Screening with Limited Information: A Dual Perspective and A Geometric Approach

Zhi Chen
Department of Management Sciences, College of Business, City University of Hong Kong, Hong Kong
zhi.chen@cityu.edu.hk

Zhenyu Hu
Department of Analytics & Operations, NUS Business School, National University of Singapore, Singapore
bizhuz@nus.edu.sg

Ruiqin Wang
Institute of Operations Research and Analytics, National University of Singapore, Singapore
ruiqin_wang@u.nus.edu

Consider a seller seeking a selling mechanism to maximize the worst-case revenue obtained from a buyer whose valuation distribution lies in a certain ambiguity set. For a generic convex ambiguity set, we show via the minimax theorem that strong duality holds between the problem of finding the optimal robust mechanism and a minimax pricing problem where the adversary first chooses a worst-case distribution and then the seller decides the best posted price mechanism. This implies that the extra value of optimizing over more sophisticated mechanisms exactly amounts to the value of eliminating distributional ambiguity under a posted price mechanism. The duality result also connects prior literature that separately studies the primal (robust mechanism design) and problems related to the dual (*e.g.*, robust pricing, buyer-optimal pricing and personalized pricing). We further provide a geometric approach to analytically solving the minimax pricing problem (as well as the robust pricing problem) for several important ambiguity sets such as the ones with mean and various dispersion measures, and with the Wasserstein metric. The solutions are then used to construct the optimal robust mechanism and to compare with the solutions to the robust pricing problem. Uniqueness of the worst-case distribution can also be established for some cases.

Key words: robust mechanism design, moment condition, mean-preserving contraction/spread, Wasserstein metric.

1. Introduction

Consider a seller who sells a product to a risk-neutral buyer. The buyer's valuation is private information and unobservable to the seller. In the classic pricing problem, the seller knows that the buyer's valuation follows a distribution and seeks a revenue maximizing price. Posting a fixed price—known as the posted price mechanism—is only one way to sell the product, and there are many other ways to sell. For example, the seller can sell a lottery: the buyer pays a fee to enter the lottery and wins the product with a certain probability.[1] Alternatively, the buyer and the seller can be involved in a dynamic bargaining process. In the seminal work of Myerson (1981), it is shown that when the virtual valuation is increasing, the posted price mechanism is optimal among all possible selling mechanisms.[2] Riley and Zeckhauser (1983) later generalize this observation to any distribution as long as it is known to the seller.

In practice, however, the distribution of the buyer's valuation is rarely known precisely. Often, the seller only has a confident prediction of certain statistics, say maximum and mean, of the buyer's valuation, or an empirical distribution based on data.[3] How does a posted price mechanism perform in such a case and is there a better mechanism that the seller can use? Let us consider the following example. The buyer values the product at \$1 or \$1/2 or no value at all. The seller is aware of the possible values and is also confident that the average valuation is \$1/2. However, the seller does not have any further information on the probability of each possible value. Suppose the seller presumes a uniform distribution over the values, then the optimal selling mechanism (under the hypothesized distribution) is to post a price of \$1/2 with the hypothesized optimal revenue of \$1/3. However, if in the true distribution the buyer values the product at \$0 or \$1 each with half probability (with the average valuation staying at \$1/2), then the seller's expected revenue would drop to \$1/4 under the posted price of \$1/2. Consider now a selling mechanism such that the seller posts a price of \$2/3 and simultaneously sells a lottery priced at \$1/3 with a winning probability of 2/3. Clearly, a buyer with a valuation zero would not purchase anything. A buyer who values the product at \$1/2 would prefer to buy the lottery while the buyer with a valuation of \$1 would pay the posted price.[4] One can show that the seller's revenue is \$1/3 regardless of the distribution of the buyer's valuation.

The example above shows that the seller can indeed do better and achieve a robust revenue guarantee of \$1/3 if using a selling mechanism beyond the posted price mechanism. This motivates us to study the general problem of finding a robust selling mechanism when the seller has limited information on the valuation distribution. In particular, we assume the seller merely knows that F belongs to an *ambiguity set* of probability distributions. The seller then seeks the selling mechanism that maximizes the worst-case expected revenue.

[1] The practice of using lotteries as a selling mechanism is commonly seen in the online gaming industry where the implementation of such mechanisms (for instance, in the form of drop rates of rare in-game items) becomes much easier compared to brick-and-mortar stores.

[2] Myerson (1981) establishes the optimality of second-price auction with a reserve price for the setting of multiple buyers. The second-price auction with a reserve price reduces to the posted price mechanism when there is one buyer.

[3] Empirical distribution plays a key role in the growing literature of revenue maximization from samples that seeks optimal price/mechanism with performance guarantees; see, *e.g.*, Allouah et al. (2021), Cole and Roughgarden (2014), Balcan et al. (2017), and the references therein.

[4] We assume that the buyer would purchase the product if he is indifferent between purchasing and no purchasing; and the buyer would pay the posted price if indifferent between buying the product and buying the lottery.

Our first contribution is an observation via the *minimax theorem* that when the ambiguity set is convex, *strong duality* holds between our problem of finding the optimal mechanism and its dual problem, referred to as the minimax pricing problem, in which an adversary selects a worst-case distribution that minimizes the maximal revenue achievable by a posted price mechanism. Although in the minimax pricing problem, the seller is restricted to a posted price mechanism, she enjoys the information advantage of first seeing the distribution chosen by the adversary before deciding her own price. Strong duality directly implies that the extra value brought by more sophisticated selling mechanisms over the simple posted price mechanism is exactly the value of distribution information under a posted price mechanism. Our minimax pricing problem is closely related to problems studied in robust pricing, buyer-optimal pricing and personalized pricing literature (we review each of them in more detail subsequently), and hence our result connects these problems with robust mechanism design via strong duality.

Our second contribution is a *geometric approach* to solving the minimax pricing problem. Geometrically, the worst-case distribution can be found by finding from the ambiguity set a distribution that is tangent to the lowest possible level curve of a bilinear function (which is a hyperbola). A similar geometric intuition also applies to the maximin pricing problem, where the optimal posted price is the tangent point of the level curve of a bilinear function and a curve we dub "adversarial selling probability". We use this geometric approach to analytically solve several important special cases: ambiguity sets with mean (Section 4.1), mean-preserving contraction (Section 4.2) and Wasserstein metric (Section 4.3), as well as ambiguity sets with other dispersion measures such as variance (Appendix C.1), mean absolute deviation (Appendix C.2) and mean-preserving spread (Appendix C.3). The dual solution enables us to construct the optimal robust mechanism and to compare with the optimal posted price mechanism. While part of our analytical results for the ambiguity sets with mean and variance/mean absolute deviation and the one based on mean-preserving contraction are known in the literature, our geometric approach provides a more systematic and intuitive proof. Our characterization of both the optimal robust mechanism and the posted price mechanism for the general Wasserstein case, to the best of our knowledge, is new to the literature. In Table 1, we summarize existing results and ours for different ambiguity sets. It is worth noting that for the multi-product mechanism design problem, Carroll (2017) shows that it is optimal to solve the robust mechanism design problem for each product separately as long as one only has information about the marginal distributions. Our comprehensive results in Table 1 can then be directly applied to the multi-product robust mechanism design problem with a rich family of ambiguity sets that can flexibly specify different distributional information for each marginal.

Ambiguity Set	Robust Screening	Minimax Pricing	Maximin Pricing
Mean	Carrasco et al. (2018) Proposition 2	Elmachtoub et al. (2020) ✓ Theorem 1	Carrasco et al. (2018) ✓ Proposition 3
Mean Preserving Contraction	Du (2018) Proposition 4	Roesler and Szentes (2017) Theorem 2	— Proposition 5
Wasserstein Distance	Li et al. (2019)[a] Proposition 6/Corollary 1[b]	— ✓ Theorem 4[b]	— ✓ Proposition 7
Mean and Variance	Carrasco et al. (2018)[c] Proposition 8	— ✓ Theorem 5	Chen et al. (2020a)[d] ✓ Proposition 9
Mean Absolute Deviation	— Proposition 10	Elmachtoub et al. (2020) ✓ Theorem 6	Roos et al. (2020) ✓ Proposition 11
Mean Preserving Spread	— Proposition 12	Condorelli and Szentes (2020) Theorem 7	— Proposition 13

[a] Li et al. (2019) consider discrete valuation; while we show that the optimal mechanism can be more complicated for continuous valuation. [b] Our characterization of the optimal mechanism does not hold for type-ρ ($1 < \rho < \infty$) Wasserstein ambiguity sets, but our characterization of the worst-case distribution Theorem 4 does hold for any $\rho \geq 1$. [c] Carrasco et al. (2018) consider possibly unbounded valuation and all distributions having variance bounded from above (as opposed to being equal to a pre-specified value in our setting). [d] Chen et al. (2020a) assume unbounded support and their analysis is based on Cantelli's inequality that does not utilize the support information.

Table 1 Summary of existing results and ours. The "✓" sign indicates that uniqueness of the worst-case distribution can be established for certain cases.

2. Literature Review

The problem of mechanism design with one product and one buyer is also known as the screening problem (Börgers 2015, chapter 2). As we mentioned in the introduction, when the distribution of the buyer's valuation is known, it is well known that the posted price mechanism is optimal. One insightful proof is by Manelli and Vincent (2007), who observe that the screening problem is essentially an infinite dimensional linear program and the posted price mechanisms constitute all the extreme points of the set of feasible mechanisms. The linearity is also the key behind our strong duality result. The screening problem is a fundamental building block for more complicated models that involve multiple products or multiple interacting agents. We refer readers to Börgers (2015) for an overview of these topics in mechanism design.

Our *primal problem* belongs to the literature of robust mechanism design. A number of papers have studied the robust screening problem under specific ambiguity sets. Bergemann and Schlag (2011) consider the ambiguity set with Lévy metric and show that the posted price mechanism is optimal. Ambiguity set with moment information is studied in Carrasco et al. (2018). They characterize the structure of the optimal mechanism for the general case when the ambiguity set contains the first N moments information (with unbounded support) and provide analytical solutions for the case with the first two moments and the case with the first moment and upper bound information. In a context of multiple buyers, Du (2018) studies the ambiguity set based on mean-preserving spread and constructs a mechanism that is optimal when there is only one buyer. For discrete valuations, Pınar and Kızılkale (2017) also characterize the optimal mechanism when the first two moments are known while Li et al. (2019) consider the ambiguity set based on the type-1 Wasserstein metric. Besides the maximin revenue objective which we focus on, the minimax regret objective has also been considered. One is referred to Bergemann and Schlag (2008) for ambiguity set with only support, Bergemann and Schlag (2011) for the one based on Lévy metric and Wang et al. (2020) for the one with moment information.

Among the aforementioned papers on the primal, Bergemann and Schlag (2011), Carrasco et al. (2018) and Du (2018) all observe that the robust mechanism can be solved by finding a saddle point (if one exists) to the objective functional. Both Bergemann and Schlag (2011) and Carrasco et al. (2018) prove the existence of saddle point (which ensures strong duality) for their specific ambiguity sets by utilizing the existence result for Nash equilibrium in Reny (1999), while Du (2018) explicitly constructs the saddle point by pairing his mechanism with the dual solution found in Roesler and Szentes (2017). In comparison, we establish strong duality for general convex ambiguity sets via a simple observation that the objective functional is bilinear so a functional version of von Neumann's classic minimax theorem can be applied. In addition, the existence result of saddle point in Bergemann and Schlag (2011) and Carrasco et al. (2018) is only used to prove the existence of the optimal solution to the primal problem, and characterizations of the optimal mechanism therein are based on analyzing the primal problem. Our characterization, however, is based on analyzing the dual problem first, which admits a more intuitive geometric interpretation.

Our *dual problem* is a minimax pricing problem, which is clearly related to the maximin pricing problem—also called the robust pricing problem—obtained from exchanging the "max" and "min" operators. The max-min inequality directly implies that the optimal objective value of the minimax pricing problem provides an upper bound to that of robust pricing—and in general, the two are not equal. Chen et al. (2020a) study the robust pricing problem with mean and variance. Roos et al. (2020) provide a closed-form solution to the optimal price under the ambiguity set with mean absolute deviation. A multi-buyer extension of robust pricing problem is finding an optimal reserve price in auctions and one is referred to Suzdaltsev (2020a) and references therein for this stream of literature. Our geometric approach also applies to solving the robust pricing problem, which further enables us to visualize the comparison between minimax pricing and maximin pricing.

The dual minimax pricing problem itself also has a more subtle connection to the buyer-optimal pricing (Roesler and Szentes 2017, Condorelli and Szentes 2020) and the value of personalized pricing (Elmachtoub et al. 2020). Both Roesler and Szentes (2017) and Condorelli and Szentes (2020) consider a Stackelberg game where the buyer moves first by choosing his valuation distribution in an ambiguity set to maximize his expected surplus and the seller moves next by deciding the price upon observing buyer's valuation distribution. It turns out that under certain conditions on the ambiguity set (*e.g.*, with mean or mean-preserving contraction) maximizing buyer's surplus is equivalent to minimizing the seller's revenue. Elmachtoub et al. (2020) examine a very different problem of bounding the ratio between the revenues of personalized pricing (or first-degree price discrimination) and of a posted price mechanism for all distributions in an ambiguity set. Yet, their problem of finding an upper bound is again equivalent to our minimax pricing problem. Hence, our paper reveals that these problems are fundamentally the dual problem of robust screening.

Robust mechanism design is also studied in more challenging multi-product or multi-buyer settings. For multiple products, Carroll (2017) and Gravin and Lu (2018) show that when only the marginals are known, it is optimal to post a price for each product separately—the problem reduces to the classic screening problem with known distribution if there is only one product. Koçyiğit et al. (2021) study the minimax regret objective and an ambiguity set with rectangular support and obtain a similar separation result. With one product, their model reduces to the one studied in Bergemann and Schlag (2008). Koçyiğit et al. (2021) also prove a strong duality result via discretizing their specific ambiguity set and explicitly constructing solutions for both primal and dual. For multi-buyer setting, besides the stream of literature that focuses on auctions with reserve prices (see Suzdaltsev 2020a), Suzdaltsev (2020b) studies the optimal deterministic mechanism by considering ambiguity sets with mean and upper bound on the valuations. Koçyiğit et al. (2019) study, among other results, the optimal second-price auction with a reserve price and a closely related randomized mechanism called "highest-bidder lottery" for the same ambiguity set. Du (2018) proposes an asymptotically optimal mechanism for ambiguity sets based on mean-preserving spread while Bandi and Bertsimas (2014) provide computational algorithms when only support information is known.

A dual perspective is also regularly pursued in the setting of Bayesian mechanism design. A common observation is that the incentive compatibility constraints in the mechanism design problem are exactly the constraints in the dual problem of a properly defined network flow problem, or equivalently, the dual variables to the incentive compatibility constraints correspond to flow decisions in a network flow problem (see chapter 4 in Vohra 2011). Cai et al. (2019) utilize such a dual perspective to provide a unified proof of several existing results that provide performance bounds for simple mechanisms. Although considering a robust mechanism design problem, Carroll (2017) uses the network flow dual of the Bayesian counterpart to help establish the separation result. Both

Giannakopoulos and Koutsoupias (2014) and Daskalakis et al. (2017) also develop dual formulations based on a reformulation of the primal multi-dimensional screening problem that uses customer's utility function as the decision variable (the incentive compatibility constraints then become the requirement that the utility function must be convex). Daskalakis et al. (2017) further connect the proposed dual problem with the optimal transport problem and prove that strong duality holds. Our dual perspective, however, is neither solely based on the dual of the mechanism design problem for a fixed distribution as the above mentioned papers nor based on the dual of the inner minimization alone as in, for example, Koçyiğit et al. (2019). Rather, we view the maximin problem together as our primal, and we derive our dual problem via the minimax theorem rather than utilizing the machinery of linear programming duality.

Finally, our work is also related to distributionally robust optimization. Delage and Ye (2010) and Wiesemann et al. (2014) study ambiguity sets based on moment information. Blanchet and Murthy (2019), Gao and Kleywegt (2016) and Mohajerin Esfahani and Kuhn (2018) study ambiguity sets based on the Wasserstein metric. In a recent work, Chen et al. (2020b) provide further generalizations in terms of both modeling uncertainty and approximations of adaptive decisions. We refer to Rahimian and Mehrotra (2019) and references therein for a comprehensive review and recent advances in distributionally robust optimization.

Notation. We use $\mathcal{P}_+$ (resp., $\mathcal{P}$) to denote the set of nonnegative measures (resp., probability distributions) supported on $\mathcal{V}=[0,1]$. We denote the cumulative distribution function (CDF) and the complementary CDF (CCDF) by F and $\bar{F}$, and their generalized inverse by F^{-1} and $\bar{F}^{-1}$. When it is clear from the content, F, $\bar{F}$, F^{-1}, and $\bar{F}^{-1}$ are all used to identify the buyer's valuation distribution.[5] The expectation under F is $\mathbb{E}_F[\cdot]$. We refer to "decreasing/increasing" (resp., "positive/negative") in the weak (resp., strong) sense.

3. The Minimax Theorem

By the revelation principle (Myerson 1981), we can focus on direct mechanisms of the form

$$\mathcal{M}=\left\{ x:\mathcal{V}\mapsto[0,1],\, t:\mathcal{V}\mapsto\mathbb{R} \;\middle|\; \begin{array}{ll} vx(v)-t(v)\geq vx(u)-t(u) & \forall u,v\in\mathcal{V} \\ vx(v)-t(v)\geq 0 & \forall v\in\mathcal{V} \end{array} \right\},$$

where $\mathcal{V}=[0,1]$ is the support of the random buyer's valuation $\tilde{v}$. In a direct mechanism, based on the realization of buyer's valuation v, the seller allocates the product to him with probability $x(v)$ and requests a transfer $t(v)$. The direct mechanism (x,t) needs to satisfy two sets of constraints. First, known as *incentive compatibility*, requires that no buyer misreports his valuation. Second, known

[5] The generalized inverses are defined as $F^{-1}(q)=\inf\{v\in\mathcal{V}\mid F(v)\geq q\}$ and $\bar{F}^{-1}(q)=\inf\{v\in\mathcal{V}\mid \bar{F}(v)\leq q\}$. In general, $\bar{F}^{-1}(\bar{F}(v))\leq v$ and $\bar{F}(\bar{F}^{-1}(q))\leq q$. We follow the convention $\bar{F}(v)=1$ for $v<0$ and $\bar{F}(v)=0$ for $v\geq 1$.

as *individual rationality*, requires that the buyer is willing to participate in the seller's mechanism. A classic result in mechanism design establishes that a direct mechanism is incentive compatible if and only if the allocation $x(v)$ is increasing in v and the transfer is specified by $t(v) = vx(v) - \int_0^v x(u)\mathrm{d}u - (0 \cdot x(0) - t(0))$. Since it is optimal to ensure the net utility of the lowest-valuation buyer to be zero, *i.e.*, $t(0) = 0$, the set of direct mechanisms can be equivalently represented as:

$$\mathcal{M} = \left\{ x : \mathcal{V} \mapsto [0,1],\, t : \mathcal{V} \mapsto \mathbb{R} \;\middle|\; \begin{array}{ll} x(u) \geq x(v) & \forall u, v \in \mathcal{V} : u \geq v \\ t(v) = vx(v) - \displaystyle\int_0^v x(u)\mathrm{d}u & \forall v \in \mathcal{V} \end{array} \right\}. \tag{1}$$

One notable class of direct mechanisms is the posted price mechanism taking the form:

$$(x(v), t(v)) = \begin{cases} (0,0) & v \in \mathcal{V} : v < p \\ (1,p) & v \in \mathcal{V} : v \geq p. \end{cases}$$

It is well known that the posted price mechanism is optimal when the seller has perfect knowledge of the buyer's valuation distribution F. In this paper, we assume that the seller merely knows that $F \in \mathcal{F}$ for some convex ambiguity set $\mathcal{F} \subseteq \mathcal{P}$. Convexity is a mild requirement satisfied by many commonly used ambiguity sets. Being aware of distributional ambiguity, the seller seeks a mechanism that maximizes the worst-case expected revenue among all distributions in the ambiguity set. That is, the seller solves the *robust screening problem* (or *robust mechanism design problem*):

$$\sup_{(x,t) \in \mathcal{M}} \inf_{F \in \mathcal{F}} \mathbb{E}_F[t(\tilde{v})]. \tag{2}$$

A closely related problem is when the seller is restricted to the posted price mechanism and solves

$$\sup_{p \in \mathcal{V}} \inf_{F \in \mathcal{F}} p\bar{F}_-(p), \tag{3}$$

where $\bar{F}_-(p) := \lim_{v \uparrow p} \bar{F}(v) = \mathbb{P}(\tilde{v} \geq p)$. Due to the simplicity of the posted price mechanism, problem (3), known as the *robust pricing problem*, is also studied in the recent literature (Chen et al. 2020a, Roos et al. 2020). We will sometimes refer to problem (3) also as *maximin pricing problem* to better connect it with a problem that we refer to as *minimax pricing problem*:

$$\inf_{F \in \mathcal{F}} \sup_{p \in \mathcal{V}} p\bar{F}_-(p). \tag{4}$$

Note that problem (4) is obtained from problem (3) by simply exchanging the "max" and "min" operators, and the max-min inequality implies that the optimal objective value of (4) provides an upper bound to that of (3). Indeed, in the minimax pricing problem, the seller has the information advantage to decide the price after seeing the distribution chosen by the adversary. Interestingly, we show next that strong duality holds between our robust screening problem (2) and the minimax pricing problem (4).

Lemma 1. *When $\mathcal{F}$ is convex, we have*

$$\sup_{(x,t)\in\mathcal{M}} \inf_{F\in\mathcal{F}} \mathbb{E}_F[t(\tilde{v})] = \inf_{F\in\mathcal{F}} \sup_{p\in\mathcal{V}} p\bar{F}_{-}(p).$$

Lemma 1 may be less surprising once one recognizes that the objective functional $\mathbb{E}_F[t(\tilde{v})]$ is bilinear in F and t, and a functional version of von Neumann's minimax theorem (see Borwein and Zhuang 1986) can be applied to obtain

$$\sup_{(x,t)\in\mathcal{M}} \inf_{F\in\mathcal{F}} \mathbb{E}_F[t(\tilde{v})] = \inf_{F\in\mathcal{F}} \sup_{(x,t)\in\mathcal{M}} \mathbb{E}_F[t(\tilde{v})].^6$$

The latter problem is then equivalent to the minimax pricing problem since given a distribution, the posted price mechanism is optimal. Similar observations as in Lemma 1 have also been made in other contexts. For example, Yao's principle (Yao 1977) connects a randomized algorithm's performance on the worst-case input and the worst-case expected performance of a deterministic algorithm. The principle is a direct consequence of von Neumann's minimax theorem and has been commonly applied to bound the performance of a randomized algorithm (see Bei et al. 2017, for an example in combinatorial auction setting).[7] In our context, by Lemma 1, the following relationships then hold among problems (2), (3) and (4):

$$\sup_{(x,t)\in\mathcal{M}} \inf_{F\in\mathcal{F}} \mathbb{E}_F[t(\tilde{v})] = \inf_{F\in\mathcal{F}} \sup_{p\in\mathcal{V}} p\bar{F}_{-}(p) \geq \sup_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}_{-}(p).$$

In other words, the extra value of optimizing over more sophisticated mechanisms exactly amounts to the value of eliminating distributional ambiguity under a posted price mechanism. This insight in some sense is also in contrast to that of Carroll (2017) in a multi-product setting. In Carroll (2017), the optimal robust mechanism when specifying only the marginal distributions becomes a simple separated posted price mechanism. That is, ignoring the fine details of some distributional information leads to a simpler optimal mechanism. In our problem, however, the optimal mechanism becomes more complicated by utilizing simpler information in a more sophisticated way.

Given a specific ambiguity set $\mathcal{F}$, it might be possible to obtain the duality result in Lemma 1 via the machinery of linear programming duality (in Section 4.1 we derive the dual in this way as an example) as commonly done for problems without the inner minimization in the literature (see, *e.g.*, Vohra 2011, Carroll 2017, Cai et al. 2019). We note, however, that depending on the primal formulation one uses, the dual problem can be in very different forms. The aforementioned papers

[6] For continuous valuations, there is also a technical issue on continuity (which we address in the proof) since linear functional is not necessarily continuous.

[7] We believe context is important. In fact, von Neumann's minimax theorem itself can be viewed as a consequence of the more general theorems on the system of linear inequalities. Yet, the same theorem stated in the context of game theory brings new interpretations and insights (see Kjeldsen 2001, for a historical account of von Neumann's minimax theorem and a discussion on the significance of the context).

all directly dualize the incentive compatibility constraints $vx(v) - t(v) \geq vx(u) - t(u)\ \forall u, v \in \mathcal{V}$. Yet, to arrive at Lemma 1, it will be much more convenient to exploit the equivalent form in (1) instead. We elaborate more on this difference in Appendix C.4 for the special case when $\mathcal{F}$ is a singleton, and in this case Lemma 1 says that the dual of a screening problem with known distribution is simply a pricing problem.

Apart from providing an alternative interpretation to the value of optimal mechanism, problem (4) is of interest in its own right. For example, Roesler and Szentes (2017) as well as Condorelli and Szentes (2020) study the following buyer-optimal pricing problem:

$$\begin{aligned} \sup_{F \in \mathcal{F}} \quad & \mathbb{E}_F[(\tilde{v} - p_0)^+] \\ \text{s.t.} \quad & p_0 \in \arg\max_{p \in \mathcal{V}} \ p\bar{F}_-(p), \end{aligned} \tag{5}$$

which can be interpreted as the buyer ex-ante first chooses his own valuation distribution from the ambiguity set $\mathcal{F}$ so as to maximize his expected ex-post surplus from transacting with the seller.[8] It turns out that under certain conditions, problem (5) is equivalent to problem (4); see Section 4.2 for more detail. As another example, Elmachtoub et al. (2020) study a problem of bounding the ratio between the revenues of personalized pricing and of a posted price mechanism:

$$\sup_{F \in \mathcal{F}} \frac{\mathbb{E}_F[\tilde{v}]}{\sup_{p \in \mathcal{V}} p\bar{F}_-(p)}.$$

Here, the numerator inside the maximization is the revenue obtained from personalized pricing (or first-degree price discrimination) and the denominator is that from posted price mechanism. The upper bound on the ratio then informs the maximal benefits obtainable from employing personalized pricing. When the mean of F is known, the above problem is equivalent to problem (4).

Finally, we conclude this section by discussing when the objective values in problems (2), (3) and (4) are all equal; that is, a posted price mechanism is optimal in problem (2). We use a performance ratio to quantify the extra value brought by the optimal mechanism over a posted price mechanism:

$$\eta = \frac{\sup_{p \in \mathcal{V}} \inf_{F \in \mathcal{F}} p\bar{F}_-(p)}{\inf_{F \in \mathcal{F}} \sup_{p \in \mathcal{V}} p\bar{F}_-(p)}.$$

[8] The ability for the buyer to change her valuation distribution is justified by the buyer's learning of her own valuation in Roesler and Szentes (2017), where the ambiguity set then naturally contains all distributions that are mean-preserving contraction of the prior valuation distribution. In particular, when the prior is a two-point distribution, the ambiguity set reduces to the one with mean and support (see Wolitzky 2016). Condorelli and Szentes (2020) provide an alternative interpretation in a setting where an innovative firm (buyer) seeks to buy a patent from a patent holder (seller) to develop her own product. In this setting, the amount of pre-development investment can potentially change the buyer's valuation for the patent and the ambiguity set describes the possible valuation distributions that can be induced by different investment strategies.

The posted price mechanism is optimal if and only if $\eta = 1$. Before providing our sufficient conditions for $\eta = 1$, we define a distribution $\Lambda^\star$ that we call *adversarial selling probability* as:

$$\Lambda^\star : \mathcal{V} \mapsto [0,1] \quad \text{such that} \quad \bar{\Lambda}^\star(p) = \inf_{F \in \mathcal{F}} \bar{F}(p) \quad \forall p \in \mathcal{V}.$$

As $\bar{\Lambda}^\star$ is decreasing, it indeed properly defines a distribution. Problem (3) can then be solved via

$$\sup_{p \in \mathcal{V}} \inf_{F \in \mathcal{F}} p\bar{F}_-(p) = \sup_{p \in \mathcal{V}} \inf_{F \in \mathcal{F}} p\bar{F}(p) = \sup_{p \in \mathcal{V}} p\bar{\Lambda}^\star(p),$$

where we have replaced $\bar{F}_-$ by $\bar{F}$ in the first identity. This is without loss of generality; see Lemma 2, Appendix A. Hence, we will sometimes use the more handy $\bar{F}$ instead in our analysis.

Proposition 1. *If $\mathcal{F}$ is convex, then $\eta = 1$ under either of the following conditions.*

(i) The adversarial selling probability $\Lambda^\star$ in the maximin pricing problem (3) *satisfies $\Lambda^\star \in \mathcal{F}$.*

(ii) There exists a worst-case distribution $F^\star$ to the minimax pricing problem (4) *such that the problem $\sup_{p \in \mathcal{V}} p\bar{F}^\star_-(p)$ admits a unique solution $p^\star$.*

Proposition 1 provides conditions on when a pure strategy is optimal for the seller. Interpreted in this way, condition (*i*) simply implies that the adversary has a unique dominant strategy $\Lambda^\star \in \mathcal{F}$. Hence, the problem reduces to the one where the seller knows the distribution and a pure strategy is optimal. This is the argument given in Bergemann and Schlag (2011) who show that the posted price mechanism is optimal for the ambiguity set with Lévy metric. The same condition can be verified for the ambiguity set based on the type-∞ Wasserstein metric (Corollary 1, Section 4.3). Condition (*ii*) shows that if the minimax pricing problem (4) admits a unique optimal price $p^\star$ (which is not necessarily a dominant strategy), then a posted price with $p^\star$ is also an optimal robust mechanism. We illustrate condition (*ii*) with the following example.

Example 1 (Mixture-Distribution Ambiguity Set). *Consider an ambiguity set that is defined as the convex hull of some given distributions $\bar{F}_1, \ldots, \bar{F}_N$:*

$$\operatorname{conv}(F_1, \ldots, F_N) = \left\{ F \in \mathcal{P} \;\middle|\; \bar{F} = \sum_{n \in [N]} \omega_n \bar{F}_n,\ \sum_{n \in [N]} \omega_n = 1,\ \boldsymbol{\omega} \in \mathbb{R}_+^N \right\}.$$

When $\bar{F}_1, \ldots, \bar{F}_N$ are all concave, any $F \in \operatorname{conv}(F_1, ..., F_N)$ is also concave. Hence, the optimal price corresponding to the worst-case distribution must be unique. By Proposition 1 (ii), $\eta = 1$.

The ambiguity set in Example 1 can be used to model a population of consumers consisting of N segments with $\bar{F}_n$ and ω_n being the valuation distribution and fraction of customers in segment n, respectively (see Chen et al. 2020b for related discussions). The condition that $\bar{F}_1, \ldots, \bar{F}_N$ are all concave, though may appear to be strong, ensures that all distributions in the ambiguity set have an increasing failure rate—a commonly imposed condition on valuation distribution in the pricing and auction literature, and it is satisfied by, for example the Beta distribution with $\alpha > 1$ and $\beta = 1$ (note that $\bar{F}$ in this case is the iso-elastic demand function commonly used in the literature; see Talluri and van Ryzin 2006, p. 325). Bergemann et al. (2020) also consider such mixture distributions and they prove that as long as $p\bar{F}_{n-}(p), n \in [N]$ are all concave, then a posted price mechanism can always achieve one half of the profit under third-degree price discrimination. Our Proposition 1 further ensures that when weights of mixtures are unknown, a posted price mechanism is optimal if price discrimination is not allowed.

Conditions (i) and (ii) do not imply each other. For example, when $\mathcal{F} = \{F\}$ and the problem $\sup_{p \in \mathcal{V}} p\bar{F}_-(p)$ admits multiple solutions, condition (i) is satisfied but condition (ii) is violated. One can also construct a counter example based on Example 1, where condition (ii) holds while condition (i) fails. Finally, conditions (i) and (ii) are imposed on the solutions to problems (3) and (4) respectively, which in many cases can be solved geometrically—an issue we discuss next.

4. A Geometric Approach

Lemma 1 opens up a new geometric perspective to derive closed-form solutions for several common ambiguity sets. Our approach to solve the dual problem (4) is inspired by the observation that the inner maximization for a fixed $\bar{F}$ is equivalent to a program with a bilinear objective:

$$\begin{aligned} \sup_{p,y} \quad & py \\ \text{s.t.} \quad & p \in \mathcal{V},\ y \le \bar{F}_-(p). \end{aligned}$$

The optimal value is then given by the level curve of py tangent to $\bar{F}$ with the optimal price being given by the tangent point. The outer minimization, on the other hand, seeks $\bar{F}$ in the ambiguity set whose corresponding tangent level curve is as low as possible. The lowest possible level curve then yields the optimal revenue of problem (4). The obtained optimal $\bar{F}$ can be used—via complementary slackness—to construct a feasible solution to the primal robust screening problem (2), and one can then certify its optimality by showing that it attains the optimal value of problem (4).

Note, however, that our geometric intuition outlined above does not provide an algorithmic procedure in finding from $\mathcal{F}$ the member that is tangent to the lowest possible level curve, and in general, solving problem (4) still requires further exploration of the special structures in a given ambiguity set $\mathcal{F}$. Nevertheless, the geometric insight would provide an intuitive understanding of the optimal solutions of several important special cases as we demonstrate below and in Appendices C.1 to C.3. In the reminder of this section, we start with the simplest and most widely studied case when only the mean of the valuation distribution is known. Section 4.2 provides an extension with additional deviation information beyond mean. Finally, Section 4.3 studies the largely unexplored ambiguity set based on the Wasserstein metric where our geometric insight leads to surprisingly simple solutions to problems (3) and (4).

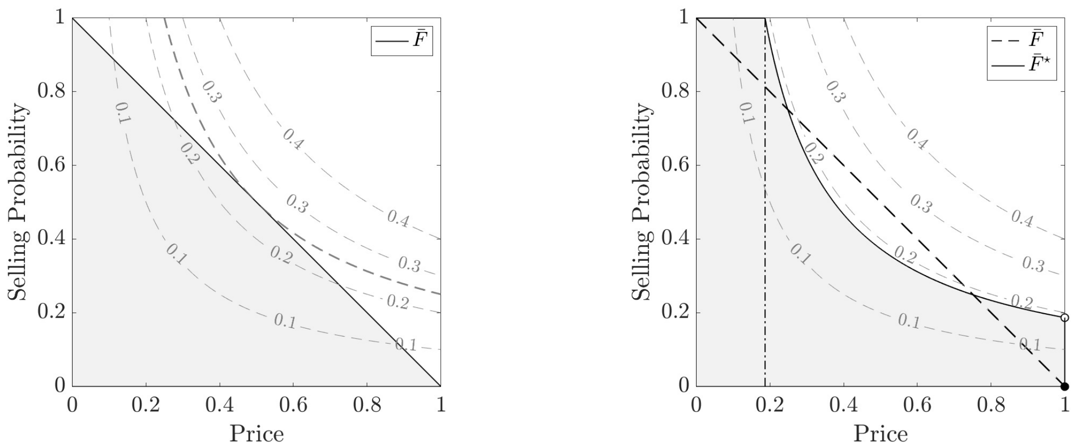

Figure 1 Compressing the shaded area in the left panel by the level curve yields the one in right panel. The threshold $v = \pi^*$ is indicated by the dash-dot line.

4.1. Mean Ambiguity Set

When the seller only knows the mean of the buyer's valuation, the ambiguity set is specified by

$$\mathcal{F}_\mathrm{M} = \left\{ F \in \mathcal{P} \;\middle|\; \int_{\mathcal{V}} v \,\mathrm{d}F(v) = \mu \right\}.$$

Using integration by parts, the mean constraint on any $F \in \mathcal{P}$ is equivalent to $\int_{\mathcal{V}} \bar{F}(v)\mathrm{d}v = \mu$. That is, the ambiguity set $\mathcal{F}_\mathrm{M}$ constitutes all the CCDFs under which the area amounts to μ. To find $F \in \mathcal{F}_\mathrm{M}$ whose corresponding tangent level curve is the lowest, one can imagine a rolling machine with the shape of the level curves in Figure 1 applied to compress a pile of dirt with mass μ (the shaded area in the left panel). The resulting shape of the dirt—keeping the mass unchanged—is then the shaded area in the right panel. We formalize these observations in the following theorem.

Theorem 1. *Given $\mathcal{F}_\mathrm{M}$, the unique worst-case distribution $F^\star$ to problem (4) is*

$$\bar{F}^\star(v) = \begin{cases} 1 & v \in [0, \pi^\star) \\ \pi^\star / v & v \in [\pi^\star, 1) \\ 0 & v = 1 \end{cases}$$

and the optimal revenue is $\pi^\star$. Here, $\pi^\star$ is the unique solution in $[0,1]$ to $\pi - \pi \ln(\pi) = \mu$ and $\pi^\star < \mu$.[9]

[9] Indeed, the nonlinear equation $\pi - \pi\ln(\pi) = \mu$ admits a unique solution $\pi^\star = -\mu / W(-\mu/e)$, where $W(\cdot)$ denotes the Lambert W function (Corless et al. 1996).

Equipped with the optimal revenue $\pi^\star$ and $\bar{F}^\star$, we next construct the optimal mechanism via complementary slackness and show that it attains the value $\pi^\star$. Given $\mathcal{F}_{\mathrm{M}}$, the inner minimization of the primal problem (2) is

$$\begin{aligned}
\inf_{F\in\mathcal{P}_+} \quad & \int_{\mathcal{V}} t(v)\,\mathrm{d}F(v) \\
\text{s.t.} \quad & \int_{\mathcal{V}} v\,\mathrm{d}F(v) = \mu \quad \cdots\ \alpha \\
& \int_{\mathcal{V}} \mathrm{d}F(v) = 1 \qquad \cdots\ \beta.
\end{aligned}$$

Combining the dual of the above problem with the outer maximization, problem (2) becomes

$$\begin{aligned}
\sup_{x,t,\alpha,\beta} \quad & \mu\alpha+\beta \\
\text{s.t.} \quad & t(v) \ge v\alpha+\beta \quad \forall v\in\mathcal{V} \\
& (x,t)\in\mathcal{M},\ \alpha\in\mathbb{R},\ \beta\in\mathbb{R}
\end{aligned}
\quad\Longleftrightarrow\quad
\begin{aligned}
\sup_{x,\alpha,\beta} \quad & \mu\alpha+\beta \\
\text{s.t.} \quad & \int_0^v u\,\mathrm{d}x(u) \ge v\alpha+\beta \quad \forall v\in\mathcal{V} \quad \cdots\ \mathrm{d}F(v) \\
& \int_{\mathcal{V}} \mathrm{d}x(v) = 1 \qquad \cdots\ \pi \\
& x\in\mathcal{P}_+,\ \alpha\in\mathbb{R},\ \beta\in\mathbb{R},
\end{aligned}
\tag{6}$$

where the right reformulation follows from representation (1) that states $t(v) = vx(v) - \int_0^v x(u)\mathrm{d}u = \int_0^v u\mathrm{d}x(u)$ and the fact that x can be interpreted as a distribution on $\mathcal{V}$. The dual of problem (6), captured by

$$\begin{aligned}
\inf_{F,\pi} \quad & \pi \\
\text{s.t.} \quad & \pi - p\int_p^1 \mathrm{d}F(v) \ge 0 \quad \forall p\in\mathcal{V} \quad \cdots\ \mathrm{d}x(p) \\
& \int_{\mathcal{V}} v\,\mathrm{d}F(v) = \mu \qquad \cdots\ \alpha \\
& \int_{\mathcal{V}} \mathrm{d}F(v) = 1 \qquad \cdots\ \beta \\
& F\in\mathcal{P}_+,\ \pi\in\mathbb{R}_+
\end{aligned}
\quad\Longleftrightarrow\quad
\begin{aligned}
\inf_{F} \quad & \sup_{p\in\mathcal{V}} p\int_p^1 \mathrm{d}F(v) \\
\text{s.t.} \quad & \int_{\mathcal{V}} v\,\mathrm{d}F(v) = \mu \\
& \int_{\mathcal{V}} \mathrm{d}F(v) = 1 \\
& F\in\mathcal{P}_+,
\end{aligned}
\tag{7}$$

is indeed problem (4), verifying Lemma 1 that the dual of a robust screening problem is a minimax pricing problem.

By Theorem 1, $F^\star$ is the worst-case distribution to problem (7). When $v\in[0,\pi^\star)$, since $v\bar{F}^\star(v) < \max_{p\in\mathcal{V}} p\bar{F}^\star(p) = \pi^\star$, the first constraint in the left reformulation of problem (7) is not binding at the optimal solution for $p\in[0,\pi^*)$. Thus by complementary slackness, in problem (6) we must have $\mathrm{d}x^\star(p) = 0$ for $p\in[0,\pi^*)$, implying $x^\star(v) = 0$ and $t^\star(v) = \int_0^v u\,\mathrm{d}x^\star(u) = 0$ for all $v\in[0,\pi^\star)$. On the other hand, $\bar{F}^\star$ in problem (7) is strictly decreasing (*i.e.*, $\mathrm{d}F^\star(v) > 0$) on $[\pi^\star,1]$. By complementary slackness, in problem (6) we must have $\int_0^v u\mathrm{d}x^\star(u) = v\alpha^\star+\beta^\star$ for $v\in[\pi^\star,1]$, yielding $v\mathrm{d}x^\star(v) =$

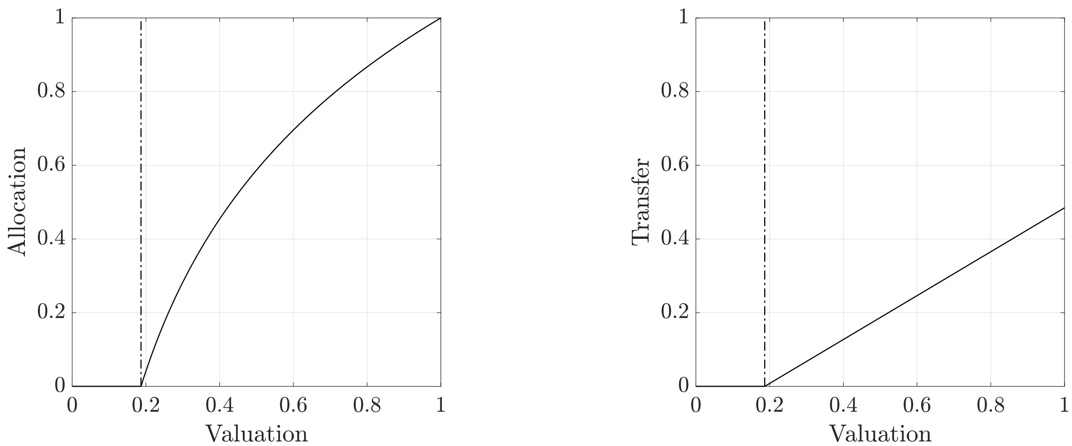

Figure 2 Allocation (left) and transfer (right) as a function of the buyer's valuation ($\mu = 0.5$ in $\mathcal{F}_{\mathrm{M}}$). The threshold $v = \pi^\star$ is indicated by the dash-dotted line.

$\alpha^\star \mathrm{d}v$. It now follows that for $v \in [\pi^\star, 1]$, $x^\star(v) = \int_{\pi^\star}^v \mathrm{d}x^\star(u) = \int_{\pi^\star}^v \frac{\alpha^\star}{u}\,\mathrm{d}u = \alpha^\star(\ln(v) - \ln(\pi^\star))$ and $t^\star(v) = \int_0^v u\,\mathrm{d}x^\star(u) = v\alpha^\star + \beta^\star$. We can solve for the constants $\alpha^\star$ and $\beta^\star$ from the equations:

$$\begin{cases} \displaystyle\int_{\mathcal{V}} \mathrm{d}x^\star(v) = \int_{\pi^\star}^1 \frac{\alpha^\star}{v}\,\mathrm{d}v = 1 \\ \displaystyle\int_0^{\pi^\star} v\,\mathrm{d}x^\star(v) = \pi^\star\alpha^\star + \beta^\star = 0 \end{cases} \quad \Longrightarrow \quad \begin{cases} \alpha^\star = \dfrac{1}{\ln(1/\pi^\star)} = \dfrac{\pi^\star}{(\mu - \pi^\star)} \\ \beta^\star = -\pi^\star\alpha^\star. \end{cases}$$

We summarize the result in the following proposition.

Proposition 2. *Given $\mathcal{F}_{\mathrm{M}}$, the optimal mechanism $(x^\star, t^\star)$ to problem (2) is*

$$(x^\star(v), t^\star(v)) = \begin{cases} (0,\,0) & v \in [0, \pi^\star) \\ (\alpha^\star(\ln(v) - \ln(\pi^\star)),\ \alpha^\star(v - \pi^\star)) & v \in [\pi^\star, 1] \end{cases} \quad \text{with} \quad \alpha^\star = \frac{\pi^\star}{\mu - \pi^\star} > 0.$$

We illustrate in Figure 2 the optimal mechanism derived in Proposition 2. Low value buyers whose valuation falls below $\pi^\star$ are never sold the product, while a high value buyer will purchase a lottery with a price $\alpha^\star(v - \pi^\star)$ that is linear in his valuation and will win the product with probability $\alpha^\star \ln(v/\pi^\star)$ that increases logarithmically with valuation.

A similar geometric approach can be applied to the maximin pricing problem (3). In the inner minimization of problem (3), for a fixed p, one seeks $\bar{F}$ that has the lowest value at p. In particular, with ambiguity set $\mathcal{F}_{\mathrm{M}}$, this is achieved by squeezing as much mass under a CCDF as possible from the area to the right of p to the area to the left of p. This results in a step function representing a two-point (or possibly a degenerated one-point) distribution. (see Figure 3 for an illustration). Again, imagine a rolling machine with the shape of a horizontal level curve applied to compress only the part of a mass μ dirt to the right of p. The optimal value then gives the adversarial selling probability $\bar{\Lambda}^\star$, and the outer maximization reduces to the classic pricing problem $\sup_{p \in \mathcal{V}} p\bar{\Lambda}^\star(p)$,

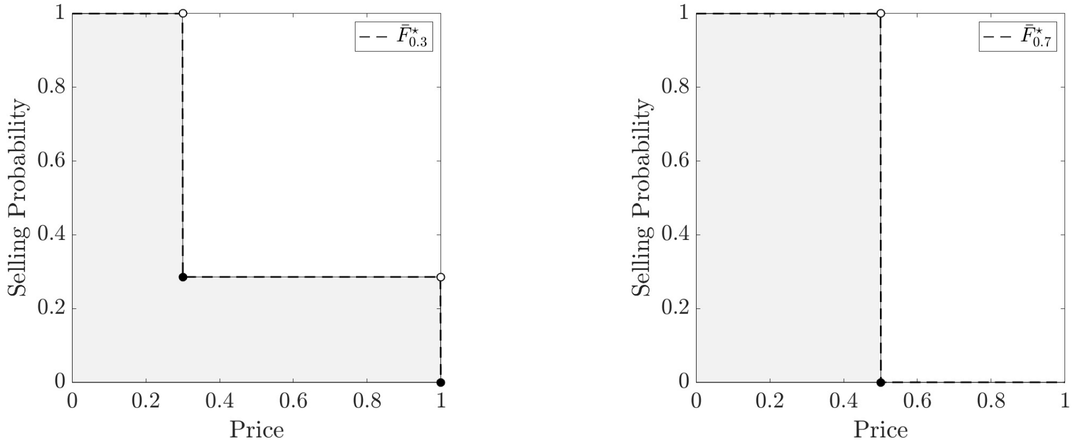

Figure 3 Illustration of the maximin pricing problem ($\mu = 0.5$ in $\mathcal{F}_{\mathrm{M}}$): $p = 0.3$ (left) and $p = 0.7$ (right).

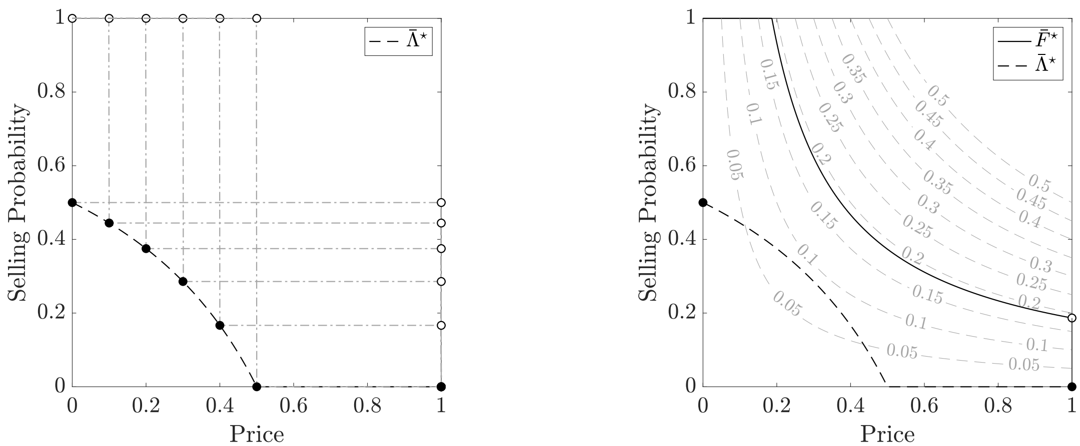

Figure 4 The left panel shows the adversarial selling probability (the dashed line) in the robust pricing problem, where the dash-dot lines indicate the worst-case distributions for a fixed price $p \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$. The right panel compares the robust mechanism design problem and the robust pricing problem ($\mu = 0.5$ in $\mathcal{F}_{\mathrm{M}}$), whose optimal revenues are 0.187 and 0.086, respectively.

whose optimal value is given by the bilinear level curve tangent to $\bar{\Lambda}^{\star}$. Both solutions to the maximin pricing and minimax pricing are illustrated and compared in Figure 4, and we formalize these observations in Proposition 3 below.

Proposition 3. *Given $\mathcal{F}_{\mathrm{M}}$, (i) when $p \in (0, \mu)$, the unique worst-case distribution $F^{\star}_p$ to the problem* $\inf_{F \in \mathcal{F}_{\mathrm{M}}} \bar{F}(p)$ *is*

$$\bar{F}^{\star}_p(v) = \begin{cases} 1 & v \in [0, p) \\ (\mu - p)/(1 - p) & v \in [p, 1) \\ 0 & v = 1; \end{cases}$$

when $p \in [\mu, 1]$*, a worst-case distribution to the problem* $\inf_{F \in \mathcal{F}_{\mathrm{M}}} \bar{F}(p)$ *is the Dirac measure at* μ*,* δ_μ*. (ii) The adversarial selling probability is* $\bar{\Lambda}^\star(p) = (\mu - p)^+/(1-p)\ \forall p \in [0,1]$*, and the optimal posted price is* $p^\star = 1 - \sqrt{1-\mu}$*. (iii) The performance ratio is* $\eta = (1 - \sqrt{1-\mu})^2/\pi^\star$*, which is increasing on* $\mu \in (0,1]$*, approaches* 0 *when* $\mu \to 0$*, and becomes* 1 *when* $\mu = 1$*.*

We conclude this subsection by making several remarks on the related literature. First, our Theorem 1 recovers theorem 1 in Elmachtoub et al. (2020). The proof, however, is quite different. Elmachtoub et al. (2020) first use what they call "pricing inequality" to obtain a lower bound on the optimal value of problem (4). They then show that the proposed lower bound is tight by constructing an $F \in \mathcal{F}_{\mathrm{M}}$ (in fact our $F^\star$) that achieves the lower bound. In comparison, we construct $F^\star$ based on our geometric insight and we prove its optimality as well as *uniqueness* by contradiction. Like Elmachtoub et al. (2020), our approach also generalizes to the more complicated ambiguity set that incorporates additional mean absolute deviation information, and we recover theorem 2 in Elmachtoub et al. (2020); see Theorem 6 and detailed discussions in Appendix C.2, where we have also solved the primal problem (2) and the robust pricing problem (3).

Both Proposition 2 and Proposition 3 can be found in Carrasco et al. (2018, propositions 5 and 6). As we mentioned in Section 2, Carrasco et al. (2018) obtained the optimal mechanism by analyzing the primal problem. Our analysis, however, starts from the dual problem of minimax pricing, which also allows us to graphically compare with the solution of maximin pricing problem (see Figure 4).[10] For ambiguity set with variance information, Carrasco et al. (2018) also analytically solve the primal problem while Elmachtoub et al. (2020) provide an explicit optimization formulation to upper bound the dual problem. We refer readers to Appendix C.1 for our solutions to both the primal and dual as well as the robust pricing problem when variance information is incorporated.

The worst-case distribution $F^\star$ in Theorem 1 (or some of its variants) is also known as the equal-revenue distribution. Roesler and Szentes (2017, lemma 1) also observe that given any valuation distribution one can construct an equal-revenue distribution from the tangent level curve that achieves the same optimal revenue. With a more general ambiguity set based on mean-preserving contraction, the authors use the observation to construct a dominating distribution for the buyer. We next discuss this ambiguity set and connect with results in Roesler and Szentes (2017).

4.2. Mean-Preserving Contraction Ambiguity Set

Given two distributions F and G, F is said to be a mean-preserving contraction of G (or equivalently, G is a mean-preserving spread of F) if

$$\int_{v \in \mathcal{V}} v \,\mathrm{d}F(v) = \int_{v \in \mathcal{V}} v \,\mathrm{d}G(v) = \mu \quad \text{and} \quad \int_0^u \bar{F}(v)\,\mathrm{d}v \geq \int_0^u \bar{G}(v)\,\mathrm{d}v \ \forall u \in \mathcal{V}.$$

[10] Although not central to our geometric approach, we present Proposition 3 (*iii*) (or proposition 6 in Carrasco et al. 2018) for completeness, and we simplify significantly the proof that the ratio $\eta \to 0$ as $\mu \to 0$ (the proof in Carrasco et al. 2018 spans around two pages).

Geometrically, F is a mean-preserving contraction of G if for all $u \in \mathcal{V}$, the area under $\bar{F}$ from 0 to u is no smaller than that under $\bar{G}$ from 0 to u and these two areas are equal when $u = 1$. The mean-preserving contraction ambiguity set is defined by all mean-preserving contractions of G:

$$\mathcal{F}_{\mathrm{MPC}} = \left\{ F \in \mathcal{P} \;\middle|\; \int_{v\in\mathcal{V}} v \,\mathrm{d}F(v) = \int_{v\in\mathcal{V}} v \,\mathrm{d}G(v) = \mu,\ \int_0^u \bar{F}(v)\,\mathrm{d}v \geq \int_0^u \bar{G}(v)\,\mathrm{d}v \ \forall u \in \mathcal{V} \right\}.$$

In particular, if the reference distribution G is a Bernoulli distribution, then any distribution with the same mean is a mean-preserving contraction of G, and $\mathcal{F}_{\mathrm{MPC}}$ reduces to $\mathcal{F}_{\mathrm{M}}$.

Clearly, if the worst-case distribution for the mean ambiguity set characterized in Theorem 1—for convenience we denote it by Q_1—is already a mean-preserving contraction of G, then Q_1 is the worst-case distribution for $\mathcal{F}_{\mathrm{MPC}}$ as well. However, if $Q_1 \notin \mathcal{F}_{\mathrm{MPC}}$, then one can find some $u \in \mathcal{V}$ such that $\int_0^u \bar{Q}_1(v)\mathrm{d}v < \int_0^u \bar{G}(v)\mathrm{d}v$. In such a case, intuitively, one can move all the mass under $\bar{Q}_1$ to the right of u and spread it over the interval $[0,u]$ so that $\int_0^u \bar{Q}_1(v)\mathrm{d}v$ can be increased; see Figure 5. This observation motivates us to consider a parametric family of distributions $\{Q_k\}_k$ in the form

$$\bar{Q}_k(v) = \begin{cases} 1 & v \in [0, \pi_k) \\ \pi_k/v & v \in [\pi_k, k) \\ 0 & v \in [k, 1]. \end{cases}$$

Here, $k \in [\mu, 1]$ and π_k is the unique solution in $[0,1]$ to the equation $\pi + \pi \ln(k/\pi) = \mu$ such that the mean of $\bar{Q}_k$ is μ.[11] Note that for any $k_1, k_2 \in [\mu, 1]$ such that $k_1 \geq k_2$, Q_{k_2} is a mean-preserving contraction of Q_{k_1} with $\bar{Q}_1$ being the least contracting distribution and $\bar{Q}_\mu$—a point mass at μ—being the most contracting distribution. Since the optimal revenue π_k under $\bar{Q}_k$ is decreasing in k, the worst-case revenue within this family can then be found by the largest k such that Q_k is a mean-preserving contraction of G. The next proposition shows that it is also the worst-case distribution in $\mathcal{F}_{\mathrm{MPC}}$.

Theorem 2. *Given $\mathcal{F}_{\mathrm{MPC}}$, the worst-case distribution to problem (4) is $F^\star = Q_{k^\star}$ with $k^\star = \max\{k \in [\mu, 1] \mid Q_k \in \mathcal{F}_{\mathrm{MPC}}\}$. The optimal revenue is $\pi^\star = \pi_{k^\star}$.*

Corresponding to the dual solution, we construct the optimal solution to the primal problem.

Proposition 4. *Given $\mathcal{F}_{\mathrm{MPC}}$, the optimal mechanism $(x^\star, t^\star)$ to problem (2) is*

$$(x^\star(v), t^\star(v)) = \begin{cases} (0,\ 0) & v \in [0, \pi^\star) \\ (\gamma^\star(\ln(v) - \ln(\pi^\star)),\ \gamma^\star(v - \pi^\star)) & v \in [\pi^\star, u^\star) \\ (1,\ \gamma^\star(u^\star - \pi^\star)) & v \in [u^\star, 1], \end{cases}$$

where $u^\star \in [\pi^\star, k^\star]$ satisfies $\int_0^{u^\star} (\bar{F}^\star(v) - \bar{G}(v))\mathrm{d}v = 0$ and $\gamma^\star = 1/\ln(u^\star/\pi^\star) > 0$.

[11] Using the Lambert W function, we have $\pi_k = -\mu/W(-\mu/(ke))$.

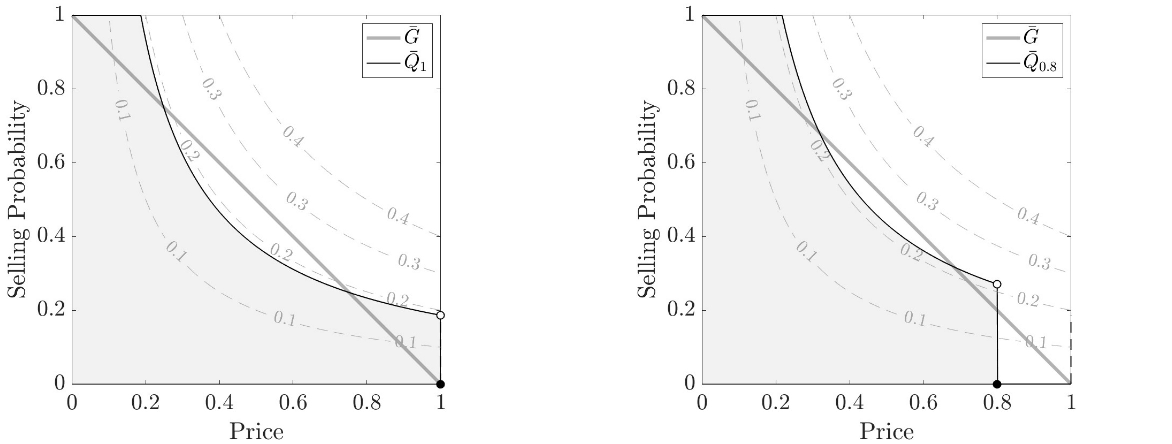

Figure 5 The worst-case distribution $\bar{Q}_1$ for $\mathcal{F}_{\mathrm{M}}$ is not in $\mathcal{F}_{\mathrm{MPC}}$. Here, $\bar{Q}_{0.8}$, a mean-preserving contraction of $\bar{Q}_1$, is formed by moving all the mass above 0.8 and spreading it along the level curves below 0.8.

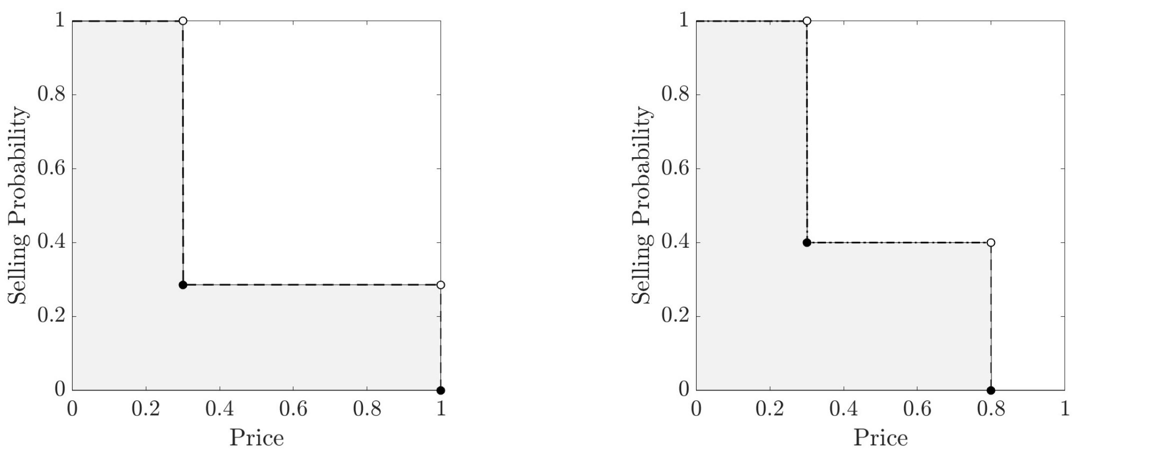

Figure 6 Illustration of the maximin pricing problem ($\mu = 0.5$ in $\mathcal{F}_{\mathrm{MPC}}$ and $p = 0.3$). The distribution on the left panel has a lower worst-case selling probability but is more spreading than that on the right panel.

Again, a similar geometric approach can be applied to the maximin pricing problem (3). For any fixed price $p \in \mathcal{V}$, if $\bar{F}_p^\star \in \arg\min_{F \in \mathcal{F}_{\mathrm{M}}} \bar{F}(p)$ belongs to $\mathcal{F}_{\mathrm{MPC}}$, then $\bar{F}_p^\star$ also solves $\inf_{F \in \mathcal{F}_{\mathrm{MPC}}} \bar{F}(p)$. Otherwise, one could squeeze the mass between $[p, 1]$ to a smaller interval $[p, k]$ with some $k < 1$ and obtain a more contracting distribution than $\bar{F}_p^\star$; see an illustration in Figure 6 for some $p < \mu$. Similarly, given $p \in [0, \mu)$, we consider the family

$$\bar{R}_{p,k}(v) = \begin{cases} 1 & v \in [0, p) \\ (\mu - p)/(k - p) & v \in [p, k) \\ 0 & v \in [k, 1] \end{cases}$$

with $k \geq p$. The worst-case distribution under a posted price mechanism is characterized below.

Proposition 5. *Given $\mathcal{F}_{\text{MPC}}$, a worst-case distribution $F_p^\star$ to the problem $\inf_{F\in\mathcal{F}_{\text{MPC}}}\bar{F}(p)$ is $\bar{R}_{p,k_p^\star}$ with $k_p^\star=\max\{k\in[p,1]\mid \bar{R}_{p,k}\in\mathcal{F}_{\text{MPC}}\}$ when $p\in[0,\mu)$; and is the Dirac measure at μ, δ_μ, when $p\in[\mu,1]$. Correspondingly, $\bar{\Lambda}^\star(p)=(\mu-p)/(k_p^\star-p)$ for $p\in[0,\mu)$ and $\bar{\Lambda}^\star(q)=0$ otherwise.*

In general, $k_p^\star$ in Proposition 5 may not admit a closed-form expression and one can only compute the optimal posted price numerically. When the reference distribution G is, for example, uniform, it is possible to analytically characterize $k_p^\star$ and the optimal posted price.

We conclude this section by providing a condition, which includes both $\mathcal{F}_{\text{M}}$ and $\mathcal{F}_{\text{MPC}}$ as special cases, to ensure that the solution to problem (4) also maximizes the buyer's payoff. To this end, we first introduce the notion of efficiency as follows. Given an ambiguity set $\mathcal{F}$ with a known mean μ, let $F^\star$ and $\pi^\star$ be respectively the optimal solution and optimal value to problem (4). We say $F^\star$ is *efficient* if there exists $p_0\in\arg\max_{p\in\mathcal{V}} p\bar{F}_-^\star(p)$ such that $\mathbb{E}_{F^\star}[(\tilde{v}-p_0)^+]=\mu-\pi^\star$. Clearly, μ is the first best social welfare, and efficiency of $F^\star$ says that under $F^\star$, there exists an optimal price for the seller such that the social welfare (seller's profit plus buyer's payoff) achieves the first best value. Note that both the worst-case distributions under $\mathcal{F}_{\text{M}}$ and $\mathcal{F}_{\text{MPC}}$ are efficient.

Theorem 3. *Given an ambiguity set $\mathcal{F}$ with a known mean μ, if the worst-case distribution $F^\star$ to problem* (4) *is efficient, then $F^\star$ is also optimal to the buyer-optimal pricing problem* (5).

Theorems 2 and 3 together recover theorem 1 in Roesler and Szentes (2017), who first establish that $Q_{k^\star}$ maximizes buyer's surplus and then note that it simultaneously minimizes the seller's revenue as well. Note that Theorem 2 also directly provides a new upper bound $\mu/\pi^\star$ on the value of personalized pricing studied in Elmachtoub et al. (2020). Proposition 4 can also be found in Du (2018) (see proposition 1 therein), who derived the optimal mechanism as a special case of the proposed exponential auction for the more general multi-buyer problem. Our characterization of the worst-case distribution for the posted price mechanism in Proposition 5, however, is new.

4.3. Wasserstein Ambiguity Set

For any $\rho\in[1,\infty]$, the type-ρ Wasserstein distance between two distributions F and G is[12]

$$d_\rho(F,G)=\|\bar{F}^{-1}-\bar{G}^{-1}\|_\rho=\left(\int_0^1|\bar{F}^{-1}(q)-\bar{G}^{-1}(q)|^\rho\,\mathrm{d}q\right)^{1/\rho}.$$

Given a reference distribution G, the type-ρ Wasserstein ambiguity set is then defined as the set of all distributions, whose type-ρ Wasserstein distance to G is bounded by $\theta>0$: [13]

$$\mathcal{F}_{(\theta,\rho)}=\left\{F\in\mathcal{P}\;\middle|\;\|\bar{F}^{-1}-\bar{G}^{-1}\|_\rho\le\theta\right\}.$$

[12] Note that the standard definition of the type-ρ Wasserstein distance is through an optimal transport problem. However, we choose to present the geometrically more intuitive form here. Indeed, one can show that the definition is equivalent to the standard one when the distributions are one-dimensional (see Appendix C.5 for detailed discussions).

[13] When $\theta=0$, $\mathcal{F}_{(\theta,\rho)}=\{G\}$ because $d_\rho(F,G)=0$. Thus the posted price mechanism is optimal.

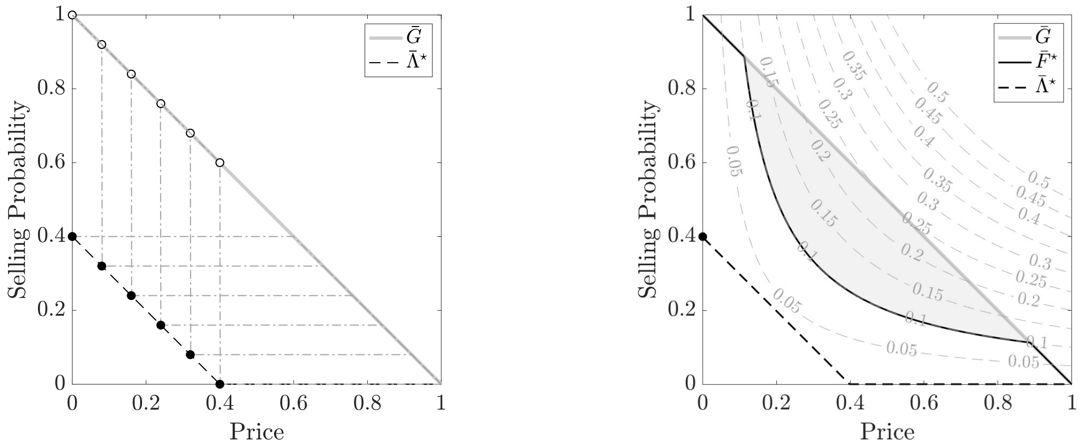

Figure 7 The reference distribution G is uniform and the Wasserstein radius is $\theta = 0.18$. The optimal revenues of robust mechanism and pricing are 0.1 and 0.04, respectively.

In particular, $\mathcal{F}_{(\theta,1)}$ is constituted of all $\bar{F}$ with the area between $\bar{F}$ and $\bar{G}$ being bounded by θ, while $\mathcal{F}_{(\theta,\infty)}$ includes all $\bar{F}$ whose horizontal distance to $\bar{G}$ is bounded by θ.

The worst-case distribution for problem (4) can be best illustrated when $\rho = 1$. Suppose the mean of G is μ and consider a subset of $\mathcal{F}_{(\theta,1)}$ that contains all $\bar{F}$ that are also first-order stochastically dominated by $\bar{G}$ (*i.e.*, $\bar{F} \le \bar{G}$). Such a subset can then be equivalently characterized as all distributions $\bar{F}$ under $\bar{G}$ with an area under $\bar{F}$ being larger than $\mu - \theta$. The same geometric intuition we used for the mean ambiguity set can then be applied: one compresses a pile of dirt with mass more than $\mu - \theta$ under the curve $\bar{G}$ using the bilinear level curves (see the right panel in Figure 7 for an illustration). We next formally characterize the worst-case distribution to problem (4) for $\mathcal{F}_{(\theta,\rho)}$.

THEOREM 4. *Given $\mathcal{F}_{(\theta,\rho)}$ with $\rho \in [1,\infty]$, the worst-case distribution $F^\star$ to problem (4) is*

$$\bar{F}^\star(v) = \min\left\{\bar{G}(v), \pi^\star/v\right\} \quad \forall v \in \mathcal{V},$$

which is unique when ρ is finite, and the optimal revenue $\pi^\star$ is the unique solution in $[0,1]$ to

$$\left(\int_0^1 ((\bar{G}^{-1}(q) - \pi/q)^+)^\rho \mathrm{d}q\right)^{1/\rho} = \theta.$$

In the following, we focus on the case of $\mathcal{F}_{(\theta,1)}$ in characterizing the optimal solutions to problems (2) and (3). We comment on the general case of $\mathcal{F}_{(\theta,\rho)}$ at the end of the section.

Following Theorem 4, we further assume that the reference distribution G is such that there are finitely many disjoint intervals: $[u_j, w_j)$, $j \in [J]$ on which $\bar{G}(v) > \pi^\star/v$. This assumption is satisfied, for example, when G is an empirical distribution. In addition, we denote $u_0 = w_0 = 0$ and $u_{J+1} = 1$. The optimal solution to problem (2) is characterized below.

PROPOSITION 6. *Given $\mathcal{F}_{(\theta,1)}$, the optimal mechanism $(x^\star, t^\star)$ to problem (2) is*

$$(x^\star(v), t^\star(v)) = \begin{cases} \alpha^\star \left(\ln\left(\frac{v}{w_j}\right) + \sum_{i=1}^{j-1} \ln\left(\frac{w_i}{u_i}\right), \ (v - u_j) + \sum_{i=1}^{j-1} (w_i - u_i) \right) & v \in [u_j, w_j),\ j \in [J] \\ \alpha^\star \left(\sum_{i=1}^{j} \ln\left(\frac{w_i}{u_i}\right), \ \sum_{i=1}^{j} (w_i - u_i) \right) & v \in [w_j, u_{j+1}),\ j \in [J], \end{cases}$$

where $\alpha^\star = 1/\big(\sum_{j\in[J]} \ln(w_j/u_j)\big) > 0$ and $(x^\star(1), t^\star(1)) = (1, \alpha^\star \sum_{j\in[J]}(w_j - u_j))$.

Together with Theorem 4, Proposition 6 shows that lotteries with logarithmic winning probabilities paired with linear prices are offered to valuations under which the worst-case distribution deviates from the reference distribution (*i.e.*, $\bar{F}^\star(v) < \bar{G}(v)$); while bunching occurs for other valuations. Li et al. (2019) also characterize the optimal mechanism for the type-1 Wasserstein ambiguity set when $\mathcal{V}$ is restricted to discrete valuations (see theorem 3.1) and they provide a conjecture for the continuous valuation case (see their informal theorem 1). Their conjecture corresponds to $J = 1$ in Proposition 6 wherein we formally establish that the optimal mechanism can be more complicated.

We next characterize the worst-case distribution under a given posted price mechanism.

PROPOSITION 7. *Given $\mathcal{F}_{(\theta,1)}$, for any $p \in (0,1]$, the unique worst-case distribution $F_p^\star$ to the problem $\inf_{F\in\mathcal{F}_{(\theta,1)}} \bar{F}(p)$ is*

$$\bar{F}_p^\star(v) = \begin{cases} \bar{G}(v) & v \in [0,p) \\ \bar{G}(k_p^\star) & v \in [p, k_p^\star) \\ \bar{G}(v) & v \in [k_p^\star, 1] \end{cases} \quad \text{with} \quad k_p^\star = \max\left\{ k \in [p,1] \ \middle| \ \int_{\bar{G}(k)}^{\bar{G}(p)} (\bar{G}^{-1}(q) - p)\mathrm{d}q \le \theta \right\}.$$

By Proposition 7, the optimal posted price mechanism can be obtained by solving $\sup_{p\in\mathcal{V}} p\bar{\Lambda}^\star(p) = \sup_{p\in\mathcal{V}} p\bar{G}(k_p^\star)$. The following example illustrates both Theorem 4 and Proposition 7.

EXAMPLE 2 (UNIFORM REFERENCE DISTRIBUTION). *Given $\mathcal{F}_{(1,\theta)}$ with $\theta \le 1/2$ and a uniform distribution G such that $G(v) = v$ for all $v \in \mathcal{V}$, the worst-case distribution $F^\star$ to problem (4) is*

$$\bar{F}^\star(v) = \begin{cases} 1 - v & v \in [0,u) \\ \pi^\star / v & v \in [u,w) \\ 1 - v & v \in [w,1] \end{cases} \quad \text{with} \quad u = \frac{1 - \sqrt{1 - 4\pi^\star}}{2} \quad \text{and} \quad w = \frac{1 + \sqrt{1 - 4\pi^\star}}{2},$$

and $\pi^\star$ is the unique solution in $[0,1]$ to the equation

$$\frac{\sqrt{1-4\pi}}{2} - \pi \ln\left(\frac{1 + \sqrt{1-4\pi}}{1 - \sqrt{1-4\pi}} \right) = \theta.$$

The adversarial selling probability is $\bar{\Lambda}^\star(p) = (1 - \sqrt{2\theta} - p)^+ \ \forall p \in [0,1]$, and the optimal posted price $p^\star = \frac{1-\sqrt{2\theta}}{2}$. It follows that

$$\eta = \frac{p^\star \bar{\Lambda}^\star(p^\star)}{\pi^\star} = \frac{(1 - \sqrt{2\theta})^2}{4\pi^\star}.$$

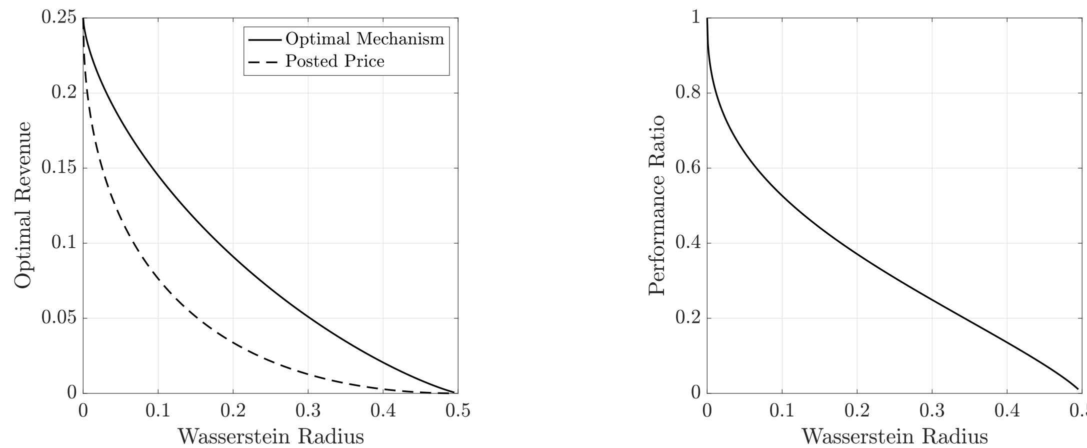

Figure 8 Optimal revenues of the minimax and maximin pricing problems (left) and performance ratio (right). The reference distribution G is uniform.

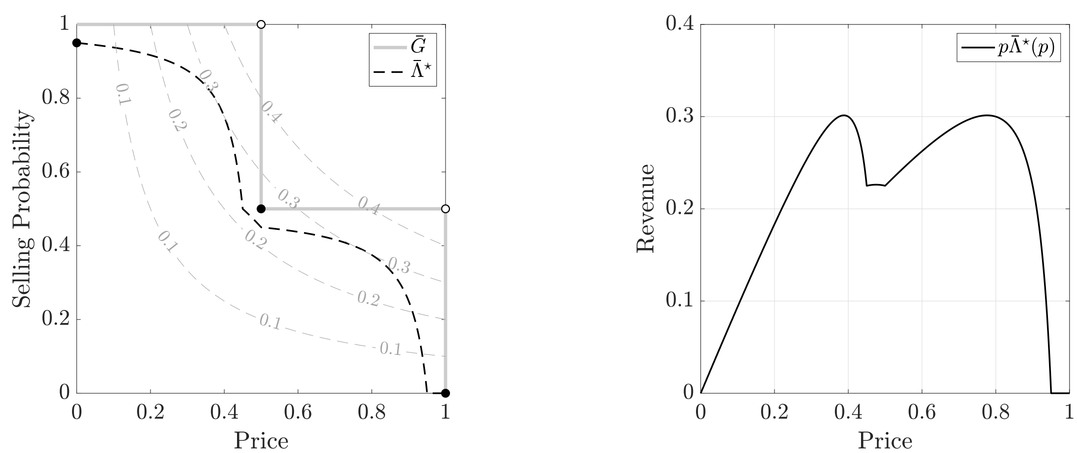

Figure 9 The reference distribution G is an empirical distribution on $v_1 = 0.5$ and $v_2 = 1$, and the radius $\theta = 0.025$. The revenue function under the adversarial selling probability is not unimodal (right).

Both solutions to problems (3) *and* (4) *are illustrated in Figure 7, and the optimal revenue for problems* (3) *and* (4) *as well as the performance ratio* η *are plotted as a function of* θ *in Figure 8. It can be seen that as ambiguity increases, the worst-case revenues under the optimal mechanism and posted price mechanism decrease; yet the value of employing the optimal mechanism increases.*

We note that in general, however, the adversarial selling probability $\Lambda^\star$ may not be well behaved and there is no general closed-form solution to the optimal posted price. In Figure 9, we illustrate such an instance where the revenue function under the adversarial selling probability is not unimodal.

Although our characterization for the worst-case distribution in Theorem 4 holds for any $\rho \geq 1$, we do not have a complete characterization of the optimal mechanism for $1 < \rho < \infty$.[14] We conjecture that bunching still occurs for v such that $\bar{G}(v) \leq \pi^\star/v$ but both the allocation probabilities and the transfer functions may involve higher order polynomials when $\bar{G}(v) > \pi^\star/v$. At the other end of the spectrum when $\rho = \infty$, however, one can use Proposition 1 to show that a posted price mechanism is optimal; see the corollary below whose proof is omitted.

Corollary 1. *Given $\mathcal{F}_{(\theta,\infty)}$, the performance ratio $\eta = 1$ and a posted price mechanism with $p^\star \in \arg\max_{p\in\mathcal{V}} p\bar{\Lambda}^\star_-(p)$, where $\bar{\Lambda}^\star(p) = \bar{G}(p+\theta)$ $\forall p \in \mathcal{V}$, is optimal.*

5. Conclusion

When the mean of an ambiguity set $\mathcal{F}$ is known (*e.g.*, $\mathcal{F}_{\mathrm{M}}$ and $\mathcal{F}_{\mathrm{MPC}}$), our approach can be directly applied to the minimax absolute (or relative) ex-post regret criterion that solves

$$\inf_{(x,t)\in\mathcal{M}} \sup_{F\in\mathcal{F}} \left\{\mathbb{E}_F[\tilde{v}] - \mathbb{E}_F[t(\tilde{v})]\right\} \quad \left(\text{or} \inf_{(x,t)\in\mathcal{M}} \sup_{F\in\mathcal{F}} \frac{\mathbb{E}_F[\tilde{v}]}{\mathbb{E}_F[t(\tilde{v})]}\right).$$

However, the above problems remain open when the mean is not fixed: *e.g.*, $\mathcal{F}_{(\theta,\rho)}$. Apart from alternative decision criteria, there are many other promising venues for future research. Although many recent works have studied robust mechanism design in the more demanding multi-product or multi-buyer settings, most of the papers focus on specific ambiguity sets such as the one with moment information. Problems with probability distance based ambiguity set, again remain largely open. Also, most of the optimal mechanisms we derived in the paper involve offering a continuum menu of lotteries, which may not be easy to implement in practice. A relevant problem is to find an approximation of the optimal mechanism that has a good performance guarantee.

References

Aliprantis, Charalambos, Kim Border. 2006. *Infinite dimensional analysis*. Springer.

Allouah, Amine, Achraf Bahamou, Omar Besbes. 2021. Pricing with samples. *SSRN 3334650*.

Balcan, Maria-Florina, Tuomas Sandholm, Ellen Vitercik. 2017. Generalization guarantees for multi-item profit maximization: Pricing, auctions, and randomized mechanisms *arXiv preprint arXiv:1705.00243*.

Bandi, Chaithanya, Dimitris Bertsimas. 2014. Optimal design for multi-item auctions: A robust optimization approach. *Mathematics of Operations Research* **39**(4) 1012–1038.

Bei, Xiaohui, Ning Chen, Nick Gravin, Pinyan Lu. 2017. Worst-case mechanism design via bayesian analysis. *SIAM Journal on Computing* **46**(4) 1428–1448.

Bergemann, Dirk, Francisco Castro, Gabriel Weintraub. 2020. Third-degree price discrimination versus uniform pricing. *SSRN 3540156*.

[14] Proposition 7, however, can still be straightforwardly generalized to the case of any (possibly infinite) $\rho > 1$.

Bergemann, Dirk, Karl Schlag. 2008. Pricing without priors. *Journal of the European Economic Association* **6**(2-3) 560–569.

Bergemann, Dirk, Karl Schlag. 2011. Robust monopoly pricing. *Journal of Economic Theory* **146**(6) 2527–2543.

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

Börgers, Tilman. 2015. *An introduction to the theory of mechanism design*. Oxford University Press.

Borwein, Jonathan, D Zhuang. 1986. On Fan's minimax theorem. *Mathematical Programming* **34**(2) 232–234.

Cai, Yang, Nikhil R Devanur, S Matthew Weinberg. 2019. A duality-based unified approach to bayesian mechanism design. *SIAM Journal on Computing* **50**(3) STOC16–160–STOC16–200.

Carrasco, Vinicius, Vitor Farinha Luz, Nenad Kos, Matthias Messner, Paulo Monteiro, Humberto Moreira. 2018. Optimal selling mechanisms under moment conditions. *Journal of Economic Theory* **177** 245–279.

Carroll, Gabriel. 2017. Robustness and separation in multidimensional screening. *Econometrica* **85**(2) 453–488.

Chen, Hongqiao, Ming Hu, Georgia Perakis. 2020a. Distribution-free pricing. *SSRN 3090002*.

Chen, Zhi, Melvyn Sim, Peng Xiong. 2020b. Robust stochastic optimization made easy with RSOME. *Management Science* **66**(8) 3329–3339.

Cole, Richard, Tim Roughgarden. 2014. The sample complexity of revenue maximization. *Proceedings of the Forty-Sixth Annual ACM symposium on Theory of Computing*. 243–252.

Condorelli, Daniele, Balázs Szentes. 2020. Information design in the holdup problem. *Journal of Political Economy* **128**(2) 681–709.

Corless, Robert, Gaston Gonnet, David Hare, David Jeffrey, Donald Knuth. 1996. On the Lambert W function. *Advances in computational mathematics* **5**(1) 329–359.

Daskalakis, Constantinos, Alan Deckelbaum, Christos Tzamos. 2017. Strong duality for a multiple-good monopolist. *Econometrica* **85**(3) 735–767.

Delage, Erick, Yinyu Ye. 2010. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research* **58**(3) 595–612.

Du, Songzi. 2018. Robust mechanisms under common valuation. *Econometrica* **86**(5) 1569–1588.

Elmachtoub, Adam, Vishal Gupta, Michael Hamilton. 2020. The value of personalized pricing. *Management Science*.

Gao, Rui, Anton Kleywegt. 2016. Distributionally robust stochastic optimization with Wasserstein distance. *arXiv preprint arXiv:1604.02199*.

Giannakopoulos, Yiannis, Elias Koutsoupias. 2014. Duality and optimality of auctions for uniform distributions. *Proceedings of the fifteenth ACM conference on Economics and computation*. 259–276.

Gravin, Nick, Pinyan Lu. 2018. Separation in correlation-robust monopolist problem with budget. *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. 2069–2080.

Kjeldsen, Tinne Hoff. 2001. John von Neumann's conception of the minimax theorem: A journey through different mathematical contexts. *Archive for history of exact sciences* **56**(1) 39–68.

Koçyiğit, Çağıl, Garud Iyengar, Daniel Kuhn, Wolfram Wiesemann. 2019. Distributionally robust mechanism design. *Management Science* **66**(1) 159–189.

Koçyiğit, Çağıl, Napat Rujeerapaiboon, Daniel Kuhn. 2021. Robust multidimensional pricing: Separation without regret. *Mathematical Programming* 1–34.

Li, Yingkai, Pinyan Lu, Haoran Ye. 2019. Revenue maximization with imprecise distribution. *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. 1582–1590.

Luenberger, David. 1997. *Optimization by vector space methods*. John Wiley & Sons.

Manelli, Alejandro, Daniel Vincent. 2007. Multidimensional mechanism design: Revenue maximization and the multiple-good monopoly. *Journal of Economic Theory* **137**(1) 153–185.

Mohajerin Esfahani, Peyman, Daniel Kuhn. 2018. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming* **171**(1-2) 115–166.

Myerson, Roger. 1981. Optimal auction design. *Mathematics of Operations Research* **6**(1) 58–73.

Pınar, Mustafa Ç, Can Kızılkale. 2017. Robust screening under ambiguity. *Mathematical Programming* **163**(1-2) 273–299.

Rahimian, Hamed, Sanjay Mehrotra. 2019. Distributionally robust optimization: A review. *arXiv preprint arXiv:1908.05659*.

Reny, Philip. 1999. On the existence of pure and mixed strategy Nash equilibria in discontinuous games. *Econometrica* **67**(5) 1029–1056.

Riley, John, Richard Zeckhauser. 1983. Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics* **98**(2) 267–289.

Roesler, Anne-Katrin, Balázs Szentes. 2017. Buyer-optimal learning and monopoly pricing. *American Economic Review* **107**(7) 2072–80.

Roos, Ernst, Ruud Brekelmans, Wouter van Eekelen, Dick den Hertog, Johan van Leeuwaarden. 2020. Tight tail probability bounds for distribution-free decision making. *arXiv preprint arXiv:2010.07784*.

Rothschild, Michael, Joseph Stiglitz. 1970. Increasing risk: I. A definition. *Journal of Economic Theory* **2**(3) 225–243.

Santambrogio, Filippo. 2015. *Optimal transport for applied mathematicians*. Springer.

Suzdaltsev, Alex. 2020a. Distributionally robust pricing in independent private value auctions. *arXiv preprint arXiv:2008.01618*.

Suzdaltsev, Alex. 2020b. An optimal distributionally robust auction. *arXiv preprint arXiv:2006.05192*.

Talluri, Kalyan, Garrett van Ryzin. 2006. *The theory and practice of revenue management*. Springer.

Vohra, Rakesh V. 2011. *Mechanism design: a linear programming approach*, vol. 47. Cambridge University Press.

Wang, Shixin, Shaoxuan Liu, Jiawei Zhang. 2020. Minimax regret mechanism design with moments information. *SSRN 3707021*.

Wiesemann, Wolfram, Daniel Kuhn, Melvyn Sim. 2014. Distributionally robust convex optimization. *Operations Research* **62**(6) 1358–1376.

Wolitzky, Alexander. 2016. Mechanism design with maxmin agents: Theory and an application to bilateral trade. *Theoretical Economics* **11**(3) 971–1004.

Yao, Andrew Chi-Chin. 1977. Probabilistic computations: Toward a unified measure of complexity. *18th Annual Symposium on Foundations of Computer Science*. 222–227.

Appendix A Technical Lemmas

Lemma 2. *Consider any $\mathcal{F} \subseteq \mathcal{P}$ and $F \in \mathcal{F}$. For each fixed p (resp., F), the functional $p\bar{F}_-(p)$ is upper semi-continuous in F (resp., p) and the functional $p\bar{F}(p)$ is lower semi-continuous in F (resp., p). In addition, we have (i) for the maximin pricing problem* (3),

$$\sup_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}_-(p) = \sup_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}(p),$$

where the supremum on the left can be attained; and (ii) for the minimax pricing problem (4),

$$\inf_{F\in\mathcal{F}} \sup_{p\in\mathcal{V}} p\bar{F}_-(p) = \inf_{F\in\mathcal{F}} \sup_{p\in\mathcal{V}} p\bar{F}(p),$$

where the supremum on the left can be attained.

Proof of Lemma 2. We focus on $p\bar{F}_-(p)$ and the proof of $p\bar{F}(p)$ is similar. Given F, the upper semi-continuity of $p\bar{F}_-(p)$ in p follows from definition. We next prove the upper semi-continuity in F. The indicator function $\mathbb{I}$, defined through $\mathbb{I}(v \geq p) = 1$ if and only if $v \geq p$, is upper semi-continuous in v for each fixed p. Since $\bar{F}_-(p) = \int_{\mathcal{V}} \mathbb{I}(v \geq p)\mathrm{d}F(v)$, using theorem 14.5 in Aliprantis and Border (2006), $\bar{F}_-(p)$ is then upper semi-continuous in F for each fixed q, and so is $p\bar{F}_-(p)$.

In view of (i), as $\bar{F}_-(p) \geq \bar{F}(p)$ for any $p \in \mathcal{V}$, we have $a := \sup_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}_-(p) \geq \sup_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}(p) =: b \geq 0$. Since $a = 0$ implies $a = b$, it is sufficient to focus on $a > 0$. By definition of supremum, for any $\varepsilon \in (0, a)$, there exists $p \in \mathcal{V}$ such that $\inf_{F\in\mathcal{F}} p\bar{F}_-(p) > a - \varepsilon$. Since $a > 0$ and $\varepsilon \in (0, a)$, it holds that $p > 0$. Choosing a small $\delta \in (0, \min\{p, \varepsilon\})$, we have $b \geq \inf_{F\in\mathcal{F}} (p-\delta)\bar{F}(p-\delta) = p \inf_{F\in\mathcal{F}} \bar{F}(p-\delta) - \delta \inf_{F\in\mathcal{F}} \bar{F}(p-\delta) > p \inf_{F\in\mathcal{F}} \bar{F}(p-\delta) - \varepsilon$, where the strict inequality follows from $0 < \delta < \varepsilon$ and $\bar{F}(p-\delta) \leq 1$. Besides, for the chosen $\delta \in (0, \min\{p, \varepsilon\})$, it holds that $\bar{F}(p-\delta) \geq \bar{F}_-(p)$ for any $F \in \mathcal{F}$. As a result, we have $b > \inf_{F\in\mathcal{F}} p\bar{F}(p-\delta) - \varepsilon \geq \inf_{F\in\mathcal{F}} p\bar{F}_-(p) - \varepsilon > a - 2\varepsilon$. As ε is arbitrarily chosen, we must have $b \geq a$. This, together with $a \geq b$, concludes $a = b$. We note that $\inf_{F\in\mathcal{F}} p\bar{F}_-(p)$, a pointwise infimum of upper semi-continuous functions in p, is also upper semi-continuous in p. Since $\mathcal{V} = [0, 1]$ is compact, by Weierstrass extreme value theorem (see, *e.g.*, Luenberger 1997, theorem 1, p.40), we have $\sup_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}_-(p) = \max_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}_-(p)$.

As for (ii), it is sufficient to note that for any $F \in \mathcal{P}$, by letting $\mathcal{F} = \{F\}$ in (i), we have $\sup_{p\in\mathcal{V}} p\bar{F}_-(p) = \sup_{p\in\mathcal{V}} p\bar{F}(p)$. Since $p\bar{F}_-(p)$ is upper semi-continuous in p, we have $\sup_{p\in\mathcal{V}} p\bar{F}_-(p) = \max_{p\in\mathcal{V}} p\bar{F}_-(p)$. □

Lemma 2 shows the smoothing effect of the supremum operator that any discontinuity of F can be safely ignored in the pricing problem. Without the supremum operator, for a given p, $\inf_{F\in\mathcal{F}} p\bar{F}_-(p)$ may not be equal to $\inf_{F\in\mathcal{F}} p\bar{F}(p)$ (*e.g.*, when $\mathcal{F} = \{F\}$ and F is discontinuous at p). We make two further remarks. First, by definition of $\bar{\Lambda}^\star$, we can equivalently solve problem (3) as

$$\sup_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}_-(p) = \sup_{p\in\mathcal{V}} \inf_{F\in\mathcal{F}} p\bar{F}(p) = \sup_{p\in\mathcal{V}} p\bar{\Lambda}^\star(p) = \max_{p\in\mathcal{V}} p\bar{\Lambda}^\star_-(p),$$

where the last identity uses Lemma 2 (i) with $\mathcal{F}=\{\Lambda^\star\}$. Second, when $\mathcal{F}$ is compact, Weierstrass extreme value theorem and $p\bar{F}(p)$ being lower semi-continuous in F ensure that the solutions to the infimum problem in $\sup_{p\in\mathcal{V}}\inf_{F\in\mathcal{F}} p\bar{F}(p)$ and $\inf_{F\in\mathcal{F}}\sup_{p\in\mathcal{V}} p\bar{F}(p)$ (or problem (4)) exist.

LEMMA 3. [von Neumann-Fan minimax theorem; Borwein and Zhuang (1986)] *Let $g(x,y)$ be a real-valued function defined on $\mathcal{X}\times\mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are nonempty convex sets and $\mathcal{X}$ is compact. Suppose $g(\cdot,y)$ is concave and upper semi-continuous for each fixed $y\in\mathcal{Y}$ and $g(x,\cdot)$ is convex for each fixed $x\in\mathcal{X}$. Then*

$$\sup_{x\in\mathcal{X}}\inf_{y\in\mathcal{Y}} g(x,y) = \max_{x\in\mathcal{X}}\inf_{y\in\mathcal{Y}} g(x,y) = \inf_{y\in\mathcal{Y}}\max_{x\in\mathcal{X}} g(x,y) = \inf_{y\in\mathcal{Y}}\sup_{x\in\mathcal{X}} g(x,y).$$

Appendix B Proofs

Proof of Lemma 1. By (1), the transfer is uniquely defined by $t(v)=vx(v)-\int_0^v x(u)\mathrm{d}u = \int_0^v u\mathrm{d}x(u)\ \forall v\in\mathcal{V}$. Hence, we have $\sup_{(x,t)\in\mathcal{M}}\inf_{F\in\mathcal{F}}\mathbb{E}_F[t(\tilde{v})]=\sup_{x\in\mathcal{X}}\inf_{F\in\mathcal{F}}\mathbb{E}_F[\int_0^{\tilde{v}} u\mathrm{d}x(u)]$, where $\mathcal{X}=\{x\mid(x,p)\in\mathcal{M}\}$ is known to be convex and compact (see Börgers 2015, lemma 2.6). Note that the objective functional $\mathbb{E}_F[\int_0^{\tilde{v}} u\mathrm{d}x(u)]=\int_{\mathcal{V}}\int_0^v u\mathrm{d}x(u)\mathrm{d}F(v)$ is bilinear in F and x, and hence is convex in F and concave in x. In addition, it is upper semi-continuous in x for each fixed $F\in\mathcal{F}$ (see Carrasco et al. 2018, lemma 8). Using von Neumann-Fan minimax theorem, we now obtain

$$\sup_{(x,t)\in\mathcal{M}}\inf_{F\in\mathcal{F}}\mathbb{E}_F[t(\tilde{v})] = \max_{x\in\mathcal{X}}\inf_{F\in\mathcal{F}}\mathbb{E}_F\left[\int_0^{\tilde{v}} u\mathrm{d}x(u)\right] = \inf_{F\in\mathcal{F}}\max_{x\in\mathcal{X}}\mathbb{E}_F\left[\int_0^{\tilde{v}} u\mathrm{d}x(u)\right].$$

Given $F\in\mathcal{F}$, Manelli and Vincent (2007) observe that the problem $\max_{x\in\mathcal{X}}\mathbb{E}_F[\int_0^{\tilde{v}} u\mathrm{d}x(u)]$ is maximizing a linear objective functional over the convex set $\mathcal{X}$. Hence, the optimum is attained at one of the extreme points of $\mathcal{X}$. Manelli and Vincent (2007) further establish that the set of extreme points of $\mathcal{X}$ is exactly the set of posted price mechanisms. Therefore, $\inf_{F\in\mathcal{F}}\max_{x\in\mathcal{X}}\mathbb{E}_F[\int_0^{\tilde{v}} u\mathrm{d}x(u)]=\inf_{F\in\mathcal{F}}\max_{p\in\mathcal{V}} p\bar{F}_-(p)$. □

Proof of Proposition 1. We first prove condition (i). For any $F\in\mathcal{F}$ and $p\in\mathcal{V}$, observe that the definition of $\Lambda^\star$ gives $p\bar{\Lambda}^\star(p)\le p\bar{F}(p)$, implying $\sup_{p\in\mathcal{V}} p\bar{\Lambda}^\star(p)\le\sup_{p\in\mathcal{V}} p\bar{F}(p)\ \forall F\in\mathcal{F}$, or equivalently, $\sup_{p\in\mathcal{V}} p\bar{\Lambda}^\star(p)\le\inf_{F\in\mathcal{F}}\sup_{p\in\mathcal{V}} p\bar{F}(p)$. Since $\Lambda^\star\in\mathcal{F}$, we arrive at $\inf_{F\in\mathcal{F}}\sup_{p\in\mathcal{V}} p\bar{F}(p)\le\sup_{p\in\mathcal{V}} p\bar{\Lambda}^\star(p)$. Hence, $\inf_{F\in\mathcal{F}}\sup_{p\in\mathcal{V}} p\bar{F}(p)=\sup_{p\in\mathcal{V}} p\bar{\Lambda}^\star(p)=\sup_{p\in\mathcal{V}}\inf_{F\in\mathcal{F}} p\bar{F}(p)$, *i.e.*, $\eta=1$.

We next prove condition (ii). Note that by Lemma 2, the supremum over $p\in\mathcal{V}$ can always be attained and we can use maximum instead. Let $\pi^\star:=\inf_{F\in\mathcal{F}}\max_{p\in\mathcal{V}} p\bar{F}_-(p)=p^\star\bar{F}_-^\star(p^\star)$. Observe that $p^\star\neq 0$ because otherwise $\pi^\star=0$ would yield a contradiction that all $p\in\mathcal{V}$ are optimal. Hence, we focus on $p^\star>0$. By uniqueness of $p^\star$ under $F^\star$, we have $\bar{F}_-^\star(p^\star)=\pi^\star/p^\star$ and $\bar{F}_-^\star(v)<\pi^\star/v$ for $v\neq p^\star$. Suppose $\eta<1$. Then $\max_{p\in\mathcal{V}}\inf_{F\in\mathcal{F}} p\bar{F}_-(p)<\pi^\star$, implying $\inf_{F\in\mathcal{F}} p^\star\bar{F}_-(p^\star)<\pi^\star$. As a result, there must exist $F\in\mathcal{F}$ that satisfies $p^\star\bar{F}_-(p^\star)<\pi^\star$, *i.e.*, $\bar{F}_-(p^\star)<\pi^\star/p^\star$. We will next show that one can then find a $\lambda^\diamond \in (0,1]$ such that for $F^\diamond = \lambda^\diamond F + (1-\lambda^\diamond)F^\star \in \mathcal{F}$, we have $\max_{p\in\mathcal{V}} p\bar{F}^\diamond_-(p) < \pi^\star$—a contradiction to the definition of $\pi^\star$.

Observe first that with $\bar{F}_-(p^\star) < \pi^\star/p^\star$ and the upper semi-continuity of $\bar{F}_-(p)$ in q, there must exist $\delta_1 > 0$ such that $\bar{F}_-(v) < \pi^\star/p^\star < \pi^\star/v$ for all $v \in (p^\star - \delta_1, p^\star)$. By the continuity of $\pi^\star/v$, there further exists $\delta_2 > 0$ such that $\bar{F}_-(v) \le \bar{F}_-(p^\star) < \pi^\star/v$ for all $v \in (p^\star, p^\star + \delta_2)$. Let $\mathcal{V}_0 := \mathcal{V} \cap (p^\star - \delta_1, p^\star + \delta_2)$. It follows that $\lambda\bar{F}_-(v) + (1-\lambda)\bar{F}^\star_-(v) < \pi^\star/v$ for any $\lambda \in (0,1]$ and $v \in \mathcal{V}_0$. Consider next the compact set $\mathcal{V}\setminus\mathcal{V}_0$. Since $\pi^\star/v - \bar{F}^\star_-(v)$ is lower semi-continuous in v, we can find $0 < \varepsilon < \min_{v\in\mathcal{V}\setminus\mathcal{V}_0}(\pi^\star/v - \bar{F}^\star_-(v))$. This yields $\bar{F}^\star_-(v) < \pi^\star/v - \varepsilon$ for all $v \in \mathcal{V}\setminus\mathcal{V}_0$. Let $\lambda^\diamond = \varepsilon/(1+\varepsilon-\pi^\star) \in (0,1]$ and consider $F^\diamond = \lambda^\diamond F + (1-\lambda^\diamond)F^\star \in \mathcal{F}$. For $v \in \mathcal{V}\setminus\mathcal{V}_0$, $F^\diamond_-(v) = \lambda^\diamond \bar{F}_-(v) + (1-\lambda^\diamond)\bar{F}^\star_-(v) < \lambda^\diamond + (1-\lambda^\diamond)(\pi^\star/v - \varepsilon) = \lambda^\diamond(1-(\pi^\star/v-\varepsilon)) + (\pi^\star/v - \varepsilon) \le \pi^\star/v$. We now obtain $vF^\diamond_-(v) < \pi^\star$ for all $v \in \mathcal{V}$ and $\max_{p\in\mathcal{V}} p\bar{F}^\diamond_-(p) < \pi^\star$, concluding the proof. □

Proof of Theorem 1. First, the uniqueness of $\pi^\star$ follows from the monotonicity of the function $z - z\ln(z)$ as $(z - z\ln(z))' = -\ln(z) > 0$ when $z \in (0,1)$. Next, since $\int_{\mathcal{V}} \bar{F}^\star(v)\mathrm{d}v = \pi^\star + \int_{\pi^\star}^1 (\pi^\star/v)\mathrm{d}v = \pi^\star + \pi^\star\ln(1/\pi^\star) = \mu$, we have $F^\star \in \mathcal{F}_{\mathrm{M}}$. In addition, it is clear that $\sup_{p\in\mathcal{V}} p\bar{F}^\star(p) = \max_{p\in\mathcal{V}} p\bar{F}^\star(p) = \pi^\star$. We next show $F^\star$ is the unique optimal solution by contradiction. Suppose there exists $F \in \mathcal{F}_{\mathrm{M}}$ such that $\sup_{p\in\mathcal{V}} p\bar{F}(p) \le \max_{p\in\mathcal{V}} p\bar{F}^\star(p) = \pi^\star$. Then for $v \in [0,\pi^\star)$, $\bar{F}(v) \le 1 = \bar{F}^\star(v)$ and for $v \in [\pi^\star, 1)$, $v\bar{F}(v) \le \sup_{p\in\mathcal{V}} p\bar{F}(p) \le \pi^\star = v\bar{F}^\star(v)$, *i.e.*, $\bar{F}(v) \le \bar{F}^\star(v)$. Since $\bar{F}, \bar{F}^\star \in F_{\mathrm{M}}$, we have $\int_{\mathcal{V}}(\bar{F}^\star(v) - \bar{F}(v))\mathrm{d}v = 0$. By right continuity of $\bar{F}^\star(v) - \bar{F}(v)$ and the fact that $\bar{F}^\star(v) - \bar{F}(v) \ge 0$ for any $v \in \mathcal{V}$, we must have $F(v) = F^\star(v)$ for any $v \in \mathcal{V}$. □

Proof of Proposition 2. We first show that the constructed $(x^\star, t^\star, \alpha^\star, \beta^\star)$ is feasible to problem (6). Since $v\alpha^\star + \beta^\star = \alpha^\star(v - \pi^\star) \le \alpha^\star(v-\pi^\star)^+ = t^\star(v)$ for all $v \in \mathcal{V}$, the first constraint in the left reformulation holds. From $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v} = v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v} = \alpha^\star$ for $v \in [\pi^\star, 1]$ and $t^\star(v) = x^\star(v) = 0$ for $v \in [0, \pi^\star)$, it follows that $t^\star(v) = vx^\star(v) - \int_0^v x^\star(u)\mathrm{d}u$. With $x^\star$ being increasing, it is now clear that $(x^\star, p^\star) \in \mathcal{M}$. Finally, the feasible solution $(x^\star, t^\star, \alpha^\star, \beta^\star)$ has an objective value $\mu\alpha^\star + \beta^\star = (\mu - \pi^\star)\alpha^\star = \pi^\star$, attaining the optimal revenue of problem (4). This certifies that $(x^\star, t^\star)$ is indeed optimal. □

Proof of Proposition 3. For any $p \in (0,\mu)$, note that $\int_{\mathcal{V}} \bar{F}^\star_p(v)\mathrm{d}v = \mu$, hence $F^\star_p \in \mathcal{F}_{\mathrm{M}}$. Suppose there exists $F \in \mathcal{F}_{\mathrm{M}}$ such that $\bar{F}(p) \le \bar{F}^\star_p(p)$. We have $\bar{F}(v) \le 1 = \bar{F}^\star_p(v)$ for $v \in [0,p)$ and $\bar{F}(v) \le \bar{F}(p) \le \bar{F}^\star_p(p) = \bar{F}^\star_p(v)$ for $v \in [p,1)$. That is, $\bar{F}(v) \le \bar{F}^\star_p(v)$ for any $v \in [0,1]$. On the other hand, since $\bar{F}, \bar{F}^\star_p \in F_{\mathrm{M}}$, we have $\int_{\mathcal{V}}(\bar{F}^\star_p(v) - \bar{F}(v))\mathrm{d}v = 0$. By right continuity, we must have $F(v) = F^\star_p(v)$ for any $v \in \mathcal{V}$. This shows that $\bar{F}^\star_p$ must be the unique worst-case distribution. For $p \in [\mu, 1]$, the Dirac distribution δ_μ must be a worst-case distribution since the revenue is 0 under δ_μ.

Now $\bar{F}^\star_p$ yields that $\bar{\Lambda}^\star(p) = (\mu-p)^+/(1-p)$ $\forall p \in [0,1]$. To solve $\sup_{p\in\mathcal{V}} p\bar{\Lambda}^\star(p)$, we only need to focus on $p \in [0,\mu]$. Since $p\bar{\Lambda}^\star(p) = p(\mu-p)/(1-p)$ is increasing on $[0, 1-\sqrt{1-\mu})$ while decreasing on $(1-\sqrt{1-\mu}, \mu]$, the optimal posted price is $p^\star = 1 - \sqrt{1-\mu}$ with an objective value $(1-\sqrt{1-\mu})^2$.

Then we look at the performance ratio η. With Theorem 1, $\eta = (1-\sqrt{1-\mu})^2/\pi^\star$. Let $y = 1/\pi^\star$, we have $\eta = (\sqrt{y}-\sqrt{y-1-\ln(y)})^2$. To prove η is increasing in μ on $(0,1]$, it is sufficient to show that $\sqrt{y}-\sqrt{y-\ln(y)-1}$ is decreasing on $y\geq 1$. Indeed, the monotonicity follows from taking the derivative: $(\sqrt{y}-\sqrt{y-\ln(y)-1})' = (1-\ln(y))(4y\sqrt{y-\ln(y)-1})\leq 0$. When $\mu\to 0$, we have $\pi^\star\to 0$ (*i.e.*, $y\to\infty$). Note that $\sqrt{y}-\sqrt{y-1-\ln(y)}\geq 0$ for any $y\geq 1$. To show η approaches 0, it remains to show $\sqrt{y}-\sqrt{y-1-\ln(y)}\to 0$ as $y\to\infty$. Indeed, we have $\sqrt{y}-\sqrt{y-1-\ln(y)}\leq (1+\ln(y))(2\sqrt{y-1-\ln(y)})$. Applying L'Hôpital's rule, we have $\lim_{y\to\infty}\frac{1+\ln(y)}{\sqrt{y}} = \lim_{y\to\infty}\frac{1/y}{1/\sqrt{y}} = 0$, *i.e.*, $1+\ln(y) = o(\sqrt{y})$. This leads to $\sqrt{y-1-\ln(y)} = \sqrt{y-o(\sqrt{y})} = O(\sqrt{y})$, concluding $\lim_{y\to\infty}\frac{(\ln(y)+1)}{2\sqrt{y-\ln(y)-1}}\to 0$. When $\mu=1$, we have $\pi^\star = 1$, implying $\eta = 1$. □

Proof of Theorem 2. Suppose, on the contrary, there exists $F\in\mathcal{F}_{\mathrm{MPC}}$ such that $\pi := \sup_{q\in\mathcal{V}} p\bar{F}(p) < \max_{p\in\mathcal{V}} p\bar{Q}_{k^\star}(p) = \pi^\star$. Since $k^\star = \pi^\star\exp((\mu-\pi^\star)/\pi^\star)$ is strictly decreasing in $\pi^\star$, it follows that $k^\star < k := \pi\exp((\mu-\pi)/\pi)$ and $\pi_k = \pi$. We next show that $Q_k\in\mathcal{F}_{\mathrm{MPC}}$—a contradiction to the definition of $k^\star$. To this end, we first observe that for $v\in[\pi,k)$, $v\bar{F}(v)\leq \sup_{p\in\mathcal{V}} p\bar{F}(p) = \pi = \max_{p\in\mathcal{V}} p\bar{Q}_k(p) = v\bar{Q}_k(v)$, *i.e.*, $\bar{F}(v)\leq\bar{Q}_k(v)$. Besides, we have $\bar{F}(v)\leq 1 = \bar{Q}_k(v)$ for $v\in[0,\pi)$ and $\bar{F}(v)\geq 0 = \bar{Q}_k(v)$ for $v\in[k,1]$. These relations imply $\int_0^u \bar{Q}_k(v)\mathrm{d}v \geq \int_0^u \bar{F}(v)\mathrm{d}v$ for $u\in[0,k)$ and $\int_0^u \bar{Q}_k(v)\mathrm{d}v = \mu - \int_u^1 \bar{Q}_k(v)\mathrm{d}v \geq \mu - \int_u^1\bar{F}(v)\mathrm{d}v = \int_0^u\bar{F}(v)\mathrm{d}v$ for $u\in[k,1]$. That is, Q_k is a mean-preserving contraction of F, and hence a mean-preserving contraction of G as well, concluding $Q_k\in\mathcal{F}_{\mathrm{MPC}}$. It now follows that the optimal revenue is $\pi^\star = \max_{p\in\mathcal{V}} p\bar{Q}_{k^\star}(p)$. □

Proof of Proposition 4. We first show that there exists $u^\star\in[\pi^\star,k^\star]$ such that $\int_0^{u^\star}(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v = 0$. If $k^\star = 1$ then we can put $u^\star = 1$. If $k^\star < 1$, suppose for all $u\in[\pi^\star,k^\star]$ we have $\int_0^u(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v > 0$. One can then define $\varepsilon = \min_{u\in[\pi^\star,k^\star]}\int_0^u(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v > 0$ and choose a small positive $\varepsilon'\in(0,\varepsilon/3)$ such that for $k = k^\star+\varepsilon'$, it holds that $k^\star < k < 1$, $0 < \pi^\star - \pi_k < \varepsilon/3$, and $0 < \int_0^u(\bar{F}^\star(v)-\bar{Q}_k(v))\mathrm{d}v = \int_0^u(\bar{Q}_{k^\star}(v)-\bar{Q}_k(v))\mathrm{d}v < \varepsilon/3$ for all $u\in\mathcal{V}$. With some tedious computations, one can then show that $\int_0^u(\bar{Q}_k(v)-\bar{G}(v))\mathrm{d}v\geq 0$ for all $u\in\mathcal{V}$, *i.e.* $Q_k\in\mathcal{F}_{\mathrm{MPC}}$—a contradiction to the definition of $k^\star$.

We now derive the optimal mechanism. Given $\mathcal{F}_{\mathrm{MPC}}$, the inner minimization of problem (2) is

$$
\begin{aligned}
\inf_{F\in\mathcal{P}_+}\quad & \int_{\mathcal{V}} t(v)\,\mathrm{d}F(v) \\
\text{s.t.}\quad & \int_{\mathcal{V}} v\,\mathrm{d}F(v) = \mu \\
& \int_{\mathcal{V}} (u-v)^+\,\mathrm{d}F(v) \leq \int_{\mathcal{V}} (u-v)^+\,\mathrm{d}G(v) \quad \forall u\in\mathcal{V} \\
& \int_{\mathcal{V}} \mathrm{d}F(v) = 1.
\end{aligned}
\tag{8}
$$

Combining the dual of problem (8) with the outer maximization, problem (2) becomes

$$\begin{aligned} \sup_{x,t,\alpha,\beta,\Gamma} \quad & \mu\alpha+\beta-\int_{\mathcal{V}}\phi(u)\,\mathrm{d}\Gamma(u) \\ \text{s.t.} \quad & p(v)\geq v\alpha+\beta-\int_{\mathcal{V}}(u-v)^{+}\,\mathrm{d}\Gamma(u) \quad \forall v\in\mathcal{V} \\ & (x,t)\in\mathcal{M},\ \alpha\in\mathbb{R},\ \beta\in\mathbb{R},\ \Gamma\in\mathcal{P}_{+}, \end{aligned} \tag{9}$$

where for any $u\in\mathcal{V}$ we put $\phi(u)=\int_{\mathcal{V}}(u-v)^{+}\mathrm{d}G(v)$.

Let $\alpha^\star=0$, $\beta^\star=\gamma^\star(u^\star-\pi^\star)$ and $\Gamma^\star$ be a step function such that $\Gamma^\star(v)=0$ for $v\in[0,u^\star)$; $\Gamma^\star(v)=\gamma^\star$ for $[u^\star,1]$. We first verify the feasiblity of $(x^\star,t^\star,\alpha^\star,\beta^\star,\Gamma^\star)$ to problem (9). For the semi-infinite constraint, it holds that $v\alpha^\star+\beta^\star-\int_{\mathcal{V}}(u-v)^{+}\,\mathrm{d}\Gamma^\star(u)=\beta^\star-\gamma^\star(u^\star-v)=\gamma^\star(v-\pi^\star)\leq t^\star(v)$ for $v\in[0,u^\star)$ and $v\alpha^\star+\beta^\star-\int_{\mathcal{V}}(u-v)^{+}\mathrm{d}\Gamma^\star(u)=\beta^\star=t^\star(v)$ for $v\in[u^\star,1]$. Since $t^\star(v)=x^\star(v)=0$ for $v\in[0,\pi^\star)$, $\frac{\mathrm{d}p^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=\gamma^\star$ for $v\in[\pi^\star,u^\star)$, and $\frac{\mathrm{d}p^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=0$ for $v\in[u^\star,1]$, we have $t^\star(v)=vx^\star(v)-\int_0^v x^\star(u)\mathrm{d}u$. Together with $x^\star$ being increasing, it follows that $(x^\star,t^\star)\in\mathcal{M}$. Finally, the objective value of $(x^\star,t^\star,\alpha^\star,\beta^\star,\Gamma^\star)$ is $\mu\alpha^\star+\beta^\star-\int_{\mathcal{V}}\phi(u)\mathrm{d}\Gamma^\star(u)=\gamma^\star(u^\star-\pi^\star)-\gamma^\star(u^\star-\int_0^{u^\star}\bar{G}(u)\mathrm{d}u)=\gamma^\star(u^\star-\pi^\star)-\gamma^\star\big(u^\star-\int_0^{u^\star}\bar{F}^\star(u)\mathrm{d}u\big)=\pi^\star$, where in the second equality we explore the fact that $u^\star$ satisfies $\int_0^{u^\star}(\bar{F}(v)-\bar{G}(v))\mathrm{d}v=0$. This concludes $(x^\star,t^\star)$ is optimal. □

Proof of Proposition 5. For any $p\geq\mu$, clearly the Dirac measure δ_μ is the worst. For $p<\mu$, suppose there exists $F\in\mathcal{F}_{\mathrm{MPC}}$ such that $\bar{F}(p)<\bar{F}_p^\star(p)$. Then we have $k:=p+(\mu-p)/\bar{F}(p)>p+(\mu-p)/\bar{F}_p^\star(p)=k_p^\star$. Consider the distribution $R_{p,k}$. For $v\in[p,k)$, we have $\bar{F}(v)\leq\bar{F}(p)=\bar{R}_{p,k}(p)=\bar{R}_{p,k}(v)$. Besides, $\bar{F}(v)\leq 1=\bar{R}_{p,k}(v)$ for $v\in[0,p)$ and $\bar{F}(v)\geq 0=\bar{R}_{p,k}(v)$ for $v\in[k,1]$. These relations imply $\int_0^u\bar{R}_{p,k}(v)\mathrm{d}v\geq\int_0^u\bar{F}(v)\mathrm{d}v$ for $u\in[0,k)$ and $\int_0^u\bar{R}_{p,k}(v)\mathrm{d}v=\mu-\int_u^1\bar{R}_{p,k}(v)\mathrm{d}v\geq\mu-\int_u^1\bar{F}(v)\mathrm{d}v=\int_0^u\bar{F}(v)\mathrm{d}v$ for $u\in[k,1]$. That is, $R_{p,k}$ is a mean-preserving contraction of F, and hence a mean-preserving contraction of G. This shows $R_{p,k}\in\mathcal{F}_{\mathrm{MPC}}$ and contradicts to the definition of $k_p^\star$. □

Proof of Theorem 3. Although the seller's problem in (5) may admit multiple optimal solutions and the choice of different prices may result in different payoffs for the buyer, Roesler and Szentes (2017) show that in equilibrium one can focus on $p_0\in\arg\max_{p\in\mathcal{V}}p\bar{F}_-(p)$ that also maximizes $\mathbb{E}_F[(\tilde{v}-p_0)^+]$.

By efficiency of $F^\star$, there exists $p^\star\in\arg\max_{p\in\mathcal{V}}p\bar{F}_-^\star(p)$ such that $\mathbb{E}_{F^\star}[(\tilde{v}-p^\star)^+]=\mu-\pi^\star$. Since $\mu-\pi^\star$ is the maximum payoff the buyer can obtain with the distribution $F^\star$, the seller's optimal price in problem (5) under $F^\star$ is $p^\star$ and the buyer's payoff under $\bar{F}^\star$ is $\mu-\pi^\star$.

Now consider any $F\in\mathcal{F}$ and $p_0\in\arg\max_{p\in\mathcal{V}}p\bar{F}_-(p)$ that also maximizes $\mathbb{E}_F[(\tilde{v}-p_0)^+]$. As $F^\star$ is an optimal solution to problem (4), we must have $p_0\bar{F}_-(p_0)\geq\sup_{p\in\mathcal{V}}p\bar{F}_-^\star(p)=\pi^\star$. It follows that $\mathbb{E}_F[(\tilde{v}-p_0)^+]=\int_{p_0}^1(v-p_0)\mathrm{d}F(v)\leq\mu-p_0\bar{F}_-(p_0)\leq\mu-\pi^\star=\mathbb{E}_{F^\star}[(\tilde{v}-p^\star)^+]$. Hence, $F^\star$ is also optimal to problem (5). □

Proof of Theorem 4. Note that $\bar{F}^{\star-1}(q) = \inf\{v \in \mathcal{V} \mid \min\{\bar{G}(v), \pi^\star/v\} \le q\} = \min\{\inf\{v \in \mathcal{V} \mid \bar{G}(v) \le q\}, \inf\{v \in \mathcal{V} \mid \pi^\star/v \le q\}\} = \min\{\bar{G}^{-1}(q), \pi^\star/q\}$, which satisfies $\bar{F}^\star(\bar{F}^{\star-1}(q)) \le q$ for $q \in [0,1]$ and $\bar{F}^{\star-1}(\bar{F}^\star(v)) \le v$ for $v \in \mathcal{V}$. Then, $|\bar{F}^{\star-1}(q) - \bar{G}^{-1}(q)| = (\bar{G}^{-1}(q) - \pi^\star/q)^+$ for $q \in [0,1]$, implying

$$d_\rho(F^\star, G) = \left(\int_0^1 |\bar{F}^{\star-1}(q) - \bar{G}^{-1}(q)|^\rho \mathrm{d}q\right)^{1/\rho} = \left(\int_0^1 ((\bar{G}^{-1}(q) - \pi^\star/q)^+)^\rho \mathrm{d}q\right)^{1/\rho} = \theta.$$

Hence, $F^\star \in \mathcal{F}(\theta, \rho)$ and it remains to show its optimality. Before proceeding, we define $\mathcal{Q}_+ := \{q \in [0,1] \mid \bar{G}^{-1}(q) > \pi^\star/q\}$ on which $\bar{F}^{\star-1}(q) = \pi^\star/q$ and $\mathcal{V}_+ := \{v \in \mathcal{V} \mid \bar{G}(v) > \pi^\star/v\}$ on which $\bar{F}^\star(v) = \pi^\star/v$. By definition of $\pi^\star$, $\mathcal{V}_+ \neq \emptyset$.

For $\rho \in [1, \infty)$, suppose there exists $F \in \mathcal{F}_{(\theta,\rho)}$ such that $\sup_{p \in \mathcal{V}} p\bar{F}(p) \le \pi^\star$. We first prove $\bar{F}^{-1}(q) \le \bar{F}^{\star-1}(q) < \bar{G}^{-1}(q)$ for $q \in \mathcal{Q}_+$. For any $v \in \mathcal{V}_+$, it holds that $v\bar{F}(v) \le \pi^\star = v\bar{F}^\star(v)$, implying $\bar{F}(v) \le \bar{F}^\star(v)$. For any $q \in \mathcal{Q}_+$, we must have $\bar{G}(\pi^\star/q) > q$ because otherwise, the definition of $\bar{G}^{-1}$ would yield $\bar{G}^{-1}(q) \le \pi^\star/q$, a contradiction to $q \in \mathcal{Q}_+$. This implies $\pi^\star/q \in \mathcal{V}_+$, which leads to $\bar{F}^\star(\pi^\star/q) = \pi^\star/(\pi^\star/q) = q$. Then $q = \bar{F}^\star(\pi^\star/q) \ge \bar{F}(\pi^\star/q)$. Since $\bar{F}^{-1}$ is decreasing, we have $\bar{F}^{-1}(q) \le \bar{F}^{-1}(\bar{F}(\pi^\star/q)) \le \pi^\star/q = \bar{F}^{\star-1}(q) < \bar{G}^{-1}(q)$ for $q \in \mathcal{Q}_+$. It now follows that

$$\theta^\rho \ge \int_{\mathcal{Q}_+} |\bar{G}^{-1}(q) - \bar{F}^{-1}(q)|^\rho \mathrm{d}q = \int_{\mathcal{Q}_+} (\bar{G}^{-1}(q) - \bar{F}^{-1}(q))^\rho \mathrm{d}q \ge \int_{\mathcal{Q}_+} (\bar{G}^{-1}(q) - \bar{F}^{\star-1}(q))^\rho \mathrm{d}q,$$

where the last inequality is due to $\bar{F}^{-1}(q) \le \bar{F}^{\star-1}(q) < \bar{G}^{-1}(q)$ for $q \in \mathcal{Q}_+$. Besides, we have

$$\int_{\mathcal{Q}_+} (\bar{G}^{-1}(q) - \bar{F}^{\star-1}(q))^\rho \mathrm{d}q = \int_{\mathcal{Q}_+} (\bar{G}^{-1}(q) - \pi^\star/q)^\rho \mathrm{d}q = \int_0^1 |(\bar{G}^{-1}(q) - \pi^\star/q)^+|^\rho \mathrm{d}q = \theta^\rho.$$

Now we obtain $\int_{\mathcal{Q}_+} (\bar{G}^{-1}(q) - \bar{F}^{-1}(q))^\rho \, \mathrm{d}q = \int_{\mathcal{Q}_+} (\bar{G}^{-1}(q) - \bar{F}^{\star-1}(q))^\rho \, \mathrm{d}q = \theta^\rho$. By right continuity and the fact that $(\bar{G}^{-1}(q) - \bar{F}^{-1}(q))^\rho \ge (\bar{G}^{-1}(q) - \bar{F}^{\star-1}(q))^\rho$, for all $q \in \mathcal{Q}_+$ we must have $(\bar{G}^{-1}(q) - \bar{F}^{-1}(q))^\rho - (\bar{G}^{-1}(q) - \bar{F}^{\star-1}(q))^\rho = 0$, which implies $\bar{F}^{-1}(q) = \bar{F}^{\star-1}(q)$. In addition, since $\theta^\rho \ge \int_0^1 |\bar{G}^{-1}(q) - \bar{F}^{-1}(q)|^\rho \, \mathrm{d}q \ge \int_{\mathcal{Q}_+} |\bar{G}^{-1}(q) - \bar{F}^{-1}(q)|^\rho \, \mathrm{d}q = \theta^\rho$, we have $\int_{[0,1] \setminus \mathcal{Q}_+} |\bar{G}^{-1}(q) - \bar{F}^{-1}(q)|^\rho \mathrm{d}q = 0$. By right continuity and the fact that $|\bar{G}^{-1}(q) - \bar{F}^{-1}(q)|^\rho \ge 0$, we have $\bar{F}^{-1}(q) = \bar{G}^{-1}(q) = \bar{F}^{\star-1}(q)$ for $q \notin \mathcal{Q}_+$. Now it is clear that $F = F^\star$. Hence, $F^\star$ is the unique worst-case distribution.

For $\rho = \infty$, suppose there exists $F \in \mathcal{F}_{(\theta,\rho)}$ such that $\sup_{p \in \mathcal{V}} p\bar{F}(p) < \pi^\star$. By moving $\bar{G}$ to the left with a distance θ, we can define a new distribution $G^\diamond(v) := \bar{G}(v + \theta)$ $\forall v \in \mathcal{V}$ and its inverse is: $\bar{G}^{\diamond-1}(q) = (\bar{G}^{-1}(q) - \theta)^+$. We first show that $\bar{F}(v) \ge \bar{G}^\diamond(v)$ for all $v \in \mathcal{V}$. Suppose, on the contrary, there exists $v' \in \mathcal{V}$ such that $\bar{F}(v') < \bar{G}^\diamond(v') \le \bar{G}^\diamond(0) = \bar{G}(\theta)$. Then we must have $\bar{G}^{\diamond-1}(\bar{F}(v')) > v'$ because otherwise if $\bar{G}^{\diamond-1}(\bar{F}(v')) \le v'$, then $\bar{F}(v') \ge \bar{G}(\bar{G}^{\diamond-1}(\bar{F}(v'))) \ge \bar{G}^\diamond(v')$, which contradicts with our hypothesis. Since $\bar{F}(v') < \bar{G}(\theta)$, we have $\theta = \bar{G}^{-1}(\bar{F}(v')) - \bar{G}^{\diamond-1}(\bar{F}(v')) < \bar{G}^{-1}(\bar{F}(v')) - v' \le \bar{G}^{-1}(\bar{F}(v')) - \bar{F}^{-1}(\bar{F}(v'))$, contradicting to $F \in \mathcal{F}_{(\theta,\infty)}$. Hence, we must have $\bar{F}(v) \ge \bar{G}^\diamond(v)$ for all $v \in \mathcal{V}$. Next, by definition of $\pi^\star$, we have $\sup_{q \in (0,1]} (\bar{G}^{-1}(q) - \pi^\star/q)^+ = \theta$. Thus, for any $\varepsilon \in (0, \pi^\star)$, there exists $q_0 \in (0,1]$ such that $\bar{G}^{-1}(q_0) - \pi^\star/q_0 > \theta - \varepsilon/2$. Define $v_0 := \pi^\star/q_0 - \varepsilon/2 > \varepsilon/2$. Then we have $v_0 < (\bar{G}^{-1}(q_0)-\theta)^+ = \bar{G}^{\diamond -1}(q_0) = \inf\{v \in \mathcal{V} \mid \bar{G}^{\diamond}(v) \le q_0\}$. Thus, for any $v \in [0, v_0)$, we have $\bar{G}^{\diamond}(v) > q_0$. It follows that $\sup_{p\in\mathcal{V}} p\bar{G}^{\diamond}(p) \ge (v_0-\varepsilon/2)\bar{G}^{\diamond}(v_0-\varepsilon/2) > (v_0-\varepsilon/2)q_0 = \pi^\star - \varepsilon q_0 \ge \pi^\star - \varepsilon$. Since $\bar{F}(v) \ge \bar{G}^{\diamond}(v)$ for $v \in \mathcal{V}$, we have $\sup_{p\in\mathcal{V}} p\bar{F}(p) \ge \sup_{p\in\mathcal{V}} p\bar{G}^{\diamond}(p) \ge \pi^\star - \varepsilon$ for any $\varepsilon \in (0,\pi^\star)$, contradicting to $\sup_{p\in\mathcal{V}} p\bar{F}(p) < \pi^\star$. Hence, $F^\star$ is a worst-case distribution. Our proof here also establishes that $G^{\diamond}$ is a worst-case distribution.

Finally, regardless of the value of ρ, it is clear that $v\bar{F}^\star(v) \le v\cdot\pi^\star/v = \pi^\star$ for any $v \in \mathcal{V}$, and the optimal revenue under $\bar{F}^\star$ is $\pi^\star$ since $\mathcal{V}_+$ is nonempty. $\square$

Proof of Proposition 6. Recall from Appendix C.5, we have $d_1(F,G) = \inf_{Q\in\mathcal{Q}(F,G,\mathcal{V})} \mathbb{E}_Q[\|\tilde{v}-\tilde{u}\|_1]$, where $\tilde{v} \sim F$, $\tilde{u} \sim G$, and $\mathcal{Q}(F,G,\mathcal{V})$ is the set of all joint distributions supported on $\mathcal{V}\times\mathcal{V}$ with marginals F and G. By the law of total probability, we can express $Q \in \mathcal{Q}(F,G,\mathcal{V})$ of $(\tilde{v},\tilde{u})$ as the marginal distribution $G(u)$ and the conditional distribution $F_u(v)$, that is, for any $(v_0,u_0) \in \mathcal{V}\times\mathcal{V}$,

$$Q(v_0,u_0) = \mathbb{P}(\tilde{v} \le v_0, \tilde{u} \le u_0) = \int_0^{u_0} \mathbb{P}(\tilde{v} \le v_0 \mid \tilde{u} = u)\, \mathrm{d}G(u) = \int_0^{u_0} F_u(v_0)\, \mathrm{d}G(u).$$

As a result, we have

$$\mathbb{E}_Q[t(\tilde{v})] = \int_{u\in\mathcal{V}} \left(\int_{v\in\mathcal{V}} p(v) \mathrm{d}F_u(v) \right) \mathrm{d}G(u) \quad \text{and} \quad \mathbb{E}_Q[\|\tilde{v}-\tilde{u}\|_1] = \int_{u\in\mathcal{V}} \left(\int_{v\in\mathcal{V}} |v-u|\, \mathrm{d}F_u(v) \right) \mathrm{d}G(u).$$

These interpretations allow us to write the inner minimization of problem (2) as follows:

$$\begin{aligned} \inf_{F_u\in\mathcal{P}_+, u\in\mathcal{V}} \quad & \int_{u\in\mathcal{V}} \left(\int_{v\in\mathcal{V}} t(v)\, \mathrm{d}F_u(v) \right) \mathrm{d}G(u) \\ \text{s.t.} \quad & \int_{u\in\mathcal{V}} \left(\int_{v\in\mathcal{V}} |v-u|\, \mathrm{d}F_u(v) \right) \mathrm{d}G(u) \le \theta \\ & \int_{\mathcal{V}} \mathrm{d}F_u(v) = 1 \qquad\qquad \forall u \in \mathcal{V}. \end{aligned} \tag{10}$$

Combining the dual of problem (10) with the outer maximization, problem (2) now becomes

$$\begin{aligned} \sup_{x,t,\alpha,\beta} \quad & \int_{u\in\mathcal{V}} \beta(u)\, \mathrm{d}G(u) - \theta\alpha \\ \text{s.t.} \quad & t(v) \ge \beta(u) - |v-u|\alpha \qquad \forall v,u \in \mathcal{V} \\ & (x,t) \in \mathcal{M},\ \alpha \in \mathbb{R}_+,\ \beta(u) \in \mathbb{R} \quad \forall u \in \mathcal{V}. \end{aligned} \tag{11}$$

One can check that $p^\star$ is continuous and $\frac{\mathrm{d}p^\star(v)}{\mathrm{d}v} \le \alpha^\star$ whenever differentiable. Thus, with $\beta^\star = t^\star$ we have $t^\star(v) - \beta^\star(u) = t^\star(v) - t^\star(u) \ge -\alpha^\star|u-v|$, that is, the first constraint in problem (11) holds. Besides, we have $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v} = v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}$ almost everywhere and $t^\star(0) = x^\star(0) = 0$. It follows that $t^\star(v) = vx^\star(v) - \int_0^v x^\star(u)\mathrm{d}u$. Finally, since $x^\star(v)$ is clearly increasing in v, we have $(x^\star,t^\star) \in \mathcal{M}$.

For the optimality of $(x^\star,t^\star)$, it remains to show $\int_{u\in\mathcal{V}} \beta^\star(u)\mathrm{d}G(u) - \theta\alpha^\star = \pi^\star$. We first calculate

$$\begin{aligned} -\theta\alpha^\star &= -\alpha^\star \sum_{j\in[J]} \int_{u_j}^{w_j} \left(\bar{G}(v) - \pi^\star/v\right)\mathrm{d}v = \alpha^\star\pi^\star \sum_{j\in[J]} \ln(w_j/u_j) - \alpha^\star \sum_{j\in[J]} \int_{u_j}^{w_j} \bar{G}(v)\mathrm{d}v \\ &= \pi^\star - \alpha^\star \sum_{j\in[J]} w_j(1-G(w_j)) + \alpha^\star \sum_{j\in[J]} u_j(1-G(u_j)) - \alpha^\star \sum_{j\in[J]} \int_{u_j}^{w_j} v\, \mathrm{d}G(v), \end{aligned}$$

where the last identity follows from

$$\int_{u_j}^{w_j} \bar{G}(v)\,\mathrm{d}v = (w_j - u_j) - \int_{u_j}^{w_j} G(v)\,\mathrm{d}v = w_j(1 - G(w_j)) - u_j(1 - G(u_j)) + \int_{u_j}^{w_j} v \mathrm{d}\,G(v).$$

We next calculate

$$\begin{aligned}
&\int_{u\in\mathcal{V}} \beta^\star(u)\,\mathrm{d}G(u) \\
&= \sum_{j\in[J]} \int_{u_j}^{w_j} \left(\alpha^\star(v - u_j) + \alpha^\star \sum_{i=1}^{j-1}(w_i - u_i)\right) \mathrm{d}G(v) + \sum_{j\in[J]} \left(\alpha^\star \sum_{i=1}^{j}(w_i - u_i)\right)(G(u_{j+1}) - G(w_j)) \\
&= \alpha^\star \sum_{j\in[J]} \int_{u_j}^{w_j} v\,\mathrm{d}G(v) + \sum_{j\in[J]} \left[\left(\alpha^\star \sum_{i=1}^{j}(w_i - u_i)\right)(G(u_{j+1}) - G(u_j)) - \alpha^\star w_j(G(w_j) - G(u_j))\right] \\
&= \alpha^\star \sum_{j\in[J]} \int_{u_j}^{w_j} v\,\mathrm{d}G(v) + \alpha^\star \sum_{j\in[J]} w_j(1 - G(w_j)) - \alpha^\star \sum_{j\in[J]} u_j(1 - G(u_j)),
\end{aligned}$$

where the last line follows from the observations:

$$\begin{aligned}
&\sum_{j\in[J]} \left[\left(\alpha^\star \sum_{i=1}^{j} w_i\right)(G(u_{j+1}) - G(u_j)) - \alpha^\star w_j(G(w_j) - G(u_j))\right] = \alpha^\star \sum_{j\in[J]} w_j(1 - G(w_j)) \\
&\sum_{j\in[J]} \left(-\alpha^\star \sum_{i=1}^{j} u_i\right)(G(u_{j+1}) - G(u_j)) = -\alpha^\star \sum_{j\in[J]} u_j(1 - G(u_j)).
\end{aligned}$$

Finally, combining the expressions of $\int_{u\in\mathcal{V}} \beta^\star(u)\mathrm{d}G(u)$ and $-\theta\alpha^\star$ concludes the proof. $\square$

Proof of Proposition 7. Since $\|\bar{G}^{-1} - \bar{F}_p^{\star-1}\|_1 = \int_{\bar{G}(k_p^\star)}^{\bar{G}(p)} (\bar{G}^{-1}(q) - p)\mathrm{d}q \le \theta$, $F_p^\star \in \mathcal{F}_{(\theta,1)}$. We next prove the optimality and uniqueness of $F_p^\star$. Suppose there exists $F \in \mathcal{F}_{(\theta,1)}$ such that $\bar{F}(p) \le \bar{F}_p^\star(p)$. Then for $v \in [p, k_p^\star)$, we have $\bar{F}(v) \le \bar{F}(p) \le \bar{F}_p^\star(p) = \bar{G}(k_p^\star)$. For $q \in [\bar{G}(k_p^\star), \bar{G}(p))$, we have $\bar{F}_p^{\star-1}(q) = p$ and $q \ge \bar{F}_p^\star(p) \ge \bar{F}(p)$, implying $\bar{F}^{-1}(q) \le \bar{F}^{-1}(\bar{F}(p)) \le p = \bar{F}_p^{\star-1}(q)$. Similarly, since $\bar{F}_p^\star(v) \le \bar{G}(v)$ for $v \in [p, k_p^\star)$, we have $\bar{F}_p^{\star-1}(q) \le \bar{G}^{-1}(q)$ for $q \in [G(\bar{k}_p^\star), \bar{G}(p))$. Besides, since $\bar{F}_p^\star(v) = \bar{G}(v)$ for $v \notin [p, k_p^\star)$, then $|\bar{G}^{-1}(q) - \bar{F}_p^{-1}(q)| \ge 0 = |\bar{G}^{-1}(q) - \bar{F}_p^{\star-1}(q)|$ for $q \notin [\bar{G}(k_p^\star), \bar{G}(p))$. Thus, we obtain

$$\theta \ge \int_0^1 |\bar{G}^{-1}(q) - \bar{F}^{-1}(q)|\,\mathrm{d}q \ge \int_{\bar{G}(k_p^\star)}^{\bar{G}(p)} (\bar{G}^{-1}(q) - \bar{F}^{-1}(q))\,\mathrm{d}q \ge \int_{\bar{G}(k_p^\star)}^{\bar{G}(p)} (\bar{G}^{-1}(q) - \bar{F}_p^{\star-1}(q))\,\mathrm{d}q = \theta,$$

where the equality is due to $\int_{\bar{G}(k_p^\star)}^{\bar{G}(p)} (\bar{G}^{-1}(q) - \bar{F}_p^{\star-1}(q))\mathrm{d}q = d(F^\star, G) = \theta$ and the inequalities are indeed tight. Hence,

$$\int_{\bar{G}(k_p^\star)}^{\bar{G}(p)} (\bar{G}^{-1}(q) - \bar{F}^{-1}(q))\mathrm{d}q = \int_{\bar{G}(k_p^\star)}^{\bar{G}(p)} (\bar{G}^{-1}(q) - \bar{F}_p^{\star-1}(q))\mathrm{d}q, \quad \textit{i.e.}, \quad \int_{\bar{G}(k_p^\star)}^{\bar{G}(p)} (\bar{F}_p^{\star-1}(q) - \bar{F}^{-1}(q))\mathrm{d}q = 0.$$

Since $\bar{F}^{-1}(q) \le \bar{F}_q^{\star-1}(q)$ for $q \in [\bar{G}(k_p^\star), \bar{G}(p))$, by right continuity, $\bar{F}^{-1}(q) = \bar{F}_p^{\star-1}(q)$ for $q \in [\bar{G}(k_p^\star), \bar{G}(p)]$. Besides, since $\int_0^1 |\bar{G}^{-1}(q) - \bar{F}^{-1}(q)|\mathrm{d}q = \int_{\bar{G}(k_p^\star)}^{\bar{G}(p)} (\bar{G}^{-1}(q) - \bar{F}^{-1}(q))\mathrm{d}q$, it holds that $\bar{F}^{-1}(q) = \bar{G}^{-1}(q) = \bar{F}_p^{\star-1}(q)$ for $q \notin [\bar{G}(k_p^\star), \bar{G}(p)]$. In summary, we have $\bar{F} = \bar{F}_p^\star$ everywhere, and $F_p^\star$ is the unique worst-case distribution. $\square$

Appendix C Supplementary Results

C.1 Mean and Variance Ambiguity Set

In this section, we consider the ambiguity set with the dispersion information specified by the second moment (or variance) of the distribution:

$$\mathcal{F}_{\mathrm{V}} = \left\{ F \in \mathcal{P} \,\middle|\, \int_{\mathcal{V}} v \,\mathrm{d}F(v) = \mu,\ \int_{\mathcal{V}} (v-\mu)^2 \,\mathrm{d}F(v) = \sigma^2 \right\},$$

where $0 \le \sigma^2 \le \mu(1-\mu)$ so that $\mathcal{F}_{\mathrm{V}}$ is nonempty. Let σ_{M}^2 be the variance of the worst-case distribution derived in Theorem 1 for $\mathcal{F}_{\mathrm{M}}$. If $\sigma^2 = \sigma_{\mathrm{M}}^2$ in $\mathcal{F}_{\mathrm{V}}$, then the worst-case distribution for the mean ambiguity set is also feasible and hence optimal for $\mathcal{F}_{\mathrm{V}}$. If $\sigma^2 < \sigma_{\mathrm{M}}^2$ ($\sigma^2 > \sigma_{\mathrm{M}}^2$), based on the worst-case distribution for the mean ambiguity set, one can move some mass from larger (smaller) valuations to smaller (larger) valuations to reduce (increase) the variance; see Figure 10. For the case of $\sigma^2 < \sigma_{\mathrm{M}}^2$, reducing the variance is essentially the same as increasing the contraction (*i.e.*, reducing the spread) of the distribution as we did in Section 4.2, and we can similarly consider a parametric family

$$\bar{Q}_{k,1}(v) = \begin{cases} 1 & v \in [0, \pi_k) \\ \pi_k / v & v \in [\pi_k, k) \\ 0 & v \in [k, 1], \end{cases}$$

where $k \in [\mu, 1]$ and π_k is the unique solution in $[0,1]$ to $\pi + \pi \ln(k/\pi) = \mu$. In comparison, to increase the variance, we further define a family

$$\bar{Q}_{1,\ell}(v) = \begin{cases} \ell & v \in [0, \pi_\ell/\ell) \\ \pi_\ell / v & v \in [\pi_\ell/\ell, 1) \\ 0 & v = 1, \end{cases}$$

where $\ell \in [\mu, 1]$ and π_ℓ is the unique solution in $[0,1]$ to $\pi + \pi \ln(\ell/\pi) = \mu$. As k increases from μ to 1, the variance of $\bar{Q}_{k,1}$ increases from 0 to σ_{M}^2. As ℓ decreases from 1 to μ, the variance of $\bar{Q}_{1,\ell}$ further increases from σ_{M}^2 to $\mu(1-\mu)$—the largest possible variance for any distribution supported on $[0,1]$ with mean μ. Hence, one can always find either k or ℓ such that the corresponding distribution within the two parametric families has variance σ^2. The next proposition confirms that the obtained distribution is indeed optimal.

THEOREM 5. *Given $\mathcal{F}_{\mathrm{V}}$, the unique worst-case distribution $F^\star$ and the optimal revenue $\pi^\star$ of problem* (4) *can be characterized as follows.*

(i) If $\sigma \le \sigma_{\mathrm{M}}$, then $\bar{F}^\star = \bar{Q}_{k^\star,1}$ and $\pi^\star = \pi_{k^\star}$, where $k^\star$ is the unique solution in $[\mu, 1]$ such that the variance of $\bar{Q}_{k,1}$ equals to σ^2.

(ii) If $\sigma > \sigma_{\mathrm{M}}$, then $\bar{F}^\star = \bar{Q}_{1,\ell^\star}$ and $\pi^\star = \pi_{\ell^\star}$, where $\ell^\star$ is the unique solution in $[\mu, 1]$ such that the variance of $\bar{Q}_{1,\ell}$ equals to σ^2.

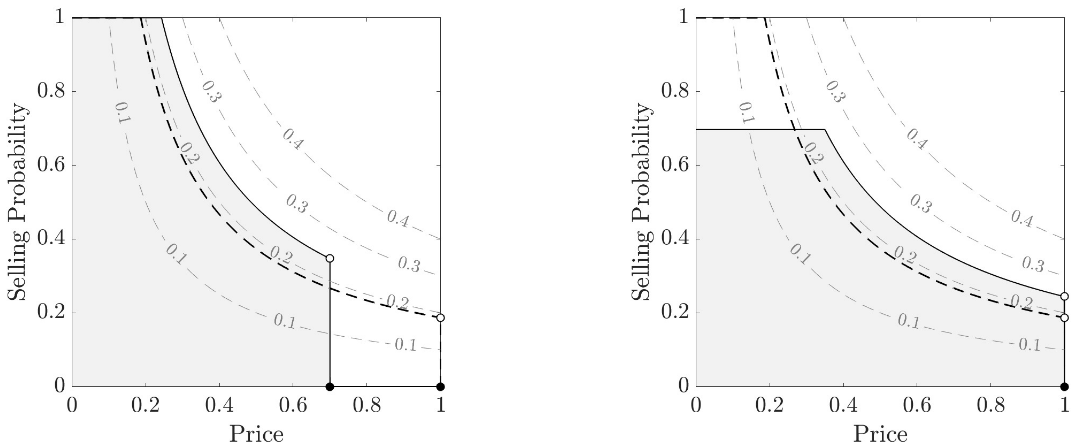

Figure 10 The worst-case distribution for the mean ambiguity set characterized in Theorem 1 is indicated by the dashed curve. In the left (right) panel, the solid curve is formed by moving mass from larger (smaller) valuations and spreading it over smaller (larger) valuations along the level curves, and the resulting variance becomes smaller (larger).

Proof of Theorem 5. We first calculate $k^\star$ (or $\ell^\star$) and $\pi^\star$, and then prove the uniqueness as well as the optimality. Note that the variance constraint is equivalent to $\int_{\mathcal{V}} v\bar{F}(v)\mathrm{d}v = (\mu^2+\sigma^2)/2$.

For case (i), the mean and variance constraints give two equations with solutions as follows:

$$\begin{cases} \mu = \displaystyle\int_{\mathcal{V}} \bar{F}^\star(v)\mathrm{d}v = \pi^\star + \pi^\star \ln\left(\frac{k^\star}{\pi^\star}\right) \\ \dfrac{\mu^2+\sigma^2}{2} = \displaystyle\int_{\mathcal{V}} v\bar{F}^\star(v)\mathrm{d}v = \frac{(\pi^\star)^2}{2} + \pi^\star(k^\star - \pi^\star) \end{cases} \Longrightarrow \begin{cases} \pi^\star = h_1^{-1}\left(\dfrac{\mu^2+\sigma^2}{2}\right) \\ k^\star = \pi^\star \exp\left(\dfrac{\mu-\pi^\star}{\pi^\star}\right), \end{cases}$$

where $h_1(z) = z^2(e^{(\mu-z)/z} - 1/2)$. Suppose there exists $F \in \mathcal{F}_{\mathrm{V}}$ such that $\sup_{p\in\mathcal{V}} p\bar{F}(p) \le \max_{p\in\mathcal{V}} p\bar{F}^\star(p)$. Then for $v \in [\pi^\star, k^\star)$, it holds that $v\bar{F}(v) \le \pi^\star = v\bar{F}^\star(v)$, *i.e.*, $\bar{F}(v) \le \bar{F}^\star(v)$. Hence,

$$\begin{aligned} \int_{\mathcal{V}} v\bar{F}(v)\,\mathrm{d}v - \int_{\mathcal{V}} v\bar{F}^\star(v)\,\mathrm{d}v &= \int_0^{k^\star} v(\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v + \int_{k^\star}^1 v(\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v \\ &\ge k^\star \int_0^{k^\star} (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v + \int_{k^\star}^1 v(\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v \\ &\ge k^\star \int_0^{k^\star} (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v + k^\star \int_{k^\star}^1 (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v = 0, \end{aligned} \tag{12}$$

where the first inequality follows from $\bar{F}(v) \le \bar{F}^\star(v) = 1$ for $v \in [0, \pi^\star)$ and $\bar{F}(v) \le \bar{F}^\star(v)$ for $v \in [\pi^\star, k^\star)$; the second inequality is due to $\bar{F}(v) \ge \bar{F}^\star(v) = 0$ for $v \in [k^\star, 1]$; and the last identity follows from the mean constraint $\int_{\mathcal{V}} \bar{F}(v)\mathrm{d}v = \int_{\mathcal{V}} \bar{F}^\star(v)\mathrm{d}v = \mu$. Now we have $\int_{\mathcal{V}} v\bar{F}(v)\mathrm{d}v \ge \int_{\mathcal{V}} v\bar{F}^\star(v)\mathrm{d}v = (\mu^2+\sigma^2)/2$. The equality holds only if $\int_0^{k^\star}(v-k^\star)(\bar{F}(v) - \bar{F}^\star(v))\mathrm{d}v = 0$ and $\int_{k^\star}^1 (v-k^\star)(\bar{F}(v) - \bar{F}^\star(v))\mathrm{d}v = 0$. Recall that $\bar{F}(v) \le \bar{F}^\star(v)$ for $v \in [0, k^\star)$, and $\bar{F}(v) \ge \bar{F}^\star(v)$ for $v \in [k^\star, 1]$. By right-continuity, the equality holds only if $F = F^\star$, implying that $F^\star$ is the unique worst-case distribution.

For case (ii), the mean and variance constraints give two equations with solutions as follows:

$$\begin{cases} \mu = \int_{\mathcal{V}} \bar{F}^\star(v)\mathrm{d}v = \pi^\star + \pi^\star \ln\left(\dfrac{\ell^\star}{\pi^\star}\right) \\ \dfrac{\mu^2+\sigma^2}{2} = \int_{\mathcal{V}} v\bar{F}^\star(v)\mathrm{d}v = \dfrac{(\pi^\star/\ell^\star)^2}{2} + \pi^\star(1-\pi^\star/\ell^\star) \end{cases} \implies \begin{cases} \pi^\star = \dfrac{\mu}{1-\ln(k_0)} \\ \ell^\star = \dfrac{\pi^\star}{k_0}, \end{cases}$$

where $k_0 = h_2^{-1}((\mu^2+\sigma^2)/2)$ and $h_2(z) = (\mu(2-z))/(2-2\ln(z))$. Suppose there exists $F \in \mathcal{F}_{\mathrm{V}}$ satisfying $\sup_{p\in\mathcal{V}} p\bar{F}(p) \le \max_{q\in\mathcal{V}} p\bar{F}^\star(p)$. Then for $v \in [k_0, 1)$, $v\bar{F}(v) \le \sup_{p\in\mathcal{V}} p\bar{F}(p) \le \max_{p\in\mathcal{V}} p\bar{F}^\star(p) = v\bar{F}^\star(v)$, *i.e.*, $\bar{F}(v) \le \bar{F}^\star(v)$. Let $v' = \sup\{v \in [0, k_0) \mid \bar{F}(v) \ge \ell^\star\}$, which must exist because otherwise, $\int_{\mathcal{V}} \bar{F}(v)\mathrm{d}v < \int_{\mathcal{V}} \bar{F}^\star(v)\mathrm{d}v = \mu$. Since $\bar{F}$ is decreasing, we have $\bar{F}(v) \ge \bar{F}^\star(v) = \ell$ for $v \in [0, v')$ and $\bar{F}(v) \le \bar{F}^\star(v)$ for $v \in [v', k_0)$. In summary, these relations yield

$$\begin{aligned} \int_{\mathcal{V}} v\bar{F}(v)\,\mathrm{d}v - \int_{\mathcal{V}} v\bar{F}^\star(v)\,\mathrm{d}v &= \int_0^{v'} v(\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v + \int_{v'}^1 v(\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v \\ &\le v' \int_0^{v'} (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v + \int_{v'}^1 v(\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v \\ &\le v' \int_0^{v'} (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v + v' \int_{v'}^1 (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v = 0. \end{aligned} \tag{13}$$

Now we obtain that $\int_{\mathcal{V}} v\bar{F}(v)\mathrm{d}v \le \int_{\mathcal{V}} v\bar{F}^\star(v)\mathrm{d}v = (\mu^2+\sigma^2)/2$. The equality holds only if $\int_0^{v'}(v - v')(\bar{F}(v) - \bar{F}^\star(v))\mathrm{d}v = 0$ and $\int_{v'}^1 (v-v')(\bar{F}(v) - \bar{F}^\star(v))\mathrm{d}v = 0$. Recall that $\bar{F}(v) \ge \bar{F}^\star(v)$ for $v \in [0, v')$, and $\bar{F}(v) \le \bar{F}^\star(v)$ for $v \in [v', 1]$. By right-continuity, the equality holds only if $F = F^\star$, implying that $F^\star$ is the unique worst-case distribution. □

We now characterize the optimal mechanism to the primal problem (2).

PROPOSITION 8. *Given $\mathcal{F}_{\mathrm{V}}$, the optimal mechanism $(x^\star, t^\star)$ to problem (2) is characterized as:*

(i) If $\sigma \le \sigma_{\mathrm{M}}$, then

$$(x^\star(v), t^\star(v)) = \begin{cases} (0,\ 0) & v \in [0, \pi^\star) \\ \left(2\gamma^\star(v-\pi^\star) + \alpha^\star \ln\left(\dfrac{v}{\pi^\star}\right),\ \gamma^\star v^2 + \alpha^\star v + \beta^\star\right) & v \in [\pi^\star, k^\star) \\ (1,\ \gamma^\star(k^\star)^2 + \alpha^\star k^\star + \beta^\star) & v \in [k^\star, 1], \end{cases}$$

where $k^\star$ and $\pi^\star$ are defined in case (i) of Theorem 5, and other parameters are given by

$$\gamma^\star = \frac{1}{2(k^\star - \pi^\star - k^\star \ln(k^\star/\pi^\star))} < 0,\ \beta^\star = \pi^\star(2k^\star - \pi^\star)\gamma^\star,\ \alpha^\star = -2k^\star\gamma^\star.$$

(ii) If $\sigma > \sigma_{\mathrm{M}}$, then

$$(x^\star(v), t^\star(v)) = \begin{cases} (0,\ 0) & v \in [0, k_0) \\ \left(2\gamma^\star(v-k_0) + \alpha^\star \ln\left(\dfrac{v}{k_0}\right),\ \gamma^\star v^2 + \alpha^\star v\right) & v \in [k_0, 1], \end{cases}$$

where $\ell^\star$ and $\pi^\star$ are defined in case (ii) of Theorem 5, and other parameters are given by

$$k_0 = \frac{\pi^\star}{\ell^\star},\ \gamma^\star = \frac{1}{2(1-k_0) + k_0\ln(k_0)} > 0,\ \alpha^\star = -k_0\gamma^\star.$$

Proof of Proposition 8. Given $\mathcal{F}_{\mathrm{V}}$, the inner minimization of problem (2) is

$$\begin{aligned} \inf_{F\in\mathcal{P}_+} \quad & \int_{\mathcal{V}} t(v)\,\mathrm{d}F(v) \\ \text{s.t.} \quad & \int_{\mathcal{V}} v\,\mathrm{d}F(v)=\mu \\ & \int_{\mathcal{V}} v^2\,\mathrm{d}F(v)=\mu^2+\sigma^2 \\ & \int_{\mathcal{V}} \mathrm{d}F(v)=1. \end{aligned} \tag{14}$$

Combining the dual of problem (14) with the outer maximization, problem (2) becomes

$$\begin{aligned} \sup_{x,t,\alpha,\beta,\gamma} \quad & \mu\alpha+\beta+(\mu^2+\sigma^2)\gamma \\ \text{s.t.} \quad & p(v)\geq \gamma v^2+\alpha v+\beta \qquad\qquad \forall v\in\mathcal{V} \\ & (x,t)\in\mathcal{M},\ \alpha\in\mathbb{R},\ \beta\in\mathbb{R},\ \gamma\in\mathbb{R}. \end{aligned} \tag{15}$$

In the remainder, for each case, we start by showing the feasibility of the proposed solution, then we verify that it attains an objective value equal to the optimal revenue of problem (4).

For case (i), note that $\pi^\star\leq k^\star$ and $k^\star-\pi^\star-k^\star\ln(k^\star/\pi^\star)=k^\star(1+\ln(\pi^\star/k^\star)-\pi^\star/k^\star)\leq 0$. Since $\pi^\star=k^\star$ holds only in the trivial situation where $\mathcal{F}_{\mathrm{V}}=\{\delta_\mu\}$, we focus on $\pi^\star<k^\star$. Thus, $\gamma^\star<0$ and it follows that the concave quadratic function $\gamma^\star v^2+\alpha^\star v+\beta^\star$ is increasing (resp., decreasing) in v when $v\leq k^\star$ (resp., $v\geq k^\star$), implying $\gamma^\star v^2+\alpha^\star v+\beta^\star\leq \gamma^\star\pi^{\star 2}+\alpha^\star\pi^\star+\beta^\star=0=t(v)$ for $v\in[0,\pi^\star)$ and $\gamma^\star v^2+\alpha^\star v+\beta^\star\leq\gamma^\star k^{\star 2}+\alpha^\star k^\star+\beta^\star=t(v)$ for $v\in[k^\star,1]$. By definition, $t(v)=\gamma^\star v^2+\alpha^\star v+\beta^\star$ for $v\in[\pi^\star,k^\star)$. Thus the first constraint in (15) holds. Besides, we have $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=2\gamma^\star v+\alpha^\star$ for $v\in[\pi^\star,k^\star)$ and $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=0$ for $v\in[k^\star,1]$. These relations, together with the fact that $t(v)=x(v)=0$ for $v\in[0,\pi^\star)$, yield $t^\star(v)=vx^\star(v)-\int_0^v x^\star(u)\mathrm{d}u$. Since $x^\star$ is increasing, we have $(x^\star,t^\star)\in\mathcal{M}$, ensuring that $(x^\star,t^\star,\alpha^\star,\beta^\star,\gamma^\star)$ is feasible. We next verify the objective value:

$$\begin{aligned} \mu\alpha^\star+\beta^\star+(\mu^2+\sigma^2)\gamma^\star &= \gamma^\star\left(-2k^\star\mu+\pi^\star(2k^\star-\pi^\star)+2(\pi^\star)^2\exp\left(\frac{\mu}{\pi^\star}-1\right)-\pi^{\star 2}\right) \\ &= 2\gamma^\star\left(2k^\star\pi^\star-(\pi^\star)^2-k^\star\left(\pi^\star+\pi^\star\ln\left(\frac{k^\star}{\pi^\star}\right)\right)\right) \\ &= 2\gamma^\star\pi^\star\left(k^\star-\pi^\star-k^\star\ln\left(\frac{k^\star}{\pi^\star}\right)\right)=\pi^\star, \end{aligned}$$

where the second identity is due to $\pi^\star+\pi^\star\ln(k^\star/\pi^\star)=\mu$. Therefore, this feasible solution is indeed the optimal mechanism to problem (2).

For case (ii), we consider the solution $(x^\star,t^\star,\alpha^\star,\beta^\star,\gamma^\star)$ with $\beta^\star=0$. Firstly, since $k_0\leq 1$ and $2z-z\ln(z)$ is strictly increasing in z, we have $2(1-k_0)+k_0\ln(k_0)=2-(2k_0-k_0\ln(k_0))\geq 2-(2-1\ln(1))=0$. Since $k_0=1$ holds only in the trivial situation where $\mathcal{F}_{\mathrm{V}}$ is a singleton containing only the Bernoulli distribution, we focus on $k_0<1$. Thus, $\gamma^\star>0$ and it follows that the convex quadratic function $\gamma^\star v^2+\alpha^\star v+\beta^\star$ takes a value of 0 when $v=0$ and $v=k_0$, implying $\gamma^\star v^2+\alpha^\star v+\beta^\star \le 0 = t(v)$ when $v\in[0,k_0)$. By definition, $t(v)=\gamma^\star v^2+\alpha^\star v+\beta^\star$ for $v\in[k_0,1]$. Hence, the first constraint in (15) holds. Secondly, we have $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=2\gamma^\star v+\alpha^\star$ for $v\in[k_0,1]$ and $t(v)=x(v)=0$ for $v\in[0,k_0)$. It follows that $t^\star(v)=vx^\star(v)-\int_0^v x^\star(u)\mathrm{d}u$. Since $x^\star(v)$ is increasing in v, we now have $(x^\star,t^\star)\in\mathcal{M}$, ensuring that $(x^\star,t^\star,\alpha^\star,\beta^\star,\gamma^\star)$ is feasible. We next verify the objective value:

$$\mu\alpha^\star+\beta^\star+(\mu^2+\sigma^2)\gamma^\star=\gamma^\star\left(-k_0\mu+\frac{\mu(2-k_0)}{1-\ln(k_0)}\right)=\gamma^\star\pi^\star((2-k_0)-k_0(1-\ln(k_0)))=\pi^\star.$$

Therefore, this feasible solution is indeed the optimal mechanism to problem (2). □

Carrasco et al. (2018) also study the ambiguity set with variance information with the difference that they consider possibly unbounded valuation and all distributions having variance bounded from above by σ^2 (as opposed to being equal to σ^2 in our setting). Nevertheless, our characterization for the case $\sigma\le\sigma_{\mathrm{M}}$ in Proposition 8 is essentially the same as that in Carrasco et al. (2018) (see proposition 4). The case of $\sigma>\sigma_{\mathrm{M}}$, however, is new.[15] The discrepancy can be explained by the fact that our worst-case revenue π^* is decreasing in σ for $\sigma\le\sigma_{\mathrm{M}}$ but is increasing in σ for $\sigma>\sigma_{\mathrm{M}}$. As a result, if one considers all possible distributions with variance bounded from above by σ^2, then the worst-case distribution is always obtained at the one with variance σ^2 or σ_{M}^2, whichever is smaller, and the resulting optimal mechanism is always a menu of lotteries with logarithmic winning probabilities paired with a *concave* price menu. In comparison, our Proposition 8 shows that when the buyer's valuation is known to have larger variability, then it is optimal to adopt a menu of lotteries with logarithmic winning probabilities but paired with a *convex* price menu.

We next apply the geometric approach to solve the maximin pricing problem (3). When $p<\mu$, let σ_p^2 be the variance of the unique worst-case distribution to $\inf_{F\in\mathcal{F}_{\mathrm{M}}}\bar{F}(p)$ (see Proposition 3). Then if $\sigma<\sigma_p$, following the same intuition for the mean-preserving contraction ambiguity set, one could squeeze the mass between $[p,1]$ to a smaller interval $[p,k]$ with some $k<1$ and obtain a distribution with a smaller variance; while if $\sigma>\sigma_p$, to obtain a distribution with a larger variance, one could move some mass between $[0,p]$ to the interval $[p,1]$. In comparison, when $p\ge\mu$, the worst-case distribution to $\inf_{F\in\mathcal{F}_{\mathrm{M}}}\bar{F}(p)$ is not necessarily unique, and one has to compare σ with the largest possible variance. We formally characterize the worst-case distribution under a posted price mechanism in the proposition below.

Proposition 9. *Given $\mathcal{F}_{\mathrm{V}}$, the worst-case distribution $F_p^\star$ to the problem $\inf_{F\in\mathcal{F}_{\mathrm{V}}}\bar{F}(p)$ can be characterized as follows.*

[15] Our analysis for the case of $\sigma>\sigma_{\mathrm{M}}$ can be similarly applied to an ambiguity set of mean-preserving spreads of the reference distribution (as opposed to the mean-preserving contraction ambiguity set analyzed in Section 4.2); see detailed discussions in Appendix C.3.

(i) If $p \in (0, \mu)$, then for $\sigma^2 \in [0, \sigma_p^2)$, the unique worst-case distribution is

$$\bar{F}_p^\star(v) = \begin{cases} 1 & v \in [0, p) \\ (\mu - p)/(k_p^\star - p) & v \in [p, k_p^\star) \\ 0 & v \in [k_p^\star, 1] \end{cases} \quad \text{with } k_p^\star = \mu + \frac{\sigma^2}{\mu - p};$$

while for $\sigma^2 \in [\sigma_p^2, \mu(1-\mu)]$, the unique worst-case distribution is

$$\bar{F}_p^\star(v) = \begin{cases} \ell_p^\star & v \in [0, p) \\ (\mu - p\ell_p^\star)/(1-p) & v \in [p, 1) \\ 0 & v = 1 \end{cases} \quad \text{with } \ell_p^\star = \frac{\mu(1+p) - (\mu^2 + \sigma^2)}{p}.$$

(ii) If $p \in [\mu, 1]$, then for $\sigma^2 \in [0, \mu(p-\mu))$, a worst-case distribution is

$$\bar{F}_p^\star(v) = \begin{cases} \mu^2/(\mu^2 + \sigma^2) & v \in [0, \mu + \sigma^2/\mu) \\ 0 & v \in [\mu + \sigma^2/\mu, 1]; \end{cases}$$

while for $\sigma^2 \in [\mu(p-\mu), \mu(1-\mu)]$, the unique worst-case distribution is

$$\bar{F}_p^\star(v) = \begin{cases} \ell_p^\star & v \in [0, p) \\ (\mu - p\ell_p^\star)/(1-p) & v \in [p, 1) \\ 0 & v = 1 \end{cases} \quad \text{with } \ell_p^\star = \frac{\mu(1+p) - (\mu^2 + \sigma^2)}{p}.$$

Proof of Proposition 9. It is not hard to verify that $F_p^\star$ in each case indeed has mean μ and variance σ^2. Hence, $F_p^\star \in \mathcal{F}_{\mathrm{V}}$, and we focus on proving optimality as well as uniqueness.

For case (i), we first consider $\sigma^2 \in [0, \sigma_p^2)$. Suppose there exists $F \in \mathcal{F}_{\mathrm{V}}$ such that $\bar{F}(p) \le \bar{F}_p^\star(p)$. Then $\bar{F}(v) \le 1 = \bar{F}_p^\star(v)$ for $v \in [0, p)$ and $\bar{F}(v) \le \bar{F}(p) \le \bar{F}_p^\star(p) = \bar{F}_p^\star(v)$ for $v \in [p, k_p^\star)$, implying $\int_0^{k_p^\star} \bar{F}(v)\mathrm{d}v \le \int_0^{k_p^\star} \bar{F}_p^\star(v)\mathrm{d}v$. Since $\int_{\mathcal{V}} \bar{F}(v)\mathrm{d}v = \int_{\mathcal{V}} \bar{F}_p^\star(v)\mathrm{d}v = \mu$, we have $\int_{k_p^\star}^1 \bar{F}(v)\mathrm{d}v \ge \int_{k_p^\star}^1 \bar{F}_p^\star(v)\mathrm{d}v$. Following similar lines as in (12), we obtain $F = F_p^\star$ by right continuity, concluding $F_p^\star$ is the unique worst-case distribution. We next consider $\sigma^2 \in [\sigma_p^2, \mu(1-\mu)]$. Suppose there exists $F \in \mathcal{F}_{\mathrm{V}}$ such that $\bar{F}(p) \le \bar{F}_p^\star(p)$. For $v \in [p, 1)$, it holds that $\bar{F}(v) \le \bar{F}(p) \le \bar{F}_p^\star(p) = \bar{F}_p^\star(v)$, *i.e.*, $\bar{F}(v) \le \bar{F}_p^\star(v)$. Let $v' = \sup\{v \in [0, p) \mid \bar{F}(v) > \ell_p^\star\}$. If v' does not exist, then $\mu = \int_{\mathcal{V}} \bar{F}(v)\mathrm{d}v \le \int_{\mathcal{V}} \bar{F}_p^\star(v)\mathrm{d}v = \mu$, leading to $F = F_p^\star$ by right continuity and concluding $F_p^\star$ is the unique worst-case distribution. If v' exists, then because $\bar{F}$ is decreasing, $\bar{F}(v) \le \bar{F}_p^\star(v)$ for $v \in [v', p)$ and $\bar{F}(v) > \bar{F}_p^\star(v) = \ell_p^\star$ for $v \in [0, v')$. Following similar lines as in (13), we arrive at $F = F_p^\star$. Hence, $F_p^\star$ is the unique worst-case distribution.

For case (ii), when $\sigma^2 \in [0, \mu(q-\mu)]$, $\bar{F}_p^\star(p) = 0$ and $\bar{F}_p^\star$ is a worst-case distribution. When $\sigma^2 \in (\mu(p-\mu), \mu(1-\mu)]$, the proof is similar to that of case (i) and is thus omitted. $\square$

With the worst-case distribution in Proposition 9, one can find the optimal posted price $p^\star \in [0, 1]$. Chen et al. (2020a) also study the robust pricing problem with a mean and variance ambiguity set but assuming unbounded support. In their setting, the optimal price will never be larger than μ. This is because when $p > \mu$, due to the unbounded support, the adversary can always choose a two-point distribution with an arbitrary large valuation so that $\bar{F}(p)$ can be arbitrarily small and at the same time satisfying the mean and variance constraints. In our setting, however, it is possible that the optimal price is larger than μ: for example, when $0<\mu<1$ and $\sigma=\sqrt{\mu(1-\mu)}$ (*i.e.*, $\mathcal{F}_{\mathrm{V}}$ is a singleton containing only the Bernoulli distribution), the optimal price $p^\star=1>\mu$. Besides, when $p\leq\mu$, Chen et al. (2020a) characterize the worst-case distribution based on Cantelli's inequality (see their lemma 1). If $\sigma\leq\sigma_p$, our bound in Proposition 9 also reduces to Cantelli's inequality. However, when $\sigma>\sigma_p$, our Proposition 9 characterizes the worst-case distribution that achieves a tighter bound than the Cantelli's inequality, which does not utilize the support information.

C.2 Mean Absolute Deviation Ambiguity Set

The mean absolute deviation (MAD) ambiguity set is given by

$$\mathcal{F}_{\mathrm{MAD}}=\left\{F\in\mathcal{P}\;\middle|\;\int_{\mathcal{V}}v\,\mathrm{d}F(v)=\mu,\;\int_{\mathcal{V}}|v-\mu|\,\mathrm{d}F(v)=\kappa\right\},$$

where $0\leq\kappa\leq2\mu(1-\mu)$ so that $\mathcal{F}_{\mathrm{MAD}}$ is nonempty. Let κ_{M} be the MAD of the worst-case distribution derived in Theorem 1 for $\mathcal{F}_{\mathrm{M}}$. Again, if $\kappa=\kappa_{\mathrm{M}}$ in $\mathcal{F}_{\mathrm{MAD}}$, then the worst-case distribution for $\mathcal{F}_{\mathrm{M}}$ is also feasible and hence optimal. If $\kappa<\kappa_{\mathrm{M}}$ ($\kappa>\kappa_{\mathrm{M}}$), similar techniques as in the case of mean and variance ambiguity set can be applied here to decrease (increase) κ_{M} to match κ. Consider the two parametric families $\{\bar{Q}_{k,1}\}_k$ (such that as k increases from μ to 1, the MAD of $\bar{Q}_{k,1}$ increases from 0 to κ_{M}) and $\{\bar{Q}_{1,\ell}\}_\ell$ (such that as ℓ decreases from 1 to μ, the MAD of $\bar{Q}_{1,\ell}$ increases from κ_{M} to $2\mu(1-\mu)$), we can then find either k or ℓ such that the corresponding distribution has MAD κ.

THEOREM 6. *Given $\mathcal{F}_{\mathrm{MAD}}$, the worst-case distribution $F^\star$ and the optimal revenue $\pi^\star$ of problem* (4) *can be characterized as follows.*

(i) If $\kappa\in[0,\kappa_{\mathrm{M}})$, then $\bar{F}^\star=\bar{Q}_{k^\star,1}$ and $\pi^\star=\pi_{k^\star}$, where $k^\star$ is the unique solution in $[\mu,1]$ such that the MAD of $\bar{Q}_{k,1}$ equals to κ.

(ii) If $\kappa\in[\kappa_{\mathrm{M}},\kappa_0)$, then $\bar{F}^\star=\bar{Q}_{1,\ell^\star}$ and $\pi^\star=\pi_{\ell^\star}$, where $\ell^\star$ is the unique solution in $[\mu,1]$ such that the MAD of $\bar{Q}_{1,\ell}$ equals to κ.

(iii) If $\kappa\in[\kappa_0,2\mu(1-\mu)]$, then $\bar{F}^\star=\bar{Q}_{1,\ell^\star}$ and $\pi^\star=\pi_{\ell^\star}$, where $\ell^\star$ is the unique solution in $[\mu,1]$ such that the MAD of $\bar{Q}_{1,\ell}$ equals to κ. In addition, $\bar{F}^\star$ is the unique worst-case distribution.

Here, $\kappa_0=-2\mu\ln(\mu)/(1-\ln(\mu))$, and when $\kappa=\kappa_0$, the corresponding $\pi^\star$ and $\ell^\star$ satisfy $\pi^\star/\ell^\star=\mu$.

Proof of Theorem 6. For each case, we first calculate $k^\star$ (or $\ell^\star$) and $\pi^\star$ so that $Q_{k^\star,1}$ (or $Q_{1,\ell^\star}$) has mean μ and MAD κ, then we prove the optimality. Note that the mean constraint is equivalent to $\int_0^\mu\bar{F}(v)\mathrm{d}v+\int_\mu^1\bar{F}(v)\mathrm{d}v=\mu$ while the MAD constraint is equivalent to $\int_{\mathcal{V}}|v-\mu|\mathrm{d}F(v)=\mu+\int_\mu^1\bar{F}(v)\mathrm{d}v-\int_0^\mu\bar{F}(v)\mathrm{d}v=\kappa$. Hence, any $F\in\mathcal{F}_{\mathrm{MAD}}$ must satisfy $\int_0^\mu\bar{F}(v)\mathrm{d}v=\mu-\kappa/2$ and $\int_\mu^1\bar{F}(v)\mathrm{d}v=\kappa/2$.

For case (i), the two constraints give

$$\begin{cases} \mu - \dfrac{\kappa}{2} = \displaystyle\int_0^{\mu} \bar{F}^\star(v)\mathrm{d}v = \pi^\star + \pi^\star \ln\left(\frac{\mu}{\pi^\star}\right) \\ \dfrac{\kappa}{2} = \displaystyle\int_{\mu}^{1} \bar{F}^\star(v)\mathrm{d}v = \pi^\star \ln\left(\frac{k^\star}{\mu}\right) \end{cases} \implies \begin{cases} \pi^\star = h^{-1}\left(\mu - \dfrac{\kappa}{2}\right) \\ k^\star = \pi^\star \exp\left(\dfrac{\mu - \pi^\star}{\pi^\star}\right), \end{cases}$$

where $h(z) = z + z\ln(\mu/z)$. If there exists $F \in \mathcal{F}_{\mathrm{MAD}}$ such that $\sup_{p\in\mathcal{V}} p\bar{F}(p) < \sup_{p\in\mathcal{V}} p\bar{F}^\star(p)$, then for $v \in [\pi^\star, \mu)$, we have $v\bar{F}(v) < \pi^\star = v\bar{F}^\star(v)$, *i.e.*, $\bar{F}(v) < \bar{F}^\star(v)$. As a result, it holds that

$$\begin{aligned} \int_0^{\mu} \bar{F}(v)\,\mathrm{d}v - \int_0^{\mu} \bar{F}^\star(v)\,\mathrm{d}v &= \int_0^{\pi^\star} (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v + \int_{\pi^\star}^{\mu} (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v \\ &= \int_0^{\pi^\star} (\bar{F}(v) - 1)\,\mathrm{d}v + \int_{\pi^\star}^{\mu} (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v \\ &\le \int_{\pi^\star}^{\mu} (\bar{F}(v) - \bar{F}^\star(v))\,\mathrm{d}v < 0. \end{aligned}$$

That is, $\int_0^{\mu} \bar{F}(v)\,\mathrm{d}v < \int_0^{\mu} \bar{F}^\star(v)\,\mathrm{d}v = \mu - \kappa/2$, contradicting to $F \in \mathcal{F}_{\mathrm{MAD}}$.

Note that when $\kappa = \kappa_0$, we have $\ell^\star = 1/(1-\ln(\mu))$, $\pi^\star = \mu/(1-\ln(\mu))$ and $\pi^\star/\ell^\star = \mu$. For case (ii), $\kappa < \kappa_0$ ensures that when $\ell \le 1/(1-\ln(\mu))$ (*i.e.*, $\pi_\ell/\ell \ge \mu$), we must have $\int_\mu^1 \bar{Q}_{1,\ell}(v)\mathrm{d}v > \kappa/2$. Thus, $\ell^\star$ must lie in $(1/(1-\ln(\mu)), 1]$ and $\pi^\star/\ell^\star \in (0, \mu)$. Then we can solve $\ell^\star$ and $\pi^\star$ via:

$$\begin{cases} \mu - \dfrac{\kappa}{2} = \displaystyle\int_0^{\mu} \bar{F}^\star(v)\mathrm{d}v = \pi^\star + \pi^\star \ln\left(\frac{\mu}{\pi^\star/\ell^\star}\right) \\ \dfrac{\kappa}{2} = \displaystyle\int_{\mu}^{1} \bar{F}^\star(v)\mathrm{d}v = \pi^\star \ln\left(\frac{1}{\mu}\right) \end{cases} \implies \begin{cases} \pi^\star = -\dfrac{\kappa}{2\ln(\mu)} \\ \ell^\star = \pi^\star \exp\left(\dfrac{\mu - \pi^\star}{\pi^\star}\right). \end{cases}$$

Suppose, on the contrary, there exists $F \in \mathcal{F}_{\mathrm{MAD}}$ such that $\sup_{p\in\mathcal{V}} p\bar{F}(p) < \sup_{p\in\mathcal{V}} p\bar{F}^\star(p)$. Then for $v \in [\mu, 1)$, we have $v\bar{F}(v) < \pi^\star = v\bar{F}^\star(v)$, *i.e.*, $\bar{F}(v) < \bar{F}^\star(v)$. As a result, $\int_\mu^1 \bar{F}(v)\mathrm{d}v < \int_\mu^1 \bar{F}^\star(v)\mathrm{d}v = \kappa/2$, contradicting to $F \in \mathcal{F}_{\mathrm{MAD}}$.

Similarly, as for case (iii), $\kappa \ge \kappa_0$ ensures that $\ell^\star$ must lie in $(0, 1/(1-\ln(\mu))]$ and $\pi^\star/\ell^\star \in [\mu, 1]$. Then we can solve $\ell^\star$ and $\pi^\star$ via:

$$\begin{cases} \mu - \dfrac{\kappa}{2} = \displaystyle\int_0^{\mu} \bar{F}^\star(v)\mathrm{d}v = \mu\ell^\star \\ \dfrac{\kappa}{2} = \displaystyle\int_{\mu}^{1} \bar{F}^\star(v)\mathrm{d}v = \pi^\star - \mu\ell^\star + \pi^\star \ln\left(\frac{1}{\pi^\star/\ell^\star}\right) \end{cases} \implies \begin{cases} \pi^\star = \pi_{\ell^\star} \\ \ell^\star = \dfrac{2\mu - \kappa}{2\mu}. \end{cases}$$

Suppose there exists $F \in \mathcal{F}_{\mathrm{MAD}}$ such that $\sup_{p\in\mathcal{V}} p\bar{F}(p) \le \sup_{p\in\mathcal{V}} p\bar{F}^\star(p)$. Then for $v \in [\pi^\star/\ell^\star, 1)$, we have $v\bar{F}(v) \le \pi^\star = v\bar{F}^\star(v)$, *i.e.*, $\bar{F}(v) \le \bar{F}^\star(v)$. Since $\bar{F}$ is decreasing and $\int_\mu^1 \bar{F}(v)\mathrm{d}v = \int_\mu^1 \bar{F}^\star(v)\mathrm{d}v = \kappa/2$, we must have $\bar{F}(\mu) \ge \bar{F}^\star(\mu) = \ell^\star$, which also implies that $\bar{F}(v) \ge \bar{F}(\mu) \ge \ell^\star = \bar{F}^\star(v)$ for $v \in [0, \mu)$. Because $\int_0^{\mu} \bar{F}(v)\mathrm{d}v = \int_0^{\mu} \bar{F}^\star(v)\mathrm{d}v = \mu - \kappa/2$, by right continuity, it follows that $\bar{F}(v) = \ell^\star = \bar{F}^\star(v)$ for $v \in [0, \mu)$. Since $\bar{F}$ is decreasing, $\bar{F}(v) \le \bar{F}_-(\mu) = \ell^\star = \bar{F}^\star(v)$ for $v \in [\mu, \pi^\star/\ell^\star)$. Since $\int_\mu^1 \bar{F}(v)\mathrm{d}v = \int_\mu^1 \bar{F}^\star(v)\mathrm{d}v = \kappa/2$, by right continuity, we have $\bar{F}(v) = \bar{F}^\star$ for $v \in [\mu, 1)$. Thus, $F^\star$ is the unique worst-case distribution. $\square$

Unlike Theorem 5 for the mean and variance ambiguity set $\mathcal{F}_{\mathrm{V}}$, the worst-case distribution for $\mathcal{F}_{\mathrm{MAD}}$ is not necessarily unique unless the MAD is sufficiently large. Hence, in Theorem 6, when $\kappa \ge \kappa_{\mathrm{M}}$, we further distinguish two cases, and only when $\kappa \ge \kappa_0$, the distribution we identified within the family $\{Q_{1,\ell}\}_\ell$ is the unique worst-case distribution.

Theorem 6 recovers theorem 2 in Elmachtoub et al. (2020), where the authors generalize their approach for the mean ambiguity set by developing different lower bounds on the optimal value of problem (4) when κ is small, intermediate and large (*i.e.*, our cases (*i*), (*ii*) and (*iii*), respectively). They then construct a distribution $F \in \mathcal{F}_{\mathrm{MAD}}$ that achieves their lower bounds for each case. Our proof, however, uses the same geometric insight that we apply to all the other ambiguity sets to construct the worst-case distribution, and we prove its optimality by contradiction. In addition, we establish that when κ is large, the worst-case distribution is unique.

For each case in Theorem 6, we characterize the corresponding optimal mechanism below.

Proposition 10. *Given $\mathcal{F}_{\mathrm{MAD}}$, the optimal mechanism $(x^\star, t^\star)$ to problem (2) is characterized as:*

(i) If $\kappa \in [0, \kappa_{\mathrm{M}})$, then

$$(x^\star(v), t^\star(v)) = \begin{cases} (0,\ 0) & v \in [0, \pi^\star) \\ (2\gamma^\star(\ln(\pi^\star) - \ln(v)),\ 2\gamma^\star(\pi^\star - v)) & v \in [\pi^\star, \mu) \\ (1,\ 2\gamma^\star(\pi^\star - \mu)) & v \in [\mu, 1], \end{cases}$$

where $\pi^\star$ is defined in case (i) of Theorem 6 and $\gamma^\star = -1/(2\ln(\mu/\pi^\star)) < 0$.

(ii) If $\kappa \in [\kappa_{\mathrm{M}}, \kappa_0)$, then

$$(x^\star(v), t^\star(v)) = \begin{cases} (0,\ 0) & v \in [0, \mu) \\ (2\gamma^\star(\ln(v) - \ln(\mu)),\ 2\gamma^\star(v - \mu)) & v \in [\mu, 1], \end{cases}$$

where $\gamma^\star = -1/(2\ln(\mu)) > 0$.

(iii) If $\kappa \in [\kappa_0, 2\mu(1-\mu)]$, then

$$(x^\star(v), t^\star(v)) = \begin{cases} (0,\ 0) & v \in [0, k_0) \\ \left(\dfrac{2\mu\gamma^\star}{k_0}(\ln(v) - \ln(k_0)),\ \dfrac{2\mu\gamma^\star}{k_0}(v - k_0)\right) & v \in [k_0, 1], \end{cases}$$

where $\pi^\star$ is defined in case (iii) of Theorem 6, $k_0 = 2\mu\pi^\star/(2\mu - \kappa) \in [\mu, 1]$, and $\gamma^\star = -k_0/(2\mu\ln(k_0)) > 0$.

Proof of Proposition 10. Given $\mathcal{F}_{\text{MAD}}$, the inner minimization of problem (2) is

$$\begin{aligned}\inf_{F\in\mathcal{P}_+} \quad & \int_{\mathcal{V}} t(v)\,\mathrm{d}F(v)\\ \text{s.t.} \quad & \int_{\mathcal{V}} v\,\mathrm{d}F(v)=\mu\\ & \int_{\mathcal{V}} |v-\mu|\,\mathrm{d}F(v)=\kappa\\ & \int_{\mathcal{V}} \mathrm{d}F(v)=1.\end{aligned}\tag{16}$$

Combining the dual of problem (16) with the outer maximization, problem (2) becomes

$$\begin{aligned}\sup_{x,t,\alpha,\beta,\gamma} \quad & \gamma\kappa+\alpha\mu+\beta\\ \text{s.t.} \quad & t(v)\geq |v-\mu|\gamma+v\alpha+\beta \quad \forall v\in\mathcal{V}\\ & (x,t)\in\mathcal{M},\ \alpha,\beta,\gamma\in\mathbb{R}.\end{aligned}\tag{17}$$

In the remainder, for each case, we construct a feasible solution to (17) and verify that it attains an objective value equal to the optimal revenue of (4). Hence, the constructed solution is optimal.

For case (i), consider the solution $(x^\star,t^\star,\alpha^\star,\beta^\star,\gamma^\star)$ with $\alpha^\star=-\gamma^\star$ and $\beta^\star=\gamma^\star(2\pi^\star-\mu)$. Observe that $|v-\mu|\gamma^\star+v\alpha^\star+\beta^\star=(\alpha^\star-\gamma^\star)v+\beta^\star+\mu\gamma^\star=2(\pi^\star-v)\gamma^\star\leq 0=t^\star(v)$ for $v\in[0,\pi^\star)$, $|v-\mu|\gamma^\star+v\alpha^\star+\beta^\star=2(\pi^\star-v)\gamma^\star=t^\star(v)$ for $v\in[\pi^\star,\mu)$, and $|v-\mu|\gamma^\star+v\alpha^\star+\beta^\star=(\alpha^\star+\gamma^\star)v+\beta^\star-\mu\gamma^\star=-2\gamma^\star(\mu-\pi^\star)=t^\star(v)$ for $v\in[\mu,1]$. Hence, the first constraint in problem (17) holds. From $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=-2\gamma^\star$ for $v\in[\pi^\star,\mu)$, $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=0$ for $v\in[\mu,1]$ and $t^\star(v)=x^\star(v)=0$ for $v\in[0,\pi^\star)$, it follows that $t^\star(v)=vx^\star(v)-\int_0^v x^\star(u)\mathrm{d}u$. With $x^\star$ being increasing, it is now clear that $(x^\star,t^\star)\in\mathcal{M}$. Finally, this feasible solution has an objective value $\gamma^\star\kappa+\alpha^\star\mu+\beta^\star=-2\gamma^\star(\mu-\kappa/2-\pi^\star)=-2\gamma^\star\pi^\star\ln(\mu/\pi^\star)=\pi^\star$, attaining the optimal revenue of problem (4).

For case (ii), consider the solution $(x^\star,t^\star,\alpha^\star,\beta^\star,\gamma^\star)$ with $\alpha^\star=\gamma^\star$ and $\beta^\star=-\mu\gamma^\star$. Since $|v-\mu|\gamma^\star+v\alpha^\star+\beta^\star=(\alpha^\star-\gamma^\star)v+\beta^\star+\mu\gamma^\star=0=t^\star(v)$ for $v\in[0,\mu)$ and $|v-\mu|\gamma^\star+v\alpha^\star+\beta^\star=(\alpha^\star+\gamma^\star)v+\beta^\star-\mu\gamma^\star=2\gamma^\star(v-\mu)=t^\star(v)$ for $v\in[\mu,1]$, the first constraint in problem (17) holds. With $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=2\gamma^\star$ for $v\in[\mu,1]$ and $t^\star(v)=x^\star(v)=0$ for $v\in[0,\mu)$, we have $t^\star(v)=vx^\star(v)-\int_0^v x^\star(u)\mathrm{d}u$. It is now clear that $(x^\star,t^\star)\in\mathcal{M}$ because $x^\star$ is increasing. This feasible solution has an objective value $\gamma^\star\kappa+\alpha^\star\mu+\beta^\star=\gamma^\star\kappa=\pi^\star$ and achieves the optimal revenue of problem (4).

For case (iii), consider the solution $(x^\star,t^\star,\alpha^\star,\beta^\star,\gamma^\star)$ with $\alpha^\star=(k_0/(2\mu)-1)/\ln(k_0)$ and $\beta^\star=k_0/(2\ln(k_0))$. Recall from the proof of Theorem 6, when $\kappa\geq\kappa_0$, we have $k_0=\pi^\star/\ell^\star\in[\mu,1]$. Then we can observe that $|v-\mu|\gamma^\star+v\alpha^\star+\beta^\star=(\alpha^\star-\gamma^\star)v+\beta^\star+\mu\gamma^\star=(v(k_0/\mu-1))/\ln(k_0)\leq 0=t^\star(v)$ for $v\in[0,\mu)$, $|v-\mu|\gamma^\star+v\alpha^\star+\beta^\star=(\alpha^\star+\gamma^\star)v+\beta^\star-\mu\gamma^\star=(k_0-v)/\ln(k_0)\leq 0=t^\star(v)$ for $v\in[\mu,k_0)$, and $|v-\mu|\gamma^\star+v\alpha^\star+\beta^\star=(\alpha^\star+\gamma^\star)v+\beta^\star-\mu\gamma^\star=(k_0-v)/\ln(k_0)=t^\star(v)$ for $v\in[k_0,1]$. Hence, the first constraint in problem (17) holds. From $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=\alpha^\star+\gamma^\star$ for $v\in[k_0,1]$ and $t^\star(v)=x^\star(v)=0$

for $v\in[0,k_0)$, it follows that $t^\star(v)=vx^\star(v)-\int_0^v x^\star(u)\mathrm{d}u$. With $x^\star$ being increasing, it is now clear that $(x^\star,t^\star)\in\mathcal{M}$. Finally, the objective value of this feasible solution is

$$\gamma^\star\kappa+\alpha^\star\mu+\beta^\star=\frac{k_0(1-\kappa/(2\mu))-\mu}{\ln(k_0)}=\frac{\pi^\star-\mu}{\ln(k_0)}=\frac{\pi^\star-\mu}{\ln(\pi^\star/\ell^\star)}=\pi^\star,$$

attaining the optimal revenue of problem (4). $\square$

We next characterize the worst-case distribution under a posted price mechanism.

Proposition 11. [theorem 2, Roos et al. (2020)] *Given $\mathcal{F}_{\mathrm{MAD}}$, let $\underline{p}=(2\mu(1-\mu)-\kappa)/(2(1-\mu)-\kappa)\le\mu$ and $\bar{p}=(2\mu^2)/(2\mu-\kappa)\ge\mu$. Then the worst-case distribution $F_p^\star$ to the problem $\inf_{F\in\mathcal{F}_{\mathrm{MAD}}}\bar{F}(p)$ can be characterized as follows.*

(i) If $p\in(0,\underline{p})$, then

$$\bar{F}_p^\star(v)=\begin{cases}1 & v\in[0,p)\\ 1-\kappa/(2(\mu-p)) & v\in[p,\mu)\\ \kappa/(2(1-\mu)) & v\in[\mu,1)\\ 0 & v=1.\end{cases}$$

(ii) If $p\in[\underline{p},\mu)$, then

$$\bar{F}_p^\star(v)=\begin{cases}1 & v\in[0,\underline{p})\\ \kappa/(2(1-\mu)) & v\in[\underline{p},1)\\ 0 & v=1.\end{cases}$$

(iii) If $p\in[\mu,\bar{p})$, then

$$\bar{F}_p^\star(v)=\begin{cases}1-\kappa/(2\mu) & v\in[0,p)\\ (2\mu^2-2\mu p+\kappa p)/(2\mu(1-p)) & v\in[p,1)\\ 0 & v=1.\end{cases}$$

In addition, $F_p^\star$ is the unique worst-case distribution.

(iv) If $p\in[\bar{p},1]$, then

$$\bar{F}_p^\star(v)=\begin{cases}1-\kappa/(2\mu) & v\in[0,\bar{p})\\ 0 & v\in[\bar{p},1].\end{cases}$$

Proof of Proposition 11. It is not hard to verify that $F_p^\star$ in each case indeed has mean μ and MAD κ. Hence, $F_p^\star\in\mathcal{F}_{\mathrm{MAD}}$, and we focus on proving the optimality as well as uniqueness. For case (iv), since $\bar{F}_p^\star(q)=0$, it is the worst-case distribution. We focus on cases (i) to (iii) and prove by contradiction. In particular, we suppose there exists $F\in\mathcal{F}_{\mathrm{MAD}}$ such that for cases (i) and (ii), $\bar{F}(p)<\bar{F}_p^\star(p)$; while for case (iii), $\bar{F}(p)\le\bar{F}_p^\star(p)$.

For case (i), it holds that for $v\in[0,p)$, $\bar{F}(v)\le 1=\bar{F}_p^\star(v)$, and for $v\in[p,\mu)$, $\bar{F}(v)\le\bar{F}(p)<\bar{F}_p^\star(p)=\bar{F}_p^\star(v)$. These relations lead to a contradiction that $\int_0^\mu\bar{F}(v)\mathrm{d}v<\int_0^\mu\bar{F}_p^\star(v)\mathrm{d}v=\mu-\kappa/2$.

For case (ii), it holds that for any $v\in[\mu,1)$, $\bar{F}(v)\le\bar{F}(p)<\bar{F}_p^\star(p)=\bar{F}_p^\star(v)$, implying $\bar{F}_p(v)<\bar{F}_p^\star(v)$ and further yielding a contradiction that $\int_\mu^1\bar{F}(v)\mathrm{d}v<\int_\mu^1\bar{F}_p^\star(v)\mathrm{d}v=\kappa/2$.

For case (*iii*), by $\mu\bar{F}(\mu) \le \int_0^\mu \bar{F}(v)\mathrm{d}v = \int_0^\mu \bar{F}_p^\star(v)\mathrm{d}v = \mu\bar{F}_p^\star(\mu)$ we have $\bar{F}(\mu) \le \bar{F}_p^\star(\mu)$, which implies $\bar{F}(v) \le \bar{F}(\mu) \le \bar{F}_p^\star(\mu) = \bar{F}_p^\star(v)$ for $v \in [\mu, p)$. In addition, for $v \in [p,1)$ we have $\bar{F}(v) \le \bar{F}(p) \le \bar{F}_p^\star(p) = \bar{F}_p^\star(v)$, *i.e.*, $\bar{F}(v) \le \bar{F}_p^\star(v)$. Now we obtain $\kappa/2 = \int_\mu^1 \bar{F}(v)\mathrm{d}v \le \int_\mu^1 \bar{F}_p^\star(v)\mathrm{d}v = \kappa/2$. By right continuity, it holds that $\bar{F}(v) = \bar{F}_p^\star(v)$ for $v \in [\mu, 1)$. With $\bar{F}(\mu) = \bar{F}_p^\star(\mu)$, for $v \in [0,\mu)$ we have $\bar{F}(v) \ge \bar{F}(\mu) = \bar{F}_p^\star(\mu) = \bar{F}_p^\star(v)$. Since $\int_0^\mu \bar{F}(v)\mathrm{d}v = \mu - \kappa/2 = \int_0^\mu \bar{F}_p^\star(v)\mathrm{d}v$, we obtain $F = F_p^\star$ by right continuity. In conclusion, $F_p^\star$ is the unique worst-case distribution. □

To solve $\inf_{F\in\mathcal{F}_{\mathrm{MAD}}} \bar{F}(p)$, Roos et al. (2020) consider its primal and dual reformulations

$$\begin{aligned}
\inf_{F\in\mathcal{P}_+} \quad & \int_{\mathcal{V}} \mathbb{I}(p>v)\,\mathrm{d}F(v) \\
\text{s.t.} \quad & \int_{\mathcal{V}} v\,\mathrm{d}F(v) = \mu \\
& \int_{\mathcal{V}} |v-\mu|\,\mathrm{d}F(v) = \kappa \\
& \int_{\mathcal{V}} \mathrm{d}F(v) = 1
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
\sup_{\alpha,\beta,\gamma} \quad & \gamma\kappa + \alpha\mu + \beta \\
\text{s.t.} \quad & \mathbb{I}(p>v) \ge |v-\mu|\gamma + v\alpha + \beta \quad \forall v \in \mathcal{V} \\
& \alpha, \beta, \gamma \in \mathbb{R}
\end{aligned}$$

and construct a pair of primal and dual optimal solutions. Their proof spans more than two pages. In contrast, our construction of the worst-case distribution is based on the same geometric insight used for all other ambiguity sets, and our proof also follows the simple contradiction arguments used in all of our other results on either minimax pricing or maximin pricing problems.

C.3 Mean-Preserving Spread Ambiguity Set

The mean-preserving spread ambiguity set is defined by all mean-preserving spreads of a reference distribution G:

$$\mathcal{F}_{\mathrm{MPS}} = \left\{ F \in \mathcal{P} \;\middle|\; \int_{v\in\mathcal{V}} v\,\mathrm{d}F(v) = \int_{v\in\mathcal{V}} v\,\mathrm{d}G(v) = \mu,\ \int_0^u \bar{F}(v)\,\mathrm{d}v \le \int_0^u \bar{G}(v)\,\mathrm{d}v\ \forall u \in \mathcal{V} \right\}.$$

It is well known that F is a mean-preserving spread (MPS) of G if and only if there are random variables $\tilde{v} \sim F$, $\tilde{v}' \sim G$ and $\tilde{\varepsilon} \sim Q$ such that $\tilde{v} = \tilde{v}' + \tilde{\varepsilon}$ with $\mathbb{E}_Q[\tilde{\varepsilon}|v'] = 0$ for all possible realization v' of $\tilde{v}'$ (see, *e.g.*, Rothschild and Stiglitz 1970). Intuitively, the mean-preserving spread F of G can be obtained by first drawing a realization from G and then adding a zero-mean noise. Hence, $\mathcal{F}_{\mathrm{MPS}}$ can be used to model the scenario when consumer's valuation consists of $\tilde{v}' \sim G$, whose distribution can be accurately estimated by the seller, and an estimation error $\tilde{\varepsilon}$ whose distribution is unknown.

If the worst-case distribution for the mean ambiguity set is already an MPS of G, then it is also the worst-case distribution for $\mathcal{F}_{\mathrm{MPS}}$. Otherwise, however, then one can move mass from smaller valuations to larger ones along the level curves; see Figure 11. Recall the family of distributions $\{Q_{1,\ell}\}_\ell$ in Appendix C.1. Note that for any $\ell_1, \ell_2 \in [\mu, 1]$ such that $\ell_1 \ge \ell_2$, Q_{1,ℓ_2} is an MPS of Q_{1,ℓ_1} with $\bar{Q}_{1,1}$ being the least spreading distribution and $\bar{Q}_{1,\mu}$—the Bernoulli distribution—being the

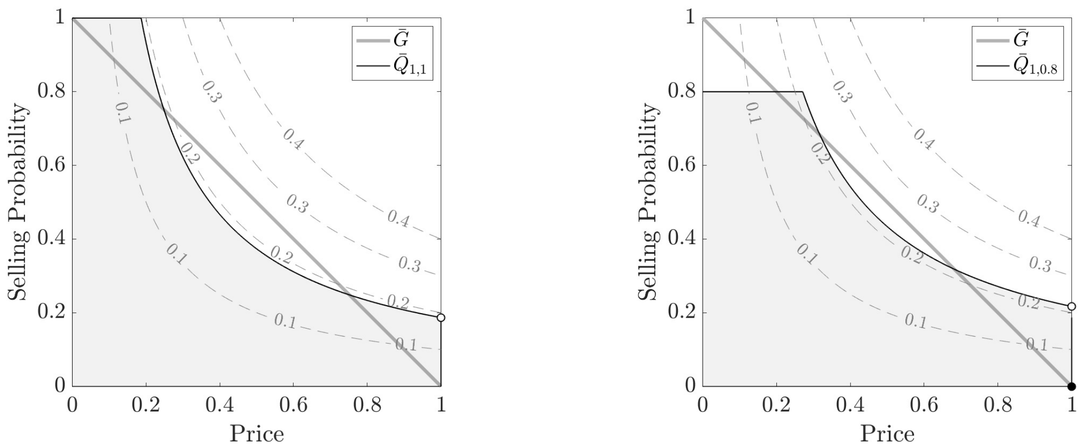

Figure 11 The worst-case distribution $\bar{Q}_{1,1}$ for $\mathcal{F}_{\mathrm{M}}$ is not in $\mathcal{F}_{\mathrm{MPS}}$. Here, $\bar{Q}_{1,0.8}$, a mean-preserving spread of $\bar{Q}_{1,1}$, is formed by moving mass from smaller valuations to larger valuations along the level curves.

most spreading one. Since the seller's optimal revenue under $\bar{Q}_{1,\ell}$ is decreasing in ℓ, the worst-case revenue within this family can then be found by the largest ℓ such that $Q_{1,\ell}$ is an MPS of G. The next theorem shows that such a distribution is also the worst-case distribution for $\mathcal{F}_{\mathrm{MPS}}$.

Theorem 7. *Given $\mathcal{F}_{\mathrm{MPS}}$, the worst-case distribution to problem (4) is $F^\star = Q_{1,\ell^\star}$ with the parameter $\ell^\star = \max\{\ell \in [\mu, 1] \mid Q_{1,\ell} \in \mathcal{F}_{\mathrm{MPS}}\}$. The optimal revenue of problem (4) is $\pi^\star = \pi_{\ell^\star}$.*

Proof of Theorem 7. Suppose, on the contrary, there exists $F \in \mathcal{F}_{\mathrm{MPS}}$ such that $\pi = \sup_{p\in\mathcal{V}} p\bar{F}(p) < \max_{p\in\mathcal{V}} p\bar{Q}_{1,\ell}(p) = \pi^\star$. Since $\ell^\star = \pi^\star \exp((\mu - \pi^\star)/\pi^\star)$ is strictly decreasing in $\pi^\star$, then $\pi^\star > \pi$ implies $\ell^\star < \ell := \pi \exp((\mu-\pi)/\pi)$. We next show that $Q_{1,\ell} \in \mathcal{F}_{\mathrm{MPS}}$, which yields a contradiction to the definition of $\ell^\star$. To this end, we first observe that for $v \in [\pi/\ell, 1]$, $v\bar{F}(v) \le \sup_{p\in\mathcal{V}} p\bar{F}(p) = \pi = \max_{p\in\mathcal{V}} p\bar{Q}_{1,\ell}(p) = v\bar{Q}_{1,\ell}(v)$, implying $\bar{F}(v) \le \bar{Q}_{1,\ell}(v)$. Let $v' = \sup\{v \in [0, \pi/\ell) \mid \bar{F}(v) \ge \ell\}$, which must exist because otherwise, we would have $\int_{\mathcal{V}} \bar{F}(v)\mathrm{d}v < \int_{\mathcal{V}} \bar{Q}_{1,\ell}(v)\mathrm{d}v = \mu$. Since $\bar{F}$ is decreasing, we have $\bar{F}(v) \ge \bar{Q}_{1,\ell}(v) = \ell$ for $v \in [0, v')$ and $\bar{F}(v) \le \bar{Q}_{1,\ell}(v)$ for $v \in (v', \pi/\ell)$. These relations imply $\int_0^u \bar{Q}_{1,\ell}(v)\mathrm{d}v \le \int_0^u \bar{F}(v)\mathrm{d}v$ for $u \in [0, v')$ and $\int_0^u \bar{Q}_{1,\ell}(v)\mathrm{d}v = \mu - \int_u^1 \bar{Q}_{1,\ell}(v)\mathrm{d}v \le \mu - \int_u^1 \bar{F}(v)\mathrm{d}v = \int_0^u \bar{F}(v)\mathrm{d}v$ for $u \in [v', 1]$. That is, $Q_{1,\ell}$ is an MPS of F, and hence an MPS of G as well, concluding $Q_{1,\ell} \in \mathcal{F}_{\mathrm{MPS}}$—a contradiction. With $Q_{1,\ell^\star}$ being the worst-case distribution, the corresponding optimal revenue is clearly $\pi^\star = \max_{q\in\mathcal{V}} p\bar{Q}_{1,\ell^\star}(p) = \pi_{\ell^\star}$. □

We characterize the corresponding solution to the primal problem next.

Proposition 12. *Given $\mathcal{F}_{\mathrm{MPS}}$ and $\ell^\star$ in Theorem 7, when $\ell^\star = 1$, the optimal mechanism under the mean ambiguity set $\mathcal{F}_{\mathrm{M}}$ (see Proposition 2) is also optimal for $\mathcal{F}_{\mathrm{MPS}}$; when $\ell^\star < 1$, we have:*

(i) *if there exists $u^\star \in [\pi^\star/\ell^\star, 1)$ such that $\int_0^{u^\star} (\bar{F}^\star(v) - \bar{G}(v))\mathrm{d}v = 0$, then*

$$(x^\star(v), t^\star(v)) = \begin{cases} (0,0) & v \in [0, u^\star) \\ (\gamma^\star \ln(v/u^\star), \gamma^\star(v-u^\star) & v \in [u^\star, 1] \end{cases} \quad \text{with} \quad \gamma^\star = -\frac{1}{\ln(u^\star)} > 0.$$

(ii) otherwise, a posted price mechanism with price $q=1$ is optimal.

Proof of Proposition 12. Given $\mathcal{F}_{\text{MPS}}$, the inner minimization of problem (2) is

$$\begin{aligned} \inf_{F\in\mathcal{P}_+} \quad & \int_{\mathcal{V}} t(v)\,\mathrm{d}F(v) \\ \text{s.t.} \quad & \int_{\mathcal{V}} v\,\mathrm{d}F(v)=\mu \\ & \int_{\mathcal{V}}(u-v)^+\,\mathrm{d}F(v)\geq \int_{\mathcal{V}}(u-v)^+\,\mathrm{d}G(v) \qquad \forall u\in\mathcal{V} \\ & \int_{\mathcal{V}} \mathrm{d}F(v)=1. \end{aligned} \tag{18}$$

Combining the dual of problem (18) with the outer maximization, problem (2) becomes

$$\begin{aligned} \sup_{x,t,\alpha,\beta,\Gamma} \quad & \mu\alpha+\beta+\int_{\mathcal{V}}\phi(u)\,\mathrm{d}\Gamma(u) \\ \text{s.t.} \quad & t(v)\geq v\alpha+\beta+\int_{\mathcal{V}}(u-v)^+\,\mathrm{d}\Gamma(u) \quad \forall v\in\mathcal{V} \\ & (x,t)\in\mathcal{M},\ \alpha\in\mathbb{R},\ \beta\in\mathbb{R},\ \Gamma\in\mathcal{P}_+, \end{aligned} \tag{19}$$

where for any $u\in\mathcal{V}$ we put $\phi(u)=\int_{\mathcal{V}}(u-v)^+\mathrm{d}G(v)$.

If $\ell^\star=1$, then $F^\star$ is exactly the worst-case distribution for the mean ambiguity set in Theorem 1. This implies that given $\mathcal{F}_{\text{MPS}}$, the worst-case revenue $\pi^\star$ of problem (4) (*i.e.*, the optimal objective value of problem (19)) coincides with that of the mean ambiguity set in Proposition 2. It is not hard to see that the optimal mechanism $(x^\star,t^\star)$ in Proposition 2, together with $\alpha^\star=\pi^\star/(\mu-\pi^\star)$, $\beta^\star=-\pi^\star\alpha^\star$ and $\Gamma^\star(v)=0$ for all $v\in\mathcal{V}$, constitutes a feasible solution to problem (19) that attains the worst-case revenue $\pi^\star$. Hence, $(x^\star,t^\star)$ in Proposition 2 remains optimal.

We next focus on $\ell^\star<1$. In view of case (i), let $\alpha^\star=\gamma^\star$, $\beta^\star=-\gamma^\star u^\star$ and $\Gamma^\star$ be a step function such that $\Gamma^\star(v)=0$ for $v\in[0,u^\star)$; $\Gamma^\star(v)=\gamma^\star$ for $v\in[u^\star,1]$. We first check that $(x^\star,t^\star,\alpha^\star,\beta^\star,\Gamma^\star)$ is feasible to problem (19). Since $v\alpha^\star+\beta^\star+\int_{\mathcal{V}}(u-v)^+\,\mathrm{d}\Gamma^\star(u)=\gamma^\star v-\gamma^\star u^\star+(u^\star-v)^+\gamma^\star=t^\star(v)$ for all $v\in[0,1]$, the first constraint of problem (19) is satisfied. Since $t^\star(v)=x^\star(v)=0$ for $v\in[0,u^\star)$ and $\frac{\mathrm{d}t^\star(v)}{\mathrm{d}v}=v\frac{\mathrm{d}x^\star(v)}{\mathrm{d}v}=\gamma^\star$ for $v\in[u^\star,1]$, it follows that $t^\star(v)=vx^\star(v)-\int_0^v x^\star(u)\mathrm{d}u$. Finally, as $x^\star(v)$ is increasing in v, we can conclude $(x^\star,t^\star)\in\mathcal{M}$. It remains to verify that the objective value of $(x^\star,t^\star,\alpha^\star,\beta^\star,\Gamma^\star)$, given by $\mu\alpha^\star+\beta^\star+\int_{\mathcal{V}}\phi(u)\mathrm{d}\Gamma^\star(u)=\gamma^\star(\mu-u^\star)+\gamma^\star\int_0^{u^\star}G(u)\mathrm{d}u$, is equal to $\pi^\star$. To this end, observe that $\int_0^{u^\star}G(u)\mathrm{d}u=u^\star-\int_0^{u^\star}\bar{G}(u)\mathrm{d}u=u^\star-\int_0^{u^\star}\bar{F}^\star(u)\mathrm{d}u=u^\star-\mu+\pi^\star\ln(1/u^\star)$, which leads to $\gamma^\star(\mu-u^\star)+\gamma^\star\int_0^{u^\star}G(u)\mathrm{d}u=\gamma^\star\pi^\star\ln(1/u^\star)=\pi^\star$—the desired result.

As for case (ii), since the posted price mechanism $(x^\star,p^\star)$ is clearly feasible, we focus on optimality.

Our first step is to show $\int_0^u(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v<0$ for all $u\in(0,1)$. Since there does not exist $u^\star$ as defined in case (i), we have $\int_0^u(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v<0$ for $u\in[\pi^\star/\ell^\star,1)$. Suppose there exists

$u_0 \in (0,\pi^\star/\ell^\star)$ such that $\int_0^{u_0}(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v = 0$. We must have $\bar{G}(u_0) \geq \bar{F}^\star(u_0) = \ell^\star$; otherwise, $\bar{G}(u_0) < \bar{F}^\star(u_0) = \ell^\star$ and $\int_0^{u}(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v > 0$ for $u \in (u_0,\pi^\star/\ell^\star)$. Since $\bar{G}$ is decreasing, $\bar{G}(v) \geq \bar{G}(u_0) = \ell^\star = \bar{F}^\star(v)$ for $v \in (0,u_0)$. By right-continuity and $\int_0^{u_0}(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v = 0$, we have $\bar{G}(v) = \bar{F}^\star(v)$ for $v \in (0,u_0)$. Because $\bar{G}$ is decreasing, we now have $\bar{G}(v) \leq \ell^\star = \bar{F}^\star(v)$ for $v \in (0,\pi^\star/\ell^\star)$, implying $\int_0^{\pi^\star/\ell^\star}(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v \geq 0$. This, together with $\int_0^{\pi^\star/\ell^\star}(\bar{G}(v)-\bar{F}^\star(v))\mathrm{d}v \geq 0$, yields $\int_0^{\pi^\star/\ell^\star}(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v = 0$, contradicting to the assumption that there does not exist $u^\star$ as defined in case (i). Therefore, $\int_0^{u}(\bar{F}^\star(v)-\bar{G}(v))\mathrm{d}v < 0$ for $u \in (0,\pi^\star/\ell^\star)$.

Our second step is to show $\pi^\star = \bar{G}_-(1)$. Note that $F^\star$ cannot be an MPS of G if $\pi^\star < \bar{G}_-(1)$. Suppose $\pi^\star > \bar{G}_-(1)$. Then there exists $\delta \in (0, 1-\pi^\star/\ell^\star)$ such that $\bar{G}(v) < (\pi^\star + \bar{G}_-(1))/2 < \pi^\star$ for all $v \in [1-\delta, 1]$. Consider a small ε_1 such that for $\ell_1 = \ell^\star + \varepsilon_1$, $\pi^\star - \pi_{\ell_1} \in (0, (\pi^\star - \bar{G}_-(1))/2)$. Then for $\ell \in [\ell^\star, \ell_1]$ and $v \in [1-\delta, 1]$, we have $\bar{Q}_{1,\ell}(v) \geq \pi_{\ell_1} > (\pi^\star + \bar{G}_-(1))/2 > \bar{G}(v)$, which further implies $\int_0^u \bar{Q}_{1,\ell}(v)\mathrm{d}v < \int_0^u \bar{G}(v)\mathrm{d}v$ for $u \in [1-\delta, 1]$. In addition, there must exist $\delta' \in (0, \pi^\star/\ell^\star)$ such that $\bar{G}(v) > \bar{F}^\star(0) = \ell^\star$ for $v \in [0, \delta']$. Otherwise, for any $v \in (0, \pi^\star/\ell^\star)$ we must have $\bar{G}(v) \leq \bar{F}(0) = \ell^\star$, which implies $\int_0^u(\bar{F}^\star(v) - \bar{G}(v))\mathrm{d}v \geq 0$ for $u \in [0, \pi^\star/\ell^\star]$. This, together with the constraint that $\int_0^u \bar{F}^\star(v)\mathrm{d}v \leq \int_0^u \bar{G}(v)\mathrm{d}v$ for $u \in [0, \pi^\star/\ell^\star]$, implies $\int_0^{\pi^\star/\ell^\star} \bar{F}^\star(v)\mathrm{d}v = \int_0^{\pi^\star/\ell^\star} \bar{G}(v)\mathrm{d}v$, a contradiction to the nonexistence of $u^\star$. Let $\varepsilon_2 = (\bar{G}(\delta') - \ell^\star)/2$ and $\ell_2 = \ell^\star + \varepsilon_2$. For any $\ell \in [\ell^\star, \ell_2]$ and any $v \in [0, \delta']$, it holds that $\bar{Q}_{1,\ell}(v) \leq \ell_2 < \bar{G}(\delta') \leq \bar{G}(v)$. This yields $\int_0^u \bar{Q}_{1,\ell}(v)\mathrm{d}v < \int_0^u \bar{G}(v)\mathrm{d}v$ for $\ell \in [\ell^\star, \ell_2]$ and $u \in [0, \delta']$. Let $\Delta := \min_{u\in[\delta',1-\delta]} \int_0^u (\bar{G}(v) - \bar{F}^\star(v)\mathrm{d}v$ and choose a small $\varepsilon_3 > 0$ such that $\max_{u\in[0,1]} \int_0^u (\bar{Q}_{1,\ell_3}(v) - \bar{F}^\star(v))\mathrm{d}v < \Delta/2$ holds for $\ell_3 = \ell^\star + \varepsilon_3$. Then $\int_0^u \bar{Q}_{1,\ell}(v)\mathrm{d}v < \int_0^u \bar{G}(v)\mathrm{d}v$ for $\ell \in [\ell^\star, \ell_3]$ and $u \in [\delta', 1-\delta]$. Let $\ell' = \min\{\ell_1, \ell_2, \ell_3\} > \ell^\star$, we now have $Q_{1,\ell'} \in \mathcal{F}_{\mathrm{MPS}}$, contradicting to the definition of $\ell^\star$. Thus, $\pi^\star = \bar{G}_-(1)$.

In the last step, we first note that the revenue of $(x^\star, t^\star)$ under F is $\bar{F}_-(1)$. Consider any $F \in \mathcal{F}_{\mathrm{MPS}}$ such that $\bar{F}_-(1) < \bar{G}_-(1)$, then there exists $\delta_F > 0$ such that $\bar{F}(v) < \bar{G}_-(1) \leq \bar{G}(v)$ for $v \in (1-\delta_F, 1)$. This implies that for all $u \in [1-\delta_F, 1]$, $\int_0^u \bar{F}(v)\mathrm{d}v > \int_0^u \bar{G}(v)\mathrm{d}v$, contradicting to $F \in \mathcal{F}_{\mathrm{MPS}}$. Hence, for all $F \in \mathcal{F}_{\mathrm{MPS}}$, $\bar{F}_-(1) \geq \bar{G}_-(1) = \pi^\star$, implying that the worst-case revenue of $(x^\star, t^\star)$ under $\mathcal{F}_{\mathrm{MPS}}$ is no smaller than $\pi^\star$. Because $\pi^\star$ is attained under G, $(x^\star, t^\star)$ is indeed optimal. $\square$

To characterize the worst-case distribution under a posted price mechanism, we can consider a family of distributions

$$\bar{W}_{p,\ell}(v) = \begin{cases} \ell & v \in [0,p) \\ (\mu - p\ell)/(1-p) & v \in [p,1) \\ 0 & v = 1, \end{cases}$$

with $\ell \in [\mu, \mu/p]$. The characterization is summarized as follows.

PROPOSITION 13. *Given $\mathcal{F}_{\mathrm{MPS}}$, for any $p \in (0,1]$, a worst-case distribution $F_p^\star$ to the problem $\inf_{F\in\mathcal{F}_{\mathrm{MPS}}} \bar{F}(p)$ is $\bar{W}_{p,\ell_p^\star}$ with $\ell_p^\star = \max\{\ell \in [\mu,\mu/p] \mid \bar{W}_{p,\ell} \in \mathcal{F}_{\mathrm{MPS}}\}$. Correspondingly, $\bar{\Lambda}^\star(p) = (\mu - p\ell_p^\star)/(1-p)$ for $p \in (0,1]$.*

Proof of Proposition 13. If $\bar{F}_p^\star(p)=0$, then $F_p^\star$ is clearly a worst-case distribution. If $\bar{F}_p^\star(p)>0$, suppose there exists $F\in\mathcal{F}_{\text{MPS}}$ such that $\bar{F}(p)<\bar{F}_p^\star(p)$. Let $\ell:=(\mu-(1-p)\bar{F}(p))/p$ such that $\bar{W}_{p,\ell}(p)=\bar{F}(p)$ and $\ell>(\mu-(1-p)\bar{F}_p^\star(p))/p=\ell_p^\star$. For $v\in[p,1)$, it holds that $\bar{F}(v)\le\bar{F}(p)=\bar{W}_{p,\ell}(p)=\bar{W}_{p,\ell}(v)$. Let $v'=\sup\{v\in[0,p)\mid \bar{F}(v)\ge\ell\}$, which must exist because otherwise, we would have $\int_{\mathcal{V}}\bar{F}(v)\mathrm{d}v<\int_{\mathcal{V}}\bar{W}_{p,\ell}(v)\mathrm{d}v=\mu$. Since $\bar{F}$ is decreasing, $\bar{F}(v)\ge\ell=\bar{W}_{p,\ell}(v)$ for $v\in[0,v')$. Thus, the continuous function $h_p(u):=\int_0^u(\bar{F}(v)-\bar{W}_{p,\ell}(v))\mathrm{d}v$ is increasing on $[0,v')$ while decreasing on $[v',1]$. Because $h_p(0)=h_p(1)=0$, it follows that $h_p(u)\ge 0$ for $u\in[0,1]$, implying $W_{p,\ell}$ is an MPS of F and $W_{p,\ell}\in\mathcal{F}_{\text{MPS}}$. This contradicts to the definition of $\ell_p^\star$. □

The mean-preserving spread ambiguity set has been considered in Condorelli and Szentes (2020) for buyer-optimal pricing problem (5), which is again equivalent to our minimax pricing problem by Theorem 3. Our results for the robust mechanism and maximin pricing problems under the mean-preserving spread ambiguity set, to the best of our knowledge, is new.

C.4 Comparison of Dual Formulations

Here, we compare the dual formulations based on different primal formulations in the special case when the ambiguity set is a singleton, *i.e.*, $\mathcal{F}=\{F\}$. A common route taken in the literature (see, *e.g.*, Vohra 2011, Carroll 2017, Cai et al. 2019) is to directly take the dual of the following (original) primal formulation:

$$
\begin{aligned}
\sup_{(x,t)\in\mathcal{M}} \mathbb{E}_F[t(\tilde{v})] = \sup_{x,t} \quad & \int_{\mathcal{V}} t(v)\,\mathrm{d}F(v) \\
\text{s.t.} \quad & (vx(u)-t(u))-(vx(v)-t(v))\le 0 \quad && \forall u,v\in\mathcal{V} \quad \cdots\ \lambda(v,u)\\
& -(vx(v)-t(v))\le 0 && \forall v\in\mathcal{V} \quad \cdots\ \kappa(v)\\
& x(v)\le 1 && \forall v\in\mathcal{V} \quad \cdots\ \nu(v)\\
& x(v)\ge 0 && \forall v\in\mathcal{V},
\end{aligned}
$$

whose dual is written as

$$
\begin{aligned}
\inf_{\lambda,\kappa,\nu} \quad & \int_{\mathcal{V}} \nu(v)\mathrm{d}v \\
\text{s.t.} \quad & \nu(v)-v\kappa(v)+\int_{\mathcal{V}}(u\lambda(u,v)-v\lambda(v,u))\,\mathrm{d}u\ge 0 \quad && \forall v\in\mathcal{V} \quad \cdots\ x(v)\\
& \mathrm{d}F(v)=\left(\kappa(v)+\int_{\mathcal{V}}(\lambda(v,u)-\lambda(u,v))\,\mathrm{d}u\right)\mathrm{d}v && \forall v\in\mathcal{V} \quad \cdots\ t(v)\\
& \lambda(v,u)\ge 0 && \forall u,v\in\mathcal{V}\\
& \kappa(v)\ge 0,\ \nu(v)\ge 0 && \forall v\in\mathcal{V}.
\end{aligned}
$$

The dual problem here can be interpreted as a network flow problem with each valuation v being a node and the dual variable $\lambda(v,u)$ representing the flow from node v to node u. Carroll (2017) constructs an optimal solution to the above problem and shows that it matches the optimal value of the primal (see Appendix C of Carroll 2017). However, it is far from clear that the above dual formulation is equivalent to a pricing problem.

To obtain Lemma 1, our approach exploits the form in (1) that states $t(v) = vx(v) - \int_0^v x(u)\mathrm{d}u = \int_0^v u\mathrm{d}x(u)$ and the fact that x can be interpreted as a distribution on $\mathcal{V}$. It follows that

$$\begin{aligned} \sup_{(x,t)\in\mathcal{M}} \mathbb{E}_F[t(\tilde{v})] \;=\; \sup_{x\in\mathcal{P}_+} \quad & \int_{\mathcal{V}} \int_0^v u\,\mathrm{d}x(u)\,\mathrm{d}F(v) \\ \text{s.t.} \quad & \int_{\mathcal{V}} \mathrm{d}x(v) = 1 \qquad\qquad \cdots\ \pi. \end{aligned}$$

Taking the dual of the above problem, we obtain

$$\begin{aligned} \inf_{\pi} \quad & \pi \\ \text{s.t.} \quad & \pi \geq p\int_p^1 \mathrm{d}F(u) \quad \forall p\in\mathcal{V} \end{aligned} \qquad \Longleftrightarrow \qquad \sup_{p\in\mathcal{V}} p\bar{F}(p),$$

which proves Lemma 1 for the case of $\mathcal{F} = \{F\}$.

C.5 Representation of Wasserstein Ambiguity Sets

For $1 \leq \rho < \infty$, the type-ρ Wasserstein metric between two distributions, F and G, is defined as

$$d_\rho(F,G) \;=\; \inf_{Q\in\mathcal{Q}(F,G,\mathcal{U})} (\mathbb{E}_Q[\|\tilde{\boldsymbol{v}} - \tilde{\boldsymbol{u}}\|^\rho])^{1/\rho} \;=\; \left(\inf_{Q\in\mathcal{Q}(F,G,\mathcal{U})} \mathbb{E}_Q[\|\tilde{\boldsymbol{v}} - \tilde{\boldsymbol{u}}\|^\rho]\right)^{1/\rho},$$

where $\tilde{\boldsymbol{v}} \sim F$, $\tilde{\boldsymbol{u}} \sim G$, and $\mathcal{Q}(F,G,\mathcal{U})$ is the set of all joint distributions supported on $\mathcal{U}\times\mathcal{U}$ with marginals F and G. The type-∞ Wasserstein metric is the limiting case as ρ approaches infinity, that is, $d_\infty(F,G) = \lim_{\rho\to\infty} d_\rho(F,G)$. A type-$\rho$ Wasserstein ambiguity set $\mathcal{F}_{(\theta,\rho)}$ is then defined as a ball of radius $\theta \geq 0$ with respect to the corresponding Wasserstein metric, centered at a prescribed reference distribution G:

$$\mathcal{F}_{(\theta,\rho)} = \{F \in \mathcal{P}(\mathcal{U}) \mid d_\rho(F,G) \leq \theta\},$$

where $\mathcal{P}(\mathcal{U})$ is the set of probability distributions supported on $\mathcal{U}$. Wasserstein metric between single-dimensional distributions can be expressed in terms of their quantile functions. In particular, for $\rho \in [1,\infty)$ and $F, G \in \mathcal{V}$, it holds that

$$d_\rho(F,G) = \left(\inf_{Q\in\mathcal{Q}(F,G,\mathcal{V})} \mathbb{E}_Q[\|\tilde{v} - \tilde{u}\|^\rho]\right)^{1/\rho} = \left(\int_0^1 \left|F^{-1}(q) - G^{-1}(q)\right|^\rho \mathrm{d}q\right)^{1/\rho} = \|\bar{F}^{-1} - \bar{G}^{-1}\|_\rho.$$

Here, the second identity follows from taking $h(\cdot) = |\cdot|^\rho$ in proposition 2.17 of Santambrogio (2015). With ρ approaching infinity, it holds that

$$d_\infty(F,G) = \lim_{\rho\to\infty} d_\rho(F,G) = \lim_{\rho\to\infty} \|\bar{F}^{-1} - \bar{G}^{-1}\|_\rho = \|\bar{F}^{-1} - \bar{G}^{-1}\|_\infty = \sup_{q\in[0,1]} |\bar{F}^{-1}(q) - \bar{G}^{-1}(q)|.$$

It now follows that for any $\rho \in [1,\infty]$, a Wasserstein ambiguity set of the corresponding type in the content of our paper can be represented as

$$\mathcal{F}_{(\theta,\rho)} = \left\{F \in \mathcal{P} \;\middle|\; \|\bar{F}^{-1} - \bar{G}^{-1}\|_\rho \leq \theta\right\}.$$