq 1$, the supremum is attained at $z = \hat{Z}$, so

$$\varphi(w;\lambda;\widehat{x}) = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[|w-\widehat{Z}|\,\middle|\,\widehat{X}=\widehat{x}\right] + \infty\mathbf{1}_{\{\lambda<1\}}.$$

Thus we reach the following reformulation,

$$\inf_{\substack{\widehat{w}:\widehat{\mathcal{X}}\to\mathbb{R}\\ \lambda\geq 1}}\left\{\lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\max_{1\leq k\leq K}\left\{\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[|\hat{f}(\widehat{x}_k)-\widehat{Z}|\,\middle|\,\widehat{X}=\widehat{x}\right] - \lambda\|\widehat{x}_k-\widehat{X}\|\right\}\right]\right\}$$

This can be transformed into a linear programming problem

$$\inf_{w_k,\lambda}\lambda\rho + \frac{1}{n}\sum_{k=1}^{K} c_j$$

$$s.t.\begin{cases} c_j \geq \sum_{i=1}^{n_j} c_{kji} - \lambda n_j\|\widehat{x}_k - \widehat{x}_j\|, \forall j,k \\ c_{kji} \geq w_k - \widehat{z}_{ji}, \forall k,j,i \\ c_{kji} \geq \widehat{z}_{ji} - w_k, \forall k,j,i \\ \lambda \geq 1 \end{cases}$$

□

Proof. Proof of Example 8. Recall that the problem could be reformulated as

$$\inf_{\substack{\widehat{w}:\widehat{\mathcal{X}}\to\mathbb{R}\\ \lambda\geq 0}} \left\{\lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\max_{1\leq k\leq K}\left\{\varphi(\hat{f}(\widehat{x}_k);\lambda,\widehat{X}) - \lambda\|\widehat{x}_k - \widehat{X}\|\right\}\right]\right\}.$$

where

$$\varphi(w;\lambda;\widehat{x}) = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\sup_{z\in\mathcal{Z}}\left\{-wz^\top\begin{pmatrix}w\\1\end{pmatrix} - \lambda\|z-\widehat{Z}\|\right\} \mid \widehat{X}=\widehat{x}\right].$$

Since the metric on Z is chosen to be infinite, the supremum over $z\in\mathcal{Z}$ is attained at $z=\widehat{Z}$, thus

$$\varphi(w;\lambda;\widehat{x}) = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[-w\widehat{Z}^\top\begin{pmatrix}w\\1\end{pmatrix} \mid \widehat{X}=\widehat{x}\right] = -w\mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\widehat{Z}\mid\widehat{X}=\widehat{x}\right]^\top\begin{pmatrix}w\\1\end{pmatrix}.$$

Denote $\bar{z}_k = \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{Z}|\widehat{X}}}\left[\widehat{Z}\mid\widehat{X}=\widehat{x}_k\right]$. Then the problem is reformulated as

$$\begin{aligned}
&\inf_{\substack{\widehat{w}:\widehat{\mathcal{X}}\to\mathbb{R}\\ \lambda\geq 0}} \left\{\lambda\rho + \mathbb{E}_{\widehat{\mathbb{P}}_{\widehat{X}}}\left[\max_{1\leq k\leq K}\left\{\varphi(\hat{f}(\widehat{x}_k);\lambda,\widehat{X}) - \lambda\|\widehat{x}_k - \widehat{X}\|\right\}\right]\right\}\\
&= \inf_{\substack{w_k\geq 0\\ \lambda\geq 0}}\left\{\lambda\rho + \sum_{j\in[K]}\hat{p}_j\left[\max_{1\leq k\leq K}\left\{\varphi(w_k;\lambda,\widehat{x}_j) - \lambda\|\widehat{x}_k-\widehat{x}_j\|\right\}\right]\right\}\\
&= \inf_{\substack{w_k\geq 0\\ \lambda\geq 0}}\left\{\lambda\rho + \sum_{j\in[K]}\hat{p}_j\left[\max_{1\leq k\leq K}\left\{-w_k\bar{z}_k^\top\begin{pmatrix}w_k\\1\end{pmatrix} - \lambda\|\widehat{x}_k-\widehat{x}_j\|\right\}\right]\right\},
\end{aligned}$$

which can be written as the following quadratic programing:

$$\begin{aligned}
\underset{\substack{w_k\geq 0\\ \lambda\geq 0}}{\text{minimize}}\quad & \lambda\rho + \sum_{j\in[K]}\hat{p}_jc_j\\
\text{subject to}\quad & c_j + w_k\bar{z}_k^\top\begin{pmatrix}w_k\\1\end{pmatrix} + \lambda\|\widehat{x}_k-\widehat{x}_j\| \geq 0, \qquad \forall k\in[K].
\end{aligned}$$

□

Proof. Proof of Example 9. By Theorem 2, the dual problem can be written as

$$\inf_{B\in\mathscr{B},\lambda\geq 0}\left\{\lambda\rho^2+\mathbb{E}_{\widehat{X}}\left[\sup_x\left\{\mathbb{E}_{\widehat{Z}}\left[\sup_z\left\{(Bx+w_0)^\top z-\lambda\|z-\widehat{Z}\|^2\right\}\right]-\lambda\|x-\widehat{X}\|^2\right\}\right]\right\}.$$

For convenience, make a change $B\rightarrow\frac{1}{2}\lambda B$ and $w_0=\frac{1}{2}\lambda w_1$, thus we have

$$\inf_{\lambda\geq 0,B\in\mathscr{B}}\left\{\lambda\rho^2+\mathbb{E}_{\widehat{X}}\left[\sup_x\left\{\mathbb{E}_{\widehat{Z}}\left[\sup_z\left\{2(Bx+w_1)^\top z-\|z-\widehat{Z}\|^2\right\}\right]-\|x-\widehat{X}\|^2\right\}\right]\right\}.$$

Note that the supremum over z can be simplified to

$$\begin{aligned}\sup_z\left\{2(Bx+w_1)^\top z-\|z-\widehat{Z}\|^2\right\}&=\sup_z\left\{-\|z\|^2+2(\widehat{Z}+Bx+w_1)^\top z-\|\widehat{Z}\|^2\right\}\\&=-\|\widehat{Z}\|^2+\|\widehat{Z}+Bx+w_1\|^2\\&=\left\langle Bx+w_1,Bx+w_1+2\widehat{Z}\right\rangle\end{aligned}$$

Taking the expectation over $\widehat{Z}$ yields

$$\mathbb{E}_{\widehat{Z}|\widehat{X}}\left[\sup_z\left\{2(Bx+w_1)^\top z-\|z-\widehat{Z}\|^2\right\}\right]=\left\langle Bx+w_1,Bx+w_1+2\mathbb{E}_{\widehat{Z}|\widehat{X}}\left[\widehat{Z}|\widehat{X}\right]\right\rangle.$$

We denote $\widehat{z}=\mathbb{E}[\widehat{Z}|\widehat{X}]$ for convenience. Next we need to calculate

$$\begin{aligned}&\sup_x\left\{\langle Bx+w_1,Bx+w_1+2\widehat{z}\rangle-\|x-\widehat{X}\|^2\right\}\\&=\sup_x\left\{x^\top\left(B^\top B-\mathrm{Id}\right)x+2(B^\top(\widehat{z}+w_1)-\widehat{X})^\top x-\|\widehat{X}\|^2+\langle w_1,w_1+2\widehat{z}\rangle\right\}.\end{aligned}$$

Observe that for the supremum to be finite, we need $-\Sigma:=B^\top B-\mathrm{Id}\leq 0$, that is, the spectral norm $\sigma(B^\top B)\leq 1$. Denote $b:=B^\top(\widehat{z}+w_1)-\widehat{X}$, $c:=-\|\widehat{X}\|^2+\langle w_1,w_1+2\widehat{z}\rangle$, we have

$$\sup_x\left\{-x^\top\Sigma x+2b^\top x+c\right\}=c+b^\top\Sigma^{-1}b.$$

Taking the expectation,

$$\begin{aligned}
\mathbb{E}[b^\top \Sigma^{-1} b] &= \mathbb{E}\left[(-\widehat{X} + B^\top \widehat{z} + B^\top w_1)^\top \Sigma^{-1}(-\widehat{X} + B^\top \widehat{z} + B^\top w_1)\right] \\
&= \mathbb{E}\left[\begin{pmatrix} \widehat{X}^\top & \widehat{Z}^\top & w_1^\top \end{pmatrix} \begin{pmatrix} -\operatorname{Id} \\ B \\ B \end{pmatrix} \Sigma^{-1} \begin{pmatrix} -\operatorname{Id} & B^\top & B^\top \end{pmatrix} \begin{pmatrix} \widehat{X} \\ \widehat{Z} \\ w_1 \end{pmatrix}\right] \\
&= \mathbb{E}\left[\begin{pmatrix} \widehat{X} \\ \widehat{z} \\ w_1 \end{pmatrix} \begin{pmatrix} \widehat{X}^\top & \widehat{z}^\top & w_1^\top \end{pmatrix}\right] : \left(\begin{pmatrix} -\operatorname{Id} \\ B \\ B \end{pmatrix} \Sigma^{-1} \begin{pmatrix} -\operatorname{Id} & B^\top & B^\top \end{pmatrix}\right) \\
&= \boldsymbol{C} : \boldsymbol{D},
\end{aligned}$$

where $\boldsymbol{C} : \boldsymbol{D} = \operatorname{tr}(\boldsymbol{C}\boldsymbol{D})$ and we denote

$$\begin{aligned}
\boldsymbol{C} &= \begin{pmatrix} \mathbb{C}\mathrm{ov}(\widehat{X}) & \mathbb{E}[\widehat{X}\widehat{Z}^\top] & \mathbb{E}[\widehat{X}]w_1^\top \\ \mathbb{E}[\widehat{Z}\widehat{X}^\top] & \mathbb{C}\mathrm{ov}(\mathbb{E}[\widehat{Z}|\widehat{X}]) & \mathbb{E}[\widehat{Z}]w_1^\top \\ w_1\mathbb{E}[\widehat{X}]^\top & w_1\mathbb{E}[\widehat{Z}]^\top & w_1 w_1^\top \end{pmatrix} \\
&= \operatorname{diag}\left(\operatorname{Id}, \operatorname{Id}, \frac{2}{\lambda}\operatorname{Id}\right) \mathbb{C}\mathrm{ov}\begin{pmatrix} \widehat{X} \\ \widehat{Z} \\ w_0 \end{pmatrix} \operatorname{diag}\left(\operatorname{Id}, \operatorname{Id}, \frac{2}{\lambda}\operatorname{Id}\right), \\
\boldsymbol{D} &= \begin{pmatrix} -\operatorname{Id} \\ B \\ B \end{pmatrix} \Sigma^{-1} \begin{pmatrix} -\operatorname{Id} & B^\top & B^\top \end{pmatrix}.
\end{aligned}$$

If we denote $S_1^2 = \mathbb{C}\mathrm{ov}\begin{pmatrix} \widehat{X} \\ \widehat{Z} \\ w_0 \end{pmatrix}$, then

$$\boldsymbol{C} : \boldsymbol{D} = \operatorname{tr}(\boldsymbol{C}\boldsymbol{D}) = \operatorname{tr}\left(S_1 \operatorname{diag}\left(\operatorname{Id}, \operatorname{Id}, \frac{2}{\lambda}\operatorname{Id}\right) \begin{pmatrix} -\operatorname{Id} \\ B \\ B \end{pmatrix} \Sigma^{-1} \begin{pmatrix} -\operatorname{Id} & B^\top & B^\top \end{pmatrix} \operatorname{diag}\left(\operatorname{Id}, \operatorname{Id}, \frac{2}{\lambda}\operatorname{Id}\right) S_1\right) = \operatorname{tr}(S_2^\top \Sigma^{-1} S$$

where $S_2 = \begin{pmatrix} -\operatorname{Id} & B^\top & \frac{2}{\lambda}B^\top \end{pmatrix} S_1$. As for the c term, we denote

$$C' := \mathbb{E}[c] = -\mathbb{E}[\|\widehat{X}\|^2] + \left\langle w_1, w_1 + 2\mathbb{E}[\widehat{Z}] \right\rangle = \boldsymbol{C} : \begin{pmatrix} -\operatorname{Id} & O & O \\ O & O & \operatorname{Id} \\ 0 & \operatorname{Id} & \operatorname{Id} \end{pmatrix}.$$

Therefore we need to compute

$$\inf_{\substack{\lambda \geq 0 \\ B \in \mathscr{B} \\ \sigma(B^\top B) \leq 1}} \lambda\left(\rho^2 + \operatorname{tr}\left(S_2^\top \Sigma^{-1} S_2\right) + C'\right).$$

Given Σ is positive definite, by Schur decomposition we have

$$\begin{aligned}\rho^2 + \operatorname{tr}\left(S_2^\top \Sigma^{-1} S_2\right) + C' &= \rho^2 + C' + \inf_{t \in \mathbb{R}} \left\{ t : t \geq \operatorname{tr}(S_2^\perp \Sigma^{-1} S_2) \right\} \\ &= \rho^2 + C' + \inf \left\{ \operatorname{tr}(Y) : \begin{pmatrix} \Sigma & S_2 \\ S_2^\top & Y \end{pmatrix} \geq 0 \right\}.\end{aligned}$$

Hence the original problem is reduced to

$$\begin{aligned}\inf_{\lambda \geq 0, B \in \mathscr{B}} \quad & \lambda \left(\rho^2 - \mathbb{E}[\|\widehat{X}\|^2] \right) + 2 w_0^\top \mathbb{E}[\widehat{Z}] + \frac{4}{\lambda} w_0^\top w_0 + \operatorname{tr}(Y) \\ \text{s.t.} \quad & \begin{cases} B^\top B \leq \operatorname{Id} \\ \begin{pmatrix} \lambda(\operatorname{Id} - B^\top B) & \begin{pmatrix} -\lambda \operatorname{Id} & \lambda B^\top & 2B^\top \end{pmatrix} S_1 \\ S_1^\top \begin{pmatrix} -\lambda \operatorname{Id} \\ \lambda B \\ 2B \end{pmatrix} & Y \end{pmatrix} \geq 0 \end{cases}\end{aligned}$$

□

Appendix C: Appendices to *"Dynamically Information Acquisition and Optimal Decision Making"*

C.1 Additional Proofs

Proof of Proposition 8. We establish monotonicity in σ; similar arguments may be used for (i) and (ii). For $\sigma_2 > \sigma_1$, let V_1, V_2 be viscosity solutions to

$$\min\left(\rho V_1(q) - \frac{1}{2}\left(\frac{h-l}{\sigma_1}\right)^2 q^2(1-q)^2 V_1''(q) + C_I(q), V_1(q) - G(q)\right) = 0, \tag{C.1}$$

$$\min\left(\rho V_2(q) - \frac{1}{2}\left(\frac{h-l}{\sigma_2}\right)^2 q^2(1-q)^2 V_2''(q) + C_I(q), V_2(q) - G(q)\right) = 0. \tag{C.2}$$

We show that V_2 is a viscosity subsolution to (C.1). To this end, let $x \in (0,1)$ and consider a test function $\varphi \in \mathcal{C}^2(0,1)$ such that

$$(V_2 - \varphi)(x) = \max_{q\in(0,1)} (V_2 - \varphi)(q) = 0.$$

Since V_2 is a viscosity subsolution to (C.2), the test function φ must satisfy

$$\min\left(\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma_2}\right)^2 x^2(1-x)^2\varphi''(x) + C_I(x), \varphi(x) - G(x)\right) \leq 0.$$

Therefore, at least one of the following situation holds:

(I) $\varphi(x) - G(x) \leq 0$.

$\varphi(x) - G(x) > 0$ and $\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma_2}\right)^2 x^2(1-x)^2\varphi''(x) + C_I(x) \leq 0$. This implies $\varphi''(x) > 0$, and consequently

$$\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma_1}\right)^2 x^2(1-x)^2\varphi''(x) + C_I(x) < \rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma_2}\right)^2 x^2(1-x)^2\varphi''(x) + C_I(x) \leq 0.$$

Combining that above inequalities and using Lemma 4, we easily conclude.

Proof of Proposition 9. We only show the limiting case of σ as the other two cases follow similarly. Recall that V is the viscosity solution to

$$\min\left(\rho V(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V''(q) + C_I(q), V(q) - G(q)\right) = 0.$$

We construct a suitable $\mathcal{C}^1(0,1)$ supersolution. For this, let

$$U(q) := \begin{cases} \mu & q \in [0, r] \\ \mu + M(q-r)^2 & q \in [r, p] \\ qh + (1-q)l & q \in [p, 1], \end{cases}$$

for some $r \in (0, \hat{p})$, $p \in (\hat{p}, 1)$ and $M > 0$. To have $U \in \mathcal{C}^1(0,1)$, we need

$$U(p) = \mu + M(p-r)^2 = ph + (1-p)l, \tag{C.3}$$

$$U'(p) = 2M(p-r) = h - l. \tag{C.4}$$

Since $U = G$ in $[0, r] \cup [p, 1]$, it suffices to verify that

$$\rho U(q) - \left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 M + C_I(q) > 0, \quad q \in [r, p].$$

For any fixed M, r, p, the above inequality will hold as long as ρ, C_I, or σ are sufficiently large. By Lemma 4, we have $r \leq \underline{q} \leq \overline{q} \leq p$.

We finish the proof by showing that r and p can be arbitrarily close to $\hat{p}$ if we choose M sufficiently large. Plugging (C.4) into (C.3) yields

$$\mu + \frac{(h-l)^2}{4M} = hp + l(1-p) = \mu + (h-l)(p - \hat{p}) \quad \Rightarrow \quad p = \hat{p} + \frac{h-l}{4M}.$$

and thus $p = \hat{p} + \frac{h-l}{4M}$. Together with (C.4) we deduce that $r = \hat{p} - \dfrac{h-l}{4M}$. Hence $p \to \hat{p}$ and $r \to \hat{p}$, as $M \to \infty$. □

Proof of Proposition 10. For $\mu_2 > \mu_1$, let

$$G_1(q) = \max\left(\mu_1, qh + (1-q)l\right) \quad \text{and} \quad G_2(q) = \max\left(\mu_2, qh + (1-q)l\right).$$

Then $G_2 \geq G_1 \geq G_2 - \mu_2 + \mu_1$. Let V_1, V_2 be the viscosity solutions to

$$\min\left(\rho V_1(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V_1''(q) + C_I(q), V_1(q) - G_1(q)\right) = 0, \tag{C.5}$$

$$\min\left(\rho V_2(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 V_2''(q) + C_I(q), V_2(q) - G_2(q)\right) = 0. \tag{C.6}$$

We first show V_1 is a viscosity subsolution to (C.6). Take $x \in (0,1)$ and a test function $\varphi \in C^2(0,1)$ such that

$$(V_1 - \varphi)(x) = \max_{q\in(0,1)} (V_1 - \varphi)(q) = 0.$$

Since V_1 is the viscosity solution to (C.5), it is also a viscosity subsolution to (C.5). Then, the test function φ satisfies

$$\min\left(\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 x^2(1-x)^2\varphi''(x) + C_I(x), \varphi(x) - G_1(x)\right) \leq 0.$$

Since $G_1(x) \leq G_2(x)$, we also have

$$\min\left(\rho\varphi(x) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 x^2(1-x)^2\varphi''(x) + C_I(x), \varphi(x) - G_2(x)\right) \leq 0,$$

which implies that V_1 is a viscosity subsolution to (C.5). By comparison, $V_1 \leq V_2$.

Next, we show that the function $V_2 - \mu_2 + \mu_1$ is a viscosity subsolution to (C.5). For this, let $y \in (0,1)$ and consider a test function $\psi \in C^2(0,1)$, such that

$$(V_2 - \mu_2 + \mu_1 - \psi)(y) = \max_{q\in(0,1)} (V_2 - \mu_2 + \mu_1 - \psi)(q) = 0.$$

Then $\hat{\psi} := \psi + \mu_2 - \mu_1$ satsifies $\hat{\psi} \in C^2(0,1)$ and

$$(V_2 - \hat{\psi})(y) = \max_{q\in(0,1)} (V_2 - \hat{\psi})(q) = 0.$$

Since V_2 is the viscosity solution to (C.6), it is also a viscosity subsolution to (C.6). Therefore, the test function $\hat{\psi}$ satisfies

$$\min\left(\rho\hat{\psi}(y) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 y^2(1-y)^2\hat{\psi}''(y) + C_I(y), \hat{\psi}(y) - G_2(y)\right) \leq 0.$$

Since $\hat{\psi} \geq \psi$, $\hat{\psi}'' = \psi''$, and $\hat{\psi} - G_2 = \psi + \mu_2 - \mu_1 - G_2 \geq \psi - G_1$, we have

$$\min\left(\rho\psi(y) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 y^2(1-y)^2\psi''(y) + C_I(y), \psi(y) - G_1(y)\right) \leq 0.$$

Therefore, $V_2 - \mu_2 + \mu_1$ is a viscosity subsolution to (C.5) and, thus, by comparison results, $V_2 - \mu_2 + \mu_1 \leq V_1$.

So far we have shown that

$$V_2(q) - \mu_2 + \mu_1 \leq V_1(q) \leq V_2(q) \qquad \text{for } q \in [0,1].$$

By Theorem 6 we deduce that there exist two pairs of cutoff points, $\left(\underline{q}_1, \overline{q}_1\right)$ and $\left(\underline{q}_2, \overline{q}_2\right)$ such that

$$\begin{cases} V_1(q) = \mu_1 & q \in [0, \underline{q}_1] \\ V_1(q) > G_1(q) & q \in (\underline{q}_1, \overline{q}_1) \\ V_1(q) = qh + (1-q)l & q \in [\overline{q}_1, 1] \end{cases}, \qquad \begin{cases} V_2(q) = \mu_2 & q \in [0, \underline{q}_2] \\ V_2(q) > G_2(q) & q \in (\underline{q}_2, \overline{q}_2) \\ V_2(q) = qh + (1-q)l & q \in [\overline{q}_2, 1] \end{cases}.$$

To compare $\overline{q}_1$ and $\overline{q}_2$, notice that

$$V_2(q) \geq V_1(q) \geq qh + (1-q)l.$$

For any $q \in [0,1]$ such that $V_2(q) = qh + (1-q)l$, we also have $V_1(q) = qh + (1-q)l$. Therefore,

$$[\overline{q}_2, 1] = \{q \in [0,1] : V_2(q) = qh+(1-q)l\} \subseteq \{q \in [0,1] : V_1(q) = qh+(1-q)l\} = [\overline{q}_1, 1],$$

which yields $\overline{q}_2 \geq \overline{q}_1$. To compare $\underline{q}_1$ and $\underline{q}_2$, notice that the inequality $V_2(q) - \mu_2 + \mu_1 \leq V_1(q)$ yields

$$0 \leq V_2(q) - \mu_2 \leq V_1(q) - \mu_1.$$

For any $q \in [0,1]$ such that $V_1(q) = \mu_1$, we also have $V_2(q) = \mu_2$. Therefore,

$$[0, \underline{q}_1] = \{q \in [0,1] : V_1(q) = \mu_1\} \subseteq \{q \in [0,1] : V_2(q) = \mu_2\} = [0, \underline{q}_2],$$

which yields $\underline{q}_2 \geq \underline{q}_1$. In conclusion, for $\mu_2 > \mu_1$, we have $\underline{q}_2 \geq \underline{q}_1$ and $\overline{q}_2 \geq \overline{q}_1$.

□

Proof of Proposition 11. We first prove **i)**. Note that $\hat{p} = \frac{\mu - l}{h - l} \downarrow 0$ when $l \uparrow \mu$. In this case, $\underline{q} \to 0$ as $\hat{p} \to 0$. Next, we construct a convex supersolution $U \in \mathcal{C}([0,1])$,

$$U(q) := \begin{cases} \mu & q = 0 \\ qh + (1-q)l + M(p-q)^2 & 0 < q \leq p \\ qh + (1-q)l & p < q \leq 1 \end{cases},$$

for some $M > 0$ and $p \in (\hat{p}, 1)$ to be determined in the sequel. Continuity at $q = 0$ requires M and p satisfy $l + p^2 M = \mu$, and thus

$$p = \sqrt{\frac{\mu - l}{M}},$$

where M will be chosen independently of l.

We verify that U is convex. Since $U''(q) = 2M > 0$, $q \in [0, p)$, and U is affine on $[p, 1]$, it suffices to compute the left derivative of U at p. We have

$$U'_{-}(p) = (h - l) + 2M(p - p) = h - l.$$

Therefore, U is convex in $[0, 1]$ and piecewise smooth.

We now verify that U is a supersolution. Notice that $U = G$, for $q = 0$ and $q \in [p, 1]$. To show U is a supersolution, it suffices to show that, for $q \in (0, p)$, it holds that

$$f(q) := \rho U(q) + C_I(q) - \frac{1}{2}\left(\frac{h - l}{\sigma}\right)^2 q^2(1-q)^2 U''(q) \geq 0.$$

Notice that

$$f(q) \geq \rho\mu + C_I(q) - M\frac{h^2}{16\sigma^2}.$$

since $U(q) \geq \mu$, $q^2(1-q)^2 \leq \frac{1}{16}$, for $q \in [0,1]$. We denote $\check{c} = \min C_I(q)$, $q \in [0,1]$. By taking $M = \frac{16(\rho\mu+\check{c})\sigma^2}{h^2}$, we have $f(q) \geq 0$.

In conclusion, by taking $p = \sqrt{\frac{\mu-l}{M}}$ and $M = \frac{16(\rho\mu+\check{c})\sigma^2}{h^2}$, we have that U is a supersolution and $U(q) > G(q)$, $q \in (0,p)$. Then, by comparison principle, we obtain $V(q) \leq U(q)$, $q \in (0,p)$ and, hence, $p \geq \overline{q}$. As $l \to \mu$, we have $\underline{q} \to 0$ and $\overline{q} \to 0$ because $\hat{p} \to 0$ and $p \to 0$.

We show **ii).** We construct a convex subsolution $\overline{U} \in \mathcal{C}([0,1])$ with $\overline{U} = \max(\mu, U)$.

For some positive constants $m, M > 0, r \in \left(0, \frac{1}{2}\right]$ and $p \in \left(\frac{1}{2}, 1\right)$ to be determined, we define

$$U(q) := \int_q^1 \Phi(t)dt + qh + (1-q)l$$

(shown in Figure C.1), where

$$U''(q) = \varphi(q) = \begin{cases} M & 0 \leq q \leq r \\ m & r < q \leq p \\ 0 & p < q \leq 1 \end{cases}, \quad \Phi(q) = \int_q^1 \varphi(t)dt = \begin{cases} m(p-r) + M(r-q) & 0 \leq q \leq r \\ m(p-q) & r < q \leq p \\ 0 & p < q \leq 1 \end{cases}.$$

We choose m, M, r independently of h. We claim that by a proper choice of m, M, p, r, U will satisfy the following properties.

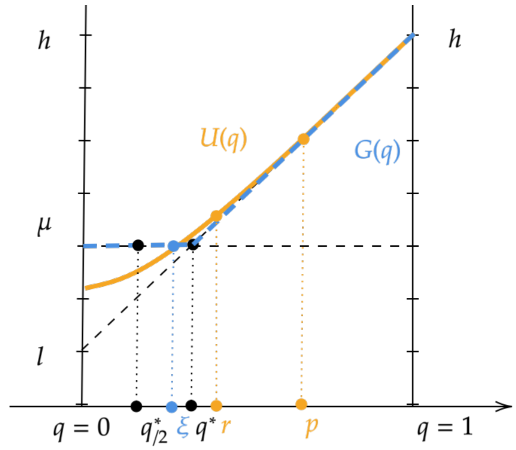


Figure C.1: Illustration of $U(q)$ (orange solid line) and $G(q) = \max\{\mu, hq + l(1-q)\}$ (blue dotted line).

First, U is increasing and convex in $[0,1]$, and $U \in \mathcal{C}^{1,1}([0,1])$.

First, U is convex since $U'(q) = -\Phi(q) + h - l$ is increasing in q. Moreover, Φ is Lipschitz, and U is $\mathcal{C}([0,1])$ since φ is bounded by $\max(m, M)$. Notice that

$$U'(0) = -\Phi(0) + h - l = -m(p-r) - Mr + h - l > -M + h - l,$$

where M will be chosen independent of h. We have $U'(0) > 0$, for h is sufficiently large.

We determine ξ at which $U(\xi) = \mu$. By direct calculation,

$$\begin{aligned} U\left(\frac{\hat{p}}{2}\right) &= \int_{\frac{\hat{p}}{2}}^{1} \Phi(t)dt + \frac{\hat{p}}{2}h + \left(1 - \frac{\hat{p}}{2}\right)l \\ &= \int_{\frac{\hat{p}}{2}}^{1} \Phi(t)dt + \hat{p}h + (1 - \hat{p}l) - \frac{\hat{p}}{2}(h-l) \\ &\le \int_0^1 \Phi(t)dt + \mu - \frac{1}{2}(\mu - l) \\ &= mr(p-r) + \frac{M}{2}r^2 + \frac{m}{2}(p-r)^2 + \mu - \frac{1}{2}(\mu - l) \\ &< \frac{1}{2}m + \frac{1}{2}Mr^2 + \mu - \frac{1}{2}(\mu - l) < \mu. \end{aligned}$$

if we choose $m = \frac{1}{2}(\mu - l)$, and $Mr^2 = \frac{1}{2}(\mu - l)$. Since $U(\hat{p}) > \hat{p}h + (1-\hat{p})l = \mu$, and U is monotonously increasing, there exists a unique $\xi \in (\frac{\hat{p}}{2}, \hat{p})$ such that $U(\xi) = \mu$.

By monotonicity of U, $U(0) < U(\xi) = \mu$. Moreover, $U(q) \le U(0) + (h-l)q \le \mu + (h-l)q$ since $U'(q) = -\Phi(q) + h - l \le h - l$, $q \in [0,1]$.

We are ready to show that $\overline{U} = \max\{\mu, U\}$ is indeed a subsolution. To this end, notice that $\overline{U} = G$, $q \in [0,\xi] \cup [p,1]$. To show $\overline{U}$ is a subsolution, it suffices to show that, for $q \in (\xi, p)$, it holds that

$$\rho U(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 U''_-(q) + C_I(q) < 0$$

For every $q \in (\xi, r] \subset (\frac{\hat{p}}{2}, r]$, since $U''_-(q) = M$ for $q < r \le \frac{1}{2}$ and $U(q) \le \mu + (h-l)q$, we have

$$\rho U(q) - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 U''_-(q) + C_I(q) \le \rho\mu + \hat{c} + \rho(h-l)q - \frac{M}{8\sigma^2}(h-l)^2 q^2 \tag{C.7}$$

$$\le \rho\mu + \hat{c} + \frac{4\rho^2\sigma^2}{M} - \frac{M}{16\sigma^2}(h-l)^2 q^2 \tag{C.8}$$

$$\le \rho\mu + \hat{c} + \frac{4\rho^2\sigma^2}{M} - \frac{M}{16\sigma^2}(h-l)^2 \left(\frac{\hat{p}}{2}\right)^2 \tag{C.9}$$

$$= \rho\mu + \hat{c} + \frac{4\rho^2\sigma^2}{M} - \frac{M}{64\sigma^2}(\mu - l)^2. \tag{C.10}$$

where $\hat{c} = \max C_I(q)$ for $q \in [0,1]$.

We choose $M = \max\left(\frac{128\sigma^2(\rho\mu+\hat{c})}{(\mu-l)^2}, \frac{32\rho\sigma^2}{\mu-l}, 2(\mu-l)\right)$ sufficiently large such that the last quantity above is negative. Then, $r = \sqrt{\frac{\mu-l}{2M}} \le \frac{1}{2}$ is chosen accordingly. We note that the choices of m, M, r are all independent of the value of h.

Since $U''(q) = m$, $r < q < p$, and $U(q) \leq U(1) = h$, by choosing

$$p = 1 - \sqrt{\frac{2\sigma^2(\rho h + \check{c})}{m(h-l)^2 r^2}},$$

with $\check{c} = \min C_I(q)$ for $q \in [0,1]$, we have

$$\rho U(q) + C_I(q) - U''(q)\frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 q^2(1-q)^2 < \rho h + \hat{c} - \frac{1}{2}\left(\frac{h-l}{\sigma}\right)^2 r^2(1-p)^2 m = 0.$$

Note that $p \to 1$ as $h \to \infty$. In conclusion, by picking $m = \frac{1}{2}(\mu - l)$, $M = \max\{\frac{128\sigma^2(\rho\mu+\hat{c})}{(\mu-l)^2}, \frac{32\rho\sigma^2}{\mu-l}, 2(\mu-l)\}$, $r = \sqrt{\frac{\mu-l}{2M}}$, and $p = 1 - \sqrt{\frac{2\sigma^2(\rho h+\check{c})}{m(h-l)^2r^2}}$, we obtain that $\overline{U}$ is a viscosity subsolution and $\overline{U}(q) > G(q)$, $q \in (\xi, p)$. By comparison, we deduce that$V(q) \geq \overline{U}(q) > G(q)$, $q \in (\xi, p)$, and hence $\underline{q} \leq \xi < p \leq \overline{q}$. As $h \to \infty$, because $\xi < \hat{p} \to 0$ and $p \to 1$, we obtain that $\underline{q} \to 0$ and $\overline{q} \to 1$. □

Works Cited

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

Bita Analui and Georg Ch Pflug. On distributionally robust multiperiod stochastic optimization. *Computational Management Science*, 11(3):197–220, 2014.

Brian DO Anderson and John B Moore. *Optimal control: linear quadratic methods*. Courier Corporation, 2007.

Kenneth J Arrow, David Blackwell, and Meyer A Girshick. Bayes and minimax solutions of sequential decision problems. *Econometrica, Journal of the Econometric Society*, pages 213–244, 1949.

Julio Backhoff, Mathias Beiglbock, Yiqing Lin, and Anastasiia Zalashko. Causal transport in discrete time and applications. *SIAM Journal on Optimization*, 27(4):2528–2562, 2017.

Gah-Yi Ban and Cynthia Rudin. The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1):90–108, 2019.

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

Güzin Bayraksan and David K Love. Data-driven stochastic programming using phi-divergences. In *INFORMS TutORials in Operations Research*, pages

1–19. INFORMS, 2015.

Thierry Bazier-Matte and Erick Delage. Generalization bounds for regularized portfolio selection with market side information. *INFOR: Information Systems and Operational Research*, 58(2):374–401, 2020.

Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton university press, 2009.

Aharon Ben-Tal, Dick Den Hertog, Anja De Waegenaere, Bertrand Melenberg, and Gijs Rennen. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357, 2013.

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

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

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

Dimitris Bertsimas and Nihal Koduri. Data-driven optimization: A reproducing kernel hilbert space approach. *Operations Research*, 70(1):454–471, 2022.

Dimitris Bertsimas and Christopher McCord. From predictions to prescriptions in multistage optimization problems. *arXiv preprint arXiv:1904.11637*, 2019.

Dimitris Bertsimas and Bart Van Parys. Bootstrap robust prescriptive analytics. *Mathematical Programming*, pages 1–40, 2021.

Dimitris Bertsimas, Dan A Iancu, and Pablo A Parrilo. Optimality of affine policies in multistage robust optimization. *Mathematics of Operations Research*, 35(2):363–394, 2010.

Dimitris Bertsimas, Dan Andrei Iancu, and Pablo A Parrilo. A hierarchy of near-optimal policies for multistage adaptive optimization. *IEEE Transactions on Automatic Control*, 56(12):2809–2824, 2011.

Dimitris Bertsimas, Christopher McCord, and Bradley Sturt. Dynamic optimization with side information. *European Journal of Operational Research*, 2022.

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

Jose Blanchet, Yang Kang, and Karthyek Murthy. Robust wasserstein profile inference and applications to machine learning. *Journal of Applied Probability*, 56(3):830–857, 2019.

Michael W Brandt, Pedro Santa-Clara, and Rossen Valkanov. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies*, 22(9):3411–3447, 2009.

Junyu Cao and Rui Gao. Contextual decision-making under parametric uncertainty and data-driven optimistic optimization. *Available at Optimization Online*, 2021.

Emilio Carrizosa, Alba V Olivares-Nadal, and Pepa Ramírez-Cobo. Robust newsvendor problem with autoregressive demand. *Computers & Operations Research*, 68:123–133, 2016.

Yeon-Koo Che and Konrad Mierendorff. Optimal dynamic allocation of attention. *American Economic Review*, 109(8):2993–3029, 2019.

Xin Chen, Melvyn Sim, Peng Sun, and Jiawei Zhang. A linear decision-based approximation approach to stochastic programming. *Operations Research*, 56 (2):344–357, 2008.

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

Michael G Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American mathematical society*, 27(1):1–67, 1992.

Bikramjit Das, Anulekha Dhara, and Karthik Natarajan. On the heavy-tail behavior of the distributionally robust newsvendor. *Operations Research*, 69 (4):1077–1099, 2021.

Erick Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations research*, 58(3):595–612, 2010.

Othman El Balghiti, Adam N Elmachtoub, Paul Grigas, and Ambuj Tewari. Generalization bounds in the predict-then-optimize framework. *Advances in Neural Information Processing Systems*, 32:14412–14421, 2019.

Adam Elmachtoub, Jason Cheuk Nam Liang, and Ryan McNellis. Decision trees for decision-making under the predict-then-optimize framework. In *International Conference on Machine Learning*, pages 2858–2867. PMLR, 2020.

Adam N Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 2021.

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

Adrián Esteban-Pérez and Juan M Morales. Distributionally robust stochastic programs with side information based on trimmings. *arXiv preprint*, page arXiv:2009.10592, 2020.

Alexander Estes. Slow rates of convergence in optimization with side information. *Available at SSRN 3803427*, 2021.

Wendell H Fleming and Raymond W Rishel. *Deterministic and stochastic optimal control*, volume 1. Springer Science & Business Media, 2012.

Mingyang Fu, Xiaobo Li, and Lianmin Zhang. Data-driven feature-based newsvendor: A distributionally robust approach. *Available at SSRN 3885663*, 2021.

Drew Fudenberg, Philipp Strack, and Tomasz Strzalecki. Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108(12):3651–84, 2018.

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

Guillermo Gallego, Jennifer K Ryan, and David Simchi-Levi. Minimax analysis for finite-horizon inventory models. *IIE Transactions*, 33(10):861–874, 2001.

Rui Gao. Finite-sample guarantees for wasserstein distributionally robust optimization: Breaking the curse of dimensionality. *arXiv preprint arXiv:2009.04382*, 2020.

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

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

Rui Gao and Anton J. Kleywegt. Distributionally Robust Stochastic Optimization with Wasserstein Distance. *arXiv e-prints*, art. arXiv:1604.02199, April 2016.

Rui Gao, Xi Chen, and Anton J Kleywegt. Wasserstein distributionally robust optimization and variation regularization. *arXiv preprint*, page arXiv:1712.06050, 2017.

Angelos Georghiou, Angelos Tsoukalas, and Wolfram Wiesemann. On the optimality of affine decision rules in robust and distributionally robust optimization. *Optimization Online*, 2021.

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

Grani A Hanasusanto, Daniel Kuhn, Stein W Wallace, and Steve Zymler. Distributionally robust multi-item newsvendor problems with multimodal demand distributions. *Mathematical Programming*, 152(1):1–32, 2015a.

Grani A Hanasusanto, Daniel Kuhn, and Wolfram Wiesemann. K-adaptability in two-stage robust binary programming. *Operations Research*, 63(4):877–891, 2015b.

Grani A Hanasusanto, Daniel Kuhn, and Wolfram Wiesemann. K-adaptability in two-stage distributionally robust binary programming. *Operations Research Letters*, 44(1):6–11, 2016.

Grani Adiwena Hanasusanto and Daniel Kuhn. Robust data-driven dynamic programming. *Advances in Neural Information Processing Systems*, 26, 2013.

Lauren Hannah, Warren Powell, and David Blei. Nonparametric density estimation for stochastic optimization with an observable state variable. *Advances in Neural Information Processing Systems*, 23:820–828, 2010.

Nam Ho-Nguyen and Fatma Kılınç-Karzan. Risk guarantees for end-to-end prediction and optimization processes. *Management Science*, 2022.

Johannes Hörner and Andrzej Skrzypacz. Learning, experimentation and information design. *Advances in Economics and Econometrics*, 1:63–98, 2017.

Omar El Housni and Vineet Goyal. On the optimality of affine policies for budgeted uncertainty sets. *arXiv preprint*, page arXiv:1807.00163, 2018.

Yichun Hu, Nathan Kallus, and Xiaojie Mao. Fast rates for contextual linear optimization. *Management Science (Forthcoming)*, 2022.

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

Nathan Kallus and Xiaojie Mao. Stochastic optimization forests. *arXiv preprint*, page arXiv:2008.07473, 2020.

Nathan Kallus and Xiaojie Mao. Stochastic optimization forests. *Management Science (Forthcoming)*, 2022.

Rohit Kannan, Güzin Bayraksan, and James R Luedtke. Data-driven sample average approximation with covariate information. *Optimization Online*, page preprint, 2020a.

Rohit Kannan, Güzin Bayraksan, and James R Luedtke. Residuals-based distributionally robust optimization with covariate information. *arXiv preprint*, page arXiv:2012.01088, 2020b.

I. Karatzas and X. Zhao. *Bayesian Adaptive Portfolio Optimization*, page 632–669. Cambridge University Press, 2001. doi: 10.1017/CBO9780511569708.018.

Godfrey Keller and Sven Rady. Strategic experimentation with poisson bandits. *Theoretical Economics*, 5(2):275–311, 2010.

Daniel Kuhn, Peyman Mohajerin Esfahani, Viet Anh Nguyen, and Soroosh Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning. In *Operations Research & Management Science in the Age of Analytics*, pages 130–166. INFORMS, 2019.

Rémi Lassalle. Causal transference plans and their monge-kantorovich problems. *arXiv preprint arXiv:1303.6925*, 2013.

Sangyoon Lee, Hyunwoo Kim, and Ilkyeong Moon. A data-driven distributionally robust newsvendor model with a Wasserstein ambiguity set. *Journal of the Operational Research Society*, pages 1–19, 2020.

Soonhui Lee, Tito Homem-de Mello, and Anton J Kleywegt. Newsvendor-type models with decision-dependent uncertainty. *Mathematical Methods of Operations Research*, 76(2):189–221, 2012.

Yuanyuan Lian, Lihe Wang, and Kai Zhang. Pointwise regularity for fully nonlinear elliptic equations in general forms. *arXiv preprint arXiv:2012.00324*, 2020.

Annie Liang, Xiaosheng Mu, and Vasilis Syrgkanis. Optimal learning from multiple information sources. *arXiv preprint arXiv:1703.06367*, 2017.

Mo Liu, Meng Qi, and Zuo-Jun Max Shen. End-to-end deep learning for inventory management with fixed ordering cost and its theoretical analysis. *Available at SSRN 3888897*, 2021.

Liwan H Liyanage and J George Shanthikumar. A practical inventory control policy using operational statistics. *Operations Research Letters*, 33(4):341–348, 2005.

Gar Goei Loke, Qinshen Tang, and Yangge Xiao. Decision-driven regularization: Harmonizing the predictive and prescriptive. *Available at SSRN*, page 3623006, 2020.

David G Luenberger. *Optimization by vector space methods*. John Wiley & Sons, 1997.

Ulrike von Luxburg and Olivier Bousquet. Distance-based classification with Lipschitz functions. *Journal of Machine Learning Research*, 5:669–695, dec 2004. ISSN 1532-4435.

Mashed. Half-and-half: What you should know before you buy. `https://www.youtube.com/watch?v=B6lIt9KFZhE`, 2022. Accessed: 2022-10-07.

Tatiana Mayskaya. Dynamic choice of information sources. *California Institute of Technology Social Science Working Paper, ICEF Working Paper WP9/2019/05*, 2022.

Qi Meng, Yuanyuan Shi, Yongzhi Qi, Chenxin Ma, Rong Yuan, Di Wu, and Zuojun (Max) Shen. A practical end-to-end inventory management model with deep learning. *Management Science*, page (Forthcoming), 2021.

Sanjoy K Mitter. Filtering and stochastic control: A historical perspective. *IEEE Control Systems Magazine*, 16(3):67–76, 1996.

J Morris. The kalman filter: A robust estimator for some classes of linear quadratic problems. *IEEE Transactions on Information Theory*, 22(5):526–534, 1976.

Giuseppe Moscarini and Lones Smith. The optimal level of experimentation. *Econometrica*, 69(6):1629–1644, 2001.

Miguel Angel Muñoz, Salvador Pineda, and Juan Miguel Morales. A bilevel framework for decision-making under uncertainty with contextual information. *Omega*, 108:102575, 2022.

Karthik Natarajan, Melvyn Sim, and Joline Uichanco. Asymmetry and ambiguity in newsvendor models. *Management Science*, 64(7):3146–3167, 2018.

Afshin Oroojlooyjadid, Lawrence V Snyder, and Martin Takáč. Applying deep learning to the newsvendor problem. *IISE Transactions*, 52(4):444–463, 2020.

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

Georg Pflug and David Wozabal. Ambiguity in portfolio selection. *Quantitative Finance*, 7(4):435–442, 2007.

Georg Ch Pflug and Alois Pichler. *Multistage stochastic optimization*. Springer, 2014.

Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

Alois Pichler and Alexander Shapiro. Mathematical foundations of distributionally robust multistage optimization. *arXiv preprint arXiv:2101.02498*, 2021.

Ioana Popescu. Robust mean-covariance solutions for stochastic optimization. *Operations Research*, 55(1):98–112, 2007.

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

Meng Qi, Yuanyuan Shi, Yongzhi Qi, Chenxin Ma, Rong Yuan, Di Wu, and Zuojun (Max) Shen. A practical end-to-end inventory management model with deep learning. *Management Science*, page (Forthcoming), 2021.

Hamed Rahimian, Güzin Bayraksan, and Tito Homem-de Mello. Distributionally robust newsvendor problems with variation distance. *Optimization Online*, page preprint, 2017.

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

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

Kristin Reikvam. Viscosity solutions of optimal stopping problems. *Stochastics and Stochastic Reports*, 62(3-4):285–301, 1998.

R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.

Ludger Rüschendorf. The wasserstein distance and approximation theorems. *Probability Theory and Related Fields*, 70(1):117–129, 1985.

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

Chuen-Teck See and Melvyn Sim. Robust approximation to multiperiod inventory management. *Operations research*, 58(3):583–594, 2010.

Soroosh Shafieezadeh-Abadeh, Daniel Kuhn, and Peyman Mohajerin Esfahani. Regularization via mass transportation. *Journal of Machine Learning Research*, 20(103):1–68, 2019.

Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

Alexander Shapiro. On duality theory of conic linear problems. *Semi-Infinite Programming: Recent Advances*, pages 135–165, 2001.

Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2014.

Lloyd Shapley. Some topics in two-person games. *Advances in game theory*, 52:1–29, 1964.

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

Aman Sinha, Hongseok Namkoong, and John Duchi. Certifying some distributional robustness with principled adversarial training. In *International Conference on Learning Representations*, 2018.

Albert Nikolaevich Sirjaev. *Statistical sequential analysis: Optimal stopping rules*, volume 38. American Mathematical Soc., 1973.

Harold W Sorenson. An overview of filtering and stochastic control in dynamic systems. *Control and dynamic systems*, 12:1–61, 1976.

Prateek R Srivastava, Yijie Wang, Grani A Hanasusanto, and Chin Pang Ho. On data-driven prescriptive analytics with side information: A regularized nadaraya-watson approach. *arXiv preprint arXiv:2110.04855*, 2021.

Robert F Stengel. *Optimal control and estimation*. Courier Corporation, 1994.

Bradley Sturt. A nonparametric algorithm for optimal stopping based on robust optimization. *arXiv preprint arXiv:2103.03300*, 2021.

Anirudh Subramanyam, Chrysanthos E Gounaris, and Wolfram Wiesemann. K-adaptability in two-stage mixed-integer robust optimization. *Mathematical Programming Computation*, pages 1–32, 2019.

Wenpin Tang, Yuming Paul Zhang, and Xun Yu Zhou. Exploratory HJB equations and their convergence. *SIAM Journal on Control and Optimization*, 60(6):3191–3216, 2022.

L Beril Toktay and Lawrence M Wein. Analysis of a forecasting-production-inventory system with stationary demand. *Management Science*, 47(9):1268–1281, 2001.

Theja Tulabandhula and Cynthia Rudin. Machine learning with operational costs. *Journal of Machine Learning Research*, 14:1989–2028, 2013.

Abraham Wald. Foundations of a general theory of sequential decision functions. *Econometrica, Journal of the Econometric Society*, pages 279–313, 1947.

Zizhuo Wang, Peter W Glynn, and Yinyu Ye. Likelihood robust optimization for data-driven problems. *Computational Management Science*, 13(2):241–261, 2016.

David Wozabal. A framework for optimization under ambiguity. *Annals of Operations Research*, 193(1):21–47, 2012.

David Wozabal. Robustifying convex risk measures for linear portfolios: A nonparametric approach. *Operations Research*, 62(6):1302–1315, 2014.

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

Linwei Xin and David Alan Goldberg. Distributionally robust inventory control when demand is a martingale. *Mathematics of Operations Research (Forthcoming)*, 2022.

Jincheng Yang, Luhao Zhang, Ningyuan Chen, Rui Gao, and Ming Hu. Decision-making with side information: a causal transport robust approach. *OptOnline e-prints*, 2022.

İhsan Yanıkoğlu, Bram L Gorissen, and Dick den Hertog. A survey of adjustable robust optimization. *European Journal of Operational Research*, 277 (3):799–813, 2019.

Luhao Zhang, Jincheng Yang, and Rui Gao. A simple duality proof for wasserstein distributionally robust optimization. *arXiv preprint arXiv:2205.00362*, 2022.

Luhao Zhang, Jincheng Yang, and Rui Gao. Optimal robust policy for feature-based newsvendor. *Management Science*, 2023. doi: 10.1287/mnsc.2023.4810.

Chaoyue Zhao and Yongpei Guan. Data-driven risk-averse stochastic optimization with Wasserstein metric. *Operations Research Letters*, 46(2):262–267, 2018.

Jianzhe Zhen, Daniel Kuhn, and Wolfram Wiesemann. Mathematical foundations of robust and distributionally robust optimization. *arXiv preprint arXiv:2105.00760*, 2021.

Weijie Zhong. Optimal dynamic information acquisition. *Econometrica*, 90(4): 1537–1582, 2022.

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

Taozeng Zhu, Jingui Xie, and Melvyn Sim. Joint estimation and robustness optimization. *Management Science*, 2021.

Steve Zymler, Daniel Kuhn, and Berç Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137, 2013.