A New Method to Solve Zero-Sum Games under Moment Conditions

Zhaolin Li
The University of Sydney Business School, Sydney, NSW2006, Australia, erick.li@sydney.edu.au

Qi Fu
The University of Macau, Taipa, Macau, China, gracefu@um.edu.mo

Chung-Piaw Teo
NUS Business School, National University of Singapore, Singapore 119245, bizteocp@nus.edu.sg

When only the moments (mean, variance or t-th moment) of the underling distribution are known, numerous max-min optimization models can be interpreted as a zero-sum game, in which the firm chooses actions to maximize her expected profit while Adverse Nature chooses a distribution subject to the moment conditions to minimize the firm's expected profit. We propose a new method to efficiently solve this class of zero-sum games under moment conditions. By applying the min-max inequality, our method reformulates the zero-sum game as a robust moral hazard model, in which Adverse Nature chooses both the distribution and actions to minimize the firm's expected profit subject to incentive compatibility (IC) constraints. Under quasi-concavity, these IC constraints are replaced by the first-order conditions, which give rise to additional moment constraints. We show that in equilibrium, these moment constraints are binding but have zero Lagrangian multipliers and thus enable us to derive closed-form solutions for several distributionally robust inventory models with different levels of complexity, including the case with mean and t-th moment, the case with multiple supply sources, and the case with component commonality.

Key words: Robust optimization, zero-sum games, moment-based ambiguity, inventory management

1. Introduction

Various practical hurdles could result in ambiguity by preventing decision makers from precisely knowing the underling distribution that governs the random environment. For example, forecasting demand for new products is notoriously difficult (see Chapter 5 in Lilien et al. 2017, page 135 to 168). Small and medium-sized enterprises usually lack the resources needed to collect and analyze data.[1] Environmental factors such as a trade embargo or flood also increase the level of ambiguity experienced by many supply chains. Therefore, ambiguity (which differs from risk) has received increasing research attention from various disciplines.

Despite ambiguity, many practitioners in supply chains are still able to obtain the mean and variance (or higher moments) of the random variable that they are interested in (e.g., Natarajan et al. 2018, Li and Kirshner 2021). To cope with ambiguity, the robust max-min decision rule is the most popular decision rule in the literature of robust operations management (see page 1931 in Lu and Shen 2021). While the practical reason is that the worst case is of special interest to risk managers (Tang 2006), the technical reason is that an ambiguity-based model can be transformed into a semi-infinite programming (SIP) model using duality, especially when the moments (such as the mean, variance, or t-th moment) are known. The practical relevance and analytical convenience make this class of max-min models under moment conditions prevalent in the literature (see Section 2.1.2 of Lu and Shen 2021, and the references therein).

For ease of exposition, hereafter we use the feminine pronoun for the firm. In a typical max-min optimization problem with known moments, the firm chooses her actions to maximize her worst-case expected profit based on the available information. In supply chains, managers often make pricing and production joint decisions, prompting us to use a vector to represent the firm's actions. It is also well-known that this max-min problem formulates a zero-sum game, in which the firm chooses an action while Adverse Nature chooses the realized value subject to the constraints on moments. In other words, the firm employs a pure strategy, while Adverse Nature employs a mixed strategy that notably translates into a distribution satisfying the moment conditions. We refer to this class of models as the zero-sum games under moment conditions.

The traditional method to solve this class of zero-sum games is based on the firm's perspective and proceeds in two stages. We first compute the firm's worst case expected profit for any given actions and then optimize the actions. When the actions are fixed, the second-stage problem is reformulated as an SIP model using duality. We label this model as the firm's SIP model, in which the second stage SIP model is linear and hence, Karush–Kuhn–Tucker (KKT) conditions are sufficient and necessary. However, the action-dependent SIP model in the second stage often produces a piece-wise or even implicit objective function, rendering the first-stage problem cumbersome or intractable. In Section 2, we thoroughly explain the significant challenge that the traditional method encounters.

We propose a new and efficient method to solve the firm's equilibrium strategy (which is also known as her robust optimal solution). The key difference is that the traditional method is based on the firm's perspective while our method is based on Adverse Nature's perspective. The optimization model of Adverse Nature becomes a min-max problem, in which Adverse Nature chooses a distribution by anticipating that the firm will choose actions that are optimal for the chosen distribution. Conceptually, the min-max problem of Adverse Nature is a robust moral hazard model where Adverse Nature is the principal and the firm is the agent. Adverse Nature jointly chooses the distribution and actions to minimize the firm's expected profit subject to the moment constraints and the firm's incentive compatibility (IC) constraints. Moral hazard interpretation is crucial because it not only enables us to apply the tools in the economics literature to simplify IC constraints but also broadens our application to moral hazard. With mean-variance (or t-th moment) ambiguity, the IC constraints become additional moment constraints that can be conveniently incorporated into Adverse Nature's SIP model.

These additional moment constraints produce two noteworthy consequences. First, the additional moment constraints effectively integrate the firm's first-order conditions (FOCs) into the min-max version of the zero-sum game, enabling us to obtain closed-form solutions without explicitly deriving the firm's objective function. Notably, we bypass the second-stage problem where the traditional method encounters significant challenges. Second, *zero* Lagrangian multipliers facilitate the characterization of the equilibrium. We assume that the zero-game has an equilibrium (otherwise, the original max-min problem has no solution). Whenever the zero-game has an equilibrium, which is a saddle point, equality must hold in the well-known min-max inequality, implying that the firm's SIP model and Adverse Nature's SIP model must produce the same optimal objective value. By contrasting the two SIP models of the firm and Adverse Nature, we find that the difference is attributable to the additional moment constraints associated with the firm's IC constraints. Thus, if equality holds in the min-max inequality, these additional moment constraints are binding but the relevant Lagrangian multipliers are zero, enabling us to solve numerous zero-sum games in closed forms that are unavailable in the extant literature.

Although we can apply the moral hazard approach from the firm's perspective, the relevant IC constraints are much more difficult to simplify. The main reason is that the firm employs a pure strategy while Adverse Nature employs a mixed strategy subject to the moment constraints. Intuitively, the IC constraints associated with pure strategy are simpler than those associated with mixed strategy. Occasionally, the SIP model of the min-max version produces multiple solutions, especially when the first derivative of the ex post payoff function is a step function (implying that the ex post payoff function is piece-wise linear). In this circumstance, we select an optimal solution that satisfies the property of zero Lagrangian multipliers.

1.1. Literature Review

Our paper relates to the literature on robust operations management. We refer readers to Lu and Shen (2021) for an updated literature review. The research on optimal ordering decisions with limited information about the demand distribution is inspired by Scarf (1958). In this literature stream, the firm knows only the mean and variance of demand and aims to find an order quantity to maximize her expected profit against the worst possible distribution. With a single supply source, the worst-case distribution is a two-point distribution. The robust optimal order quantity and the resulting profit can be derived in closed form (see Corollary 3 for details). Gallego and Moon (1993) provide a more concise proof of Scarf's result by using the Cauchy-Schwartz inequality. Natarajan et al. (2018) examine the impact of an asymmetric demand distribution by using second-order partitioned statistics to measure distributional asymmetry. Das et al. (2021) analyze the impact of heavy-tailed demand distributions by assuming that the newsvendor firm knows only the mean and the t-th moment of the demand distribution (where $t > 1$ is a real number). We refer to this model as the "$1+t$" model. While Das et al. (2021) encounter challenge in deriving the objective function and are unable to solve the robust inventory level in closed forms, we overcome this challenge. Minimax regret is another robust decision rule commonly used in the literature. Yue et al. (2006) define the value of information (VOI) as the difference between knowing and not knowing the underlying demand distribution, while Perakis and Roels (2008) refer to VOI as regret (which measures forgone profit in the absence of full information on the underlying demand distribution). Yue et al. (2006) focus on mean-variance ambiguity, whereas Perakis and Roels (2008) consider a variety of partial information on the distribution such as its mode, range, mean, and variance.

The robust optimization literature distinguishes between risk aversion and ambiguity aversion. A risk-averse decision maker prefers an order quantity that avoids profit volatility in addition to the expected profit, whereas an ambiguity-averse decision maker does not have complete knowledge of the demand distribution and thus prefers an order quantity that is distributionally robust (Han et al. 2014). Incorporating the variance of the profit, Han et al. (2014) study a distributionally robust newsvendor model by combining both risk aversion and ambiguity aversion. Several recent articles (Yang et al. 2018, Kouvelis et al. 2021, Yang et al. 2021) employ the conditional value at risk as a measure of risk tolerance. With advances in machine learning, several recent articles have proposed data-driven methods to determine the robust order quantity. For instance, Chen and Xie (2021) consider an unknown joint distribution for demand and yield, while He and Lu (2021) consider a price-setting newsvendor firm that is partially informed about the demand distribution and has limited data on a few historical prices. It is possible that multiple decision makers could simultaneously face ambiguity. For instance, Fu et al. (2018) consider an agricultural supply chain where both the upstream and downstream entities face mean-variance ambiguity. Li and Kirshner (2021) label this ambiguous environment "two-sided ambiguity" and consider the contracting issue between the firm and her sales agent.

1.2. Our Contributions

We make the following contributions to the literature.

- We develop a new and efficient method to solve a class of zero-sum games under moment conditions. By solving the zero-sum games, we solve the corresponding max-min optimization models, which are popular in the literature of robust operations management.

• We demonstrate that using min-max inequality and the property of zero Lagrangian multipliers, we can efficiently identify the equilibrium due to the following reasons. First, the IC constraints associated with pure strategy are much simpler than those associated with mixed strategy. Second, our method integrates the firm's FOCs into the new formulation such that we can solve the equilibrium without explicitly deriving the firm's objective function.

• We develop closed-form solutions for three important application examples. In the first example (i.e., the $1+t$ model), the available information includes the mean and the t-th moment. This example considers the impact of heavy (when $1<t<2$) or light (when $t>2$) tail on inventory planning. In the second example, the ex post payoff function is piece-wise linear (with $n \geq 1$ pieces) and concave. This example prescribes how to use multiple supply sources or option contracts to cope with random demand and has broad applications in the electronic appliance, energy, and remanufacturing industries. In the third example, we consider the impact of component commonality on inventory planning. By solving these examples in closed forms, we demonstrate the scalability and efficiency of our method and generate new results that are unavailable in the extant literature.

The remaining sections are organized as follows. Section 2 introduces the new method. Section 3 solves the $1+t$ newsvendor model. Section 4 investigates a capacity planning model with multiple supply sources under mean-variance conditions. Section 5 analyzes the impact of component commonality. Section 6 discusses a few technical issues and Section 7 concludes the paper. We present all the technical proofs in Online Appendix.

2. The New Method

2.1. The Firm's Perspective

Let $\tilde{\theta}$ be a random variable that affects the firm's ex post payoff. In supply chains, this random variable $\tilde{\theta}$ can represent the random demand or yield. We use θ to indicate the realization and $F(\theta)$ to represent the cumulative distribution function. However, the firm does not know the exact functional form of $F(\theta)$ except the mean and variance. Specifically, let $E\left(\tilde{\theta}\right)=\mu>0$ be the mean and $Var\left(\tilde{\theta}\right)=\sigma^2>0$ be the variance of $\tilde{\theta}$. We also let $\rho=\frac{\sigma}{\mu}>0$ be the coefficient of variation. We use σ or $\rho\mu$ to denote the standard deviation whenever convenient. In Section 3, we generalize the analysis by replacing variance with the t-th moment, where $t>1$. A wide range of max-min optimization models can be abstracted as the following distributionally robust optimization model:

$$Z=\max_{\mathbf{Q}\geq\mathbf{0}}\left\{\inf_{F\in\Omega}\int_0^\infty Z\left(\theta|\mathbf{Q}\right)dF\left(\theta\right)\right\}, \tag{2.1}$$

in which Ω represents the ambiguity set (or the feasible action space of Adverse Nature) and $Z\left(\theta|\mathbf{Q}\right)$ is the firm's ex post payoff function when the realized random variable is θ and her action vector is $\mathbf{Q} = (Q_1, Q_2, ...Q_n)$, implying that the firm can take multiple actions such as choosing prices or production quantities. The mathematical properties of the ex post payoff function $Z(\theta|\mathbf{Q})$ vary from case to case. On the other hand, the partial information about $F(\theta)$ determines the ambiguity set Ω. With mean and variance, we define the ambiguity set Ω as follows:

$$\Omega = \left\{ F(\theta) \,|\, \int_0^\infty dF(\theta) = 1, \int_0^\infty \theta dF(\theta) = \mu, \int_0^\infty \theta^2 dF(\theta) = \mu^2 + \sigma^2 \right\}. \tag{2.2}$$

The theoretical and practical parallels to equation (2.1) are abundant. For example, when Q is a single variable (rather than a vector) and $Z(\theta|Q)$ represents the newsvendor payoff function, equation (2.1) reduces to the model proposed by Scarf (1958) in his pioneering article that inspires a large and growing stream of literature on robust inventory management.

The characteristics of the ambiguity set could determine whether the worst-case expected profit is attained or approached. Observe that the ambiguity set Ω in equation (2.2) includes an infinite number of probability distributions (which can be continuous, discrete or mixed) satisfying mean-variance constraints. Many of these distributions may not satisfy log-concave, increasing failure rate, monotonic local likelihood ratio, or convex distribution function conditions, which are common in the economics and supply chain literature. Including some of these conditions could have an undesirable consequence by rendering the worst-case expected profit approached rather than attained. For example, Carroll (2015, page 542) mentions that "the worst-case payoff may be approached, but not actually attained for any technology. This is why we defined it as an infimum and not a minimum." In contrast, we impose only the moment constraints on Ω, enabling us to broaden the application and attain a minimum.

It is well-known that equation (2.1) formulates a zero-sum game between the firm and Adverse Nature. While the firm wishes to maximize her expected profit by choosing a vector $\mathbf{Q}$, Adverse Nature wishes to minimize the firm's expected profit by choosing a distribution from Ω due to the mean-variance constraints. Let $(\mathbf{Q}^*, F^*)$ be the equilibrium of the zero-sum game in equation (2.1). We refer to $\mathbf{Q}^*$ as the firm's equilibrium strategy (or her robust optimal solution), F^* as Adverse Nature's equilibrium strategy, and $Z(\mathbf{Q}^*, F^*) = Z^*$ as the value of the zero-sum game (or the firm's optimized worst-case expected profit). An important feature of the zero-sum game in equation (2.1) is that the firm employs a pure strategy (which occurs in supply chains) but Adverse Nature employs a mixed strategy. Both the firm and Adverse Nature have infinite feasible actions while the finite zero-sum game (pioneered by Nash 1951) restricts the number of feasible actions to be finite.

To cope with the infinite nature of the zero-sum game, the traditional method applies SIP tools to reformulate equation (2.1) as follows:

$$\begin{aligned} P = & \max_{\mathbf{Q}\geq\mathbf{0}} \max_{y_0,y_1,y_2} \left\{y_0 + y_1\mu + y_2\left(\mu^2+\sigma^2\right)\right\} \\ \text{s.t.} \quad & y_0 + y_1\theta + y_2\theta^2 \leq Z\left(\theta|\mathbf{Q}\right), \forall\theta\geq 0. \end{aligned} \tag{2.3}$$

For any given $\mathbf{Q}$, the inner maximization problem in equation (2.3) is a *linear* SIP model, in which probabilistic resources are being traded. In the literature of linear programming, the notion of trading resources is commonly used when interpreting a primal-dual relationship. Because our model involves probabilistic concepts (such as mean), we introduce probabilistic resources. The three decision variables y_0, y_1, and y_2 are the shadow prices of the total probability, the mean, and the variance resources, respectively. When Adverse Nature chooses a realization θ, the firm obtains her ex post payoff $Z(\theta|\mathbf{Q})$ and simultaneously purchases 1 unit of the total probability resource, θ units of the mean resource, and θ^2 units of the variance resource from Adverse Nature. By supplying these probabilistic resources, Adverse Nature generates an ex post income that equals $y_0+y_1\theta+y_2\theta^2$. However, the firm is protected by a limited liability such that the ex post payment cannot exceed the firm's ex post payoff, giving rise the SIP constraints in equation (2.3). Observe that the expected income of Adverse Nature equals $E\left(y_0+y_1\theta+y_2\theta^2\right) = y_0 + y_1\mu + y_2\left(\mu^2+\sigma^2\right)$, which is the objective function in equation (2.3). In the outer maximization problem, while the firm chooses her action vector $\mathbf{Q}$, Adverse Nature determines the shadow prices to maximize the expected income subject to the limited liability constraints. The notion of trading probabilistic resources becomes helpful when we explain our new method in the next subsection.

We use the subscript wst to indicate the worst case. When the firm plays an arbitrary strategy $\mathbf{Q}$, let $F_{wst}\left(\mathbf{Q}\right)$ be the best response of Adverse Nature (or the firm's most unfavorable distribution). When the firm plays her equilibrium strategy $\mathbf{Q}^*$, it must hold that $F_{wst}\left(\mathbf{Q}^*\right) = F^*$. However, for any $\mathbf{Q}\neq\mathbf{Q}^*$, $F_{wst}(\mathbf{Q})$ may not be identical to F^*. Therefore, $Z_{wst}(\mathbf{Q}) = Z(\mathbf{Q}, F_{wst}(\mathbf{Q}))$ formulates the firm's objective function, in which the firm anticipates that Adverse Nature plays the best response $F_{wst}(\mathbf{Q})$ if she plays $\mathbf{Q}$. The first step of the traditional method is to explicitly derive $Z_{wst}(\cdot)$ and the second step is to optimize $Z_{wst}(\cdot)$. We briefly describe the first step. For any given $\mathbf{Q}$, the inner maximization problem in equation (2.3) is a linear SIP model such that KKT conditions are sufficient and necessary. If the number of binding constraints in equation (2.3) is finite, we can obtain the most unfavorable distribution using complementary slackness and generalized finite sequence. If the number of binding constraints is infinite, we can solve a differential equation to determine the most unfavorable distribution (Carrasco et al. 2018). Thus, we can guarantee that the equilibrium $(\mathbf{Q}^*, F^*)$ is attained rather than approached.

The success of the first step critically depends on the context of the model but inevitably affects the tractability of the second step. Unfortunately, in a range of circumstances, the objective function $Z_{wst}(\cdot)$ could be either analytically unavailable or overly complex. For example, Section 3 studies the $1+t$ model, in which Das et al. (2021) apply the traditional method and find that the explicit form of $Z_{wst}(q)$ is unavailable when q is large. Guo et al. (2022) advance the analysis by deriving a semi-closed form of $Z_{wst}(q)$ for large q. Whenever the analytical form of the objective function $Z_{wst}(q)$ is absent, the traditional method is unable to solve the robust solution q^* in closed form. In contrast, our new method overcomes this challenge. In Section 4, we develop a capacity planning model with $n \geq 1$ option contracts. When $n = 1$, the model reduces to Scarf's model where the ex post payoff function has two pieces. When $n = 2$, the ex post payoff function has three pieces. In the context of strangle option contracts, Natarajan and Zhou (2007) apply the traditional method and complete the first step to obtain the objective function $Z_{wst}(q_1, q_2)$ for any given pair of (q_1, q_2). However, Natarajan and Zhou have not optimized (q_1, q_2) because their goal is to evaluate the performance of any given strangle option contract. Figure 2 of Natarajan and Zhou (2007) illustrates the piece-wise objective function $Z_{wst}(q_1, q_2)$, which has 4 cases and the boundary of each case is given by non-trivial quadratic curves. When n increases, the number of cases will grow to 2^n, making the objective function $Z_{wst}(\mathbf{Q})$ too complex to be tractable. Without explicitly deriving $Z_{wst}(\mathbf{Q})$, the traditional method encounter difficulty in analyzing the case with an arbitrary $n \geq 1$.

2.2. Adverse Nature's Perspective

The innovative feature of our method is that we bypass the obstacles in deriving $Z_{wst}(\cdot)$ and directly attack the equilibrium $(\mathbf{Q}^*, F^*)$. The crucial difference is that we solve the equilibrium from the perspective of Adverse Nature as follows:

$$Z_1 = \inf_{F \in \Omega} \left\{ \max_{\mathbf{Q} \geq \mathbf{0}} \int_0^\infty Z(\theta|\mathbf{Q})\, dF(\theta) \right\}. \tag{2.4}$$

Equation (2.4) formulates a robust moral hazard model, in which Adverse Nature jointly chooses a distribution and an action vector to minimize the firm's expected profit subject to the firm's IC constraints, which state that the chosen action vector must maximize the firm's expected profit. Under the assumption that $Z(\mathbf{Q}|F)$ is quasi-concave with respect to $\mathbf{Q}$, we can simplify the firm's IC constraint by using her FOCs: $\int_0^\infty \frac{\partial Z(\theta|\mathbf{Q})}{\partial Q_i} dF(\theta) = 0$, for $i = 1, 2, ..., n$. This first-order approach is prevalent in economics. We observe that the firm's FOCs become additional moment constraints and equation (2.4) changes to:

$$Z_1 = \inf_{\mathbf{Q} \geq \mathbf{0}, F \in \Omega} \left\{ \int_0^\infty Z(\theta|\mathbf{Q})\, dF(\theta) \right\},$$

$$\text{s.t.} \quad \begin{cases} \int_0^\infty dF(\theta) = 1, \\ \int_0^\infty \theta dF(\theta) = \mu, \\ \int_0^\infty \theta^2 dF(\theta) = \mu^2 + \sigma^2, \\ \int_0^\infty \frac{\partial Z(\theta|\mathbf{Q})}{\partial Q_i} dF(\theta) = 0, \text{ for } i = 1, 2, ..., n. \end{cases} \tag{2.5}$$

Using duality, we can conveniently reformulate model Z_1 as the following SIP model:

$$\begin{aligned} P_1 = & \max_{y_0, y_1, y_2, \mathbf{Q}, \mathbf{a}} \left\{ y_0 + y_1\mu + y_2\left(\mu^2 + \sigma^2\right) \right\} \\ \text{s.t.} \quad & y_0 + y_1\theta + y_2\theta^2 + \sum_{i=1}^n a_i \frac{\partial Z(\theta|\mathbf{Q})}{\partial Q_i} \leq Z(\theta|\mathbf{Q}), \forall \theta \geq 0, \end{aligned} \tag{2.6}$$

where $\mathbf{a} = (a_1, a_2, ..., a_n)$ is the vector of Lagrangian multipliers associated with the firm's FOCs. We use $*$ to indicate the optimal objective value or optimal solution.

Contrasting model P in equation (2.3) with model P_1 in equation (2.6), we find that they both have the same objective function $y_0 + y_1\mu + y_2(\mu^2 + \sigma^2)$. Any solution that is feasible in model P is also feasible in model P_1 but the opposite is not true unless all $a_i = 0$. Thus, it must hold that $Z^* \leq Z_1^*$, which is consistent with the well-known min-max inequality (e.g., Sion 1958). The equal sign holds if and only if model P_1 encompasses an optimal solution satisfying $a_i^* = 0$ for all $i = 1, 2, ..., n$. Furthermore, it is well-known that if $Z^* = Z_1^*$, then the zero-sum game in equation (2.1) has an equilibrium.

Theorem 1 *If the zero-sum game in equation (2.1) has an equilibrium, the following results hold: i) $P^* = Z^* = Z_1^* = P_1^*$ and ii) the SIP model P_1 in equation (2.6) must encompass an optimal solution satisfying the property that $a_i^* = 0$ for $i = 1, 2, ..., n$.*

Theorem 1 describes how to directly attack the equilibrium $(\mathbf{Q}^*, F^*)$ by avoiding the intermediate step of deriving $Z_{wst}(\cdot)$. The first derivative $\frac{\partial Z(\theta|\mathbf{Q})}{\partial Q_i}$ (or the marginal payoff), which appears in equation (2.6), plays a crucial role. The intuition becomes clearer under the notion of trading probabilistic resources. In the new model Z_1 in equation (2.5), we regard $\int_0^\infty \frac{\partial Z(\theta|\mathbf{Q})}{\partial Q_i} dF(\theta) = 0$ as generalized moment conditions such that n types of new resources are added to the trade. We refer to the new resource associated with Q_i as type-i new resource. When Adverse Nature chooses a realization θ, in addition to the three existing probabilistic resources, the firm purchases $\frac{\partial Z(\theta|\mathbf{Q})}{\partial Q_i}$ units of type-i new resource by paying the shadow price a_i per unit, giving rise to the summation at the left-hand-side of equation (2.6). Because the purchasing quantity of type-i new resource equals the firm's marginal payoff, the trading of new resources effectively integrates the firm's FOCs into the SIP model in equation (2.6). As the gain of the firm is the loss of Adverse Nature, $a_i = 0$ must hold in equilibrium; otherwise, the summation in equation (2.6) is non-zero and Adverse Nature cannot break even by supplying the new resources. This observation confirms the property of zero Lagrangian multipliers that Theorem 1 indicates.

3. The First and t-th Moment

To assist readers in deepening their understanding of our method and to prepare for the capacity planning model in Section 4, we consider the following newsvendor model with a single capacity. The price of the product is $p > 0$ and the capacity reservation cost is $c \geq 0$ per unit. When the realized demand θ arrives, the firm uses the available capacity to satisfy the demand by incurring a processing cost $r \geq 0$. Any unused capacity expires and the reservation cost is not refunded. The demand in excess of the available capacity is lost. We find that the ex post payoff function equals $Z(\theta|q) = (p-r)\min(\theta, q) - cq$ and the first derivative satisfies $\frac{\partial Z(\theta|q)}{\partial q} = -c$ if $\theta < q$ and $\frac{\partial Z(\theta|q)}{\partial q} = p - r$ if $\theta > q$. When $\theta = q$, $Z(\theta|q)$ is not differentiable. By default, we let $\frac{\partial Z(\theta|q)}{\partial q} = (p-r)$ if $\theta = q$ so that the first derivative and the cumulative distribution function are both right continuous. We slightly change the ambiguity set in equation (2.2) as follows:

$$\Omega = \left\{ F(\theta) \mid \int_0^\infty dF(\theta) = 1, \int_0^\infty \theta dF(\theta) = \mu, \int_0^\infty \theta^t dF(\theta) = m_t \right\}, \tag{3.1}$$

where m_t is known and represents the t-th moment of demand. In other words, we broaden the analysis to the case in which the mean and the t-th moment of the demand distribution (where $t > 1$ is a real number) are known.

Based on Theorem 1, we formulate the following SIP model:

$$\begin{aligned} P_1 = & \max_{q \geq 0, y_0, y_1, y_2} \{y_0 + y_1\mu + y_2 m_t\} \\ \text{s.t.} \quad & y_0 + y_1\theta + y_2\theta^t - ac \leq (p-r)\theta - cq, \forall \theta \in [0, q), \\ & y_0 + y_1\theta + y_2\theta^t + a(p-r-c) \leq (p-r-c)q, \forall \theta \geq q. \end{aligned} \tag{3.2}$$

Our new method enables us to rapidly solve the robust capacity level q^*. Before characterizing the equilibrium, we establish the following nonlinear equation:

$$H(x) = \frac{p-r-c}{p-r}x^t + \frac{c}{p-r}\left[x + \frac{p-r}{c}(\mu - x)\right]^t = m_t. \tag{3.3}$$

With $t > 1$, Jensen's inequality implies that $m_t \geq \mu^t$.

Lemma 1 *a) If*

$$\mu^t \leq m_t \leq \frac{\mu^t}{\left(\frac{c}{p-r}\right)^{t-1}}, \tag{3.4}$$

then equation (3.3) has a root θ_1 satisfying $0 \leq \theta \leq \mu$. Let $\theta_2 = \theta_1 + \left(\frac{p-r}{c}\right)(\mu - \theta_1)$ and define the following two-point distribution:

$$\begin{cases} \Pr\left(\tilde{\theta} = \theta_1\right) = \frac{p-r-c}{p-r}, \\ \Pr\left(\tilde{\theta} = \theta_2\right) = \frac{c}{p-r}. \end{cases} \tag{3.5}$$

The distribution in equation (3.5) satisfies the conditions on the mean and t-th moment.

b) If

$$m_t \geq \frac{\mu^t}{\left(\frac{c}{p-r}\right)^{t-1}} \tag{3.6}$$

then the root of equation (3.3) is negative. We let $\theta_h = \left(\frac{m_t}{\mu}\right)^{\frac{1}{t-1}}$ *and define the following two-point distribution:*

$$\begin{cases} \Pr\left(\tilde{\theta}=0\right) = 1-\frac{\mu}{\theta_h}, \\ \Pr\left(\tilde{\theta}=\theta_h\right) = \frac{\mu}{\theta_h}. \end{cases}, \tag{3.7}$$

which also satisfies the conditions on the mean and t-th moment.

When the newsvendor ratio is sufficiently high, condition (3.4) holds.

3.1. Robust Inventory Level

The extant literature regards the $1+t$ model as notoriously difficult because the objective function $Z_{wst}(q)$ cannot be explicitly derived by the traditional method. For example, Das et al. (2021, pages 1097-1098) explain the "implausibility" in deriving $Z_{wst}(q)$. Similarly, Guo et al. (2022, page 14) suggest that a semi-closed form for $Z_{wst}(q)$ is "probably the best one can hope for". When the explicit expression of the objective function $Z_{wst}(q)$ remains analytically unavailable, the first derivative $\frac{\partial Z_{wst}(q)}{\partial q}$ is unknown, preventing the extant literature from solving the equilibrium strategy q^* in closed form. In contrast, our new method overcomes the relevant technical challenge.

Proposition 1 *In the $1+t$ model formulated in equation (3.2), the equilibrium is one of the following two cases:*

a) If condition (3.4) holds, the firm's equilibrium strategy q^* satisfies that

$$q^* = \theta_1 + \frac{\theta_2^{t-1}(\theta_2-\theta_1)}{\theta_2^{t-1}-\theta_1^{t-1}} - \frac{\theta_2^t - \theta_1^t}{t\left(\theta_2^{t-1}-\theta_1^{t-1}\right)}, \tag{3.8}$$

while Adverse Nature's equilibrium strategy satisfies equation (3.5). Thus, the value of the zero-sum game equals $P_1^* = (p-r-c)\,\theta_1$.

b) If condition (3.6) holds, the firm's equilibrium strategy satisfies $q^*=0$ while Adverse Nature's equilibrium strategy satisfies equation (3.7). Thus, the value of the zero-sum game equals $P_1^*=0$.

Proposition 1 underscores the key advantage of our method over the extant literature. We can derive q^* despite that the objective function $Z_{wst}(q)$ could be analytically unavailable. Proposition 1 also enriches our understanding about the effect of heavy-tail (or light-tail) behaviors on inventory planning. For example, because $H(\theta_1)=m_t$ and $H(\cdot)$ is decreasing, we observe that the value of the zero-sum game decreases with respect to m_t. Furthermore, we can expand the ambiguity set in equation (3.1) by letting $\int_0^\infty \theta^t dF(\theta) \leq m_t$ so that we include regularly heavy-tailed distributions (Das et al. 2021). When the sign associated with m_t changes from "=" to "$\leq$", the shadow price y_2 must be nonpositive. However, the proof of Proposition 1 confirms that $y_2^* \leq 0$ and hence, Proposition 1 continues to hold.

Proposition 1 reveals an important property that the firm's worst-case distribution F^* exhibits. If the ex post profit function is piece-wise linear and has *two* pieces, then F^* must be a *two*-point distribution with realizations θ_1^* and θ_2^* (where $\theta_1^* \leq \theta_2^*$). This two-point distribution not only satisfies the moment constraints but also allocates a mass probability to the low realization θ_1^* according to the firm's newsvendor ratio (i.e., $\Pr\left(\tilde{\theta} = \theta_1^*\right)$ equals her newsvendor ratio). In general, to identify a two-point distribution, we need to determine four parameters (i.e., two probability masses and two realized values). Because the two probability masses are based on the newsvendor ratio, we only need to determine the two realized values by applying the moment conditions. The relevant intuition becomes clearer if we examine the equilibrium from the perspective of Adverse Nature. When the ex post profit function is piece-wise linear and has only two pieces, the marginal ex post payoff (i.e., the first derivative $\frac{\partial Z(\theta|q)}{\partial q}$) takes only two possible values. To randomize the demand, the optimal strategy for Adverse Nature must be choosing between two realized values θ_1 and θ_2 such that θ_1 gives the firm a lower marginal ex post payoff and θ_2 gives the firm a higher marginal ex post payoff. This explains why Proposition 1 converges to a two-point distribution. On the other hand, Adverse Nature anticipates that the firm must play her optimal response, which follows her newsvendor ratio. Therefore, when determining how frequently to choose the low realized value θ_1, Adverse Nature also follows the firm's newsvendor ratio, explaining the equilibrium strategy F^* in equation (3.5). When $t = 2$, Proposition 1 becomes Scarf's model. We explain the relevant details in Part A of Online Appendix.

An early criticism of Scarf's result is that the worst-case distribution is a two-point distribution. However, a two-point distribution is the most natural response of Adverse Nature because the marginal payoff takes only two values in the standard newsvendor model. The relevant intuition can also be generalized to the capacity model in Section 4. Specifically, when the ex post payoff function is concave piece-wise linear and has $(n+1)$ pieces (suggesting that the marginal payoff takes $(n+1)$ values), the equilibrium strategy played by Adverse Nature is characterized by a $(n+1)$-point distribution. The probability mass to be allocated to each point depends on the firm's best response. Certainly, mean and variance alone are insufficient to specify the relevant $(n+1)$-point distribution when $n \geq 2$. We need to perform additional analysis to derive the closed-form solution.

4. Capacity Planning Model

4.1. Industrial Background

Sourcing strategy is critical for many firms to excel in the current volatile business environment. Despite low administrative costs, single sourcing can expose firms to significant risks caused by demand or supply uncertainties. To gain a competitive advantage, many firms diversify their supply chains to enhance their flexibility to cope with uncertainties. For example, Hewlett-Packard is a pioneer that uses option contracts to manage the supply of memory devices (Fu et al. 2010). Each option contract specifies the premium (or the reservation fee) and the exercising fee for each memory device. After observing the realized demand, Hewlett-Packard determines which option contracts to exercise. Any unexercised option contract expires but the exercising fee is avoided. If demand exceeds the total quantities that can be satisfied under the purchased option contracts, the excess demand is lost when the required parts cannot be replaced by the standard parts from the spot market.[2] The second example is E.ON (`www.eonenergy.com`), one of the leading energy companies in the United Kingdom. E.ON reserves capacity from fossil fuel generators, wind farms, and nuclear generators. When electricity demand arrives, E.ON uses the reserved capacity to generate electricity. There could be additional costs of generating electricity from the use of reserved generators and the transmission of the generated electricity to external customers. These additional costs could be source dependent and are incurred only in the fulfillment stage. A minor difference is that the excess demand for electricity can be satisfied by the spot market rather than lost. The third example is Xerox Australia, which recycles and remanufactures photocopiers (Kerr and Ryan 2001). When photocopiers are returned to Xerox after the end of lease contracts, Xerox inspects their condition and sorts them into four grades: grade 1 (suitable for refurbishment), grade 2 (suitable for reprocessing), grade 3 (suitable for remanufacturing), and grade 4 (suitable for asset recovery or disposal). After an order for used photocopiers arrives, Xerox implements a priority rule that exhausts an alphabetically lower grade (which has a better condition) before using another grade. Grade 1 and grade 2 are cleaned and repaired. High-frequency-service parts are replaced regardless of their condition or use. Other parts are replaced depending on their condition and expected remaining life. Grade 3 and grade 4 photocopiers must undergo a labor-intensive disassembly process. Good-quality parts are cleaned, tested, and reconditioned. Some photocopiers are then reassembled, while others are sent without reassembly to a disposal facility. Customers who purchase used photocopiers are unable to discern the initial grade of the machines. If the quantity of used photocopiers is insufficient, the demand is lost or the customer is persuaded to buy other machines.

4.2. The Robust Capacity Planning Model

The operations of Hewlett-Packard, E.On, and Xerox Australia share two common characteristics: i) the firm can access $n \geq 1$ supply sources that are substitutes for each other, and ii) the firm makes the procurement and fulfillment decisions in two stages. Accordingly, we develop the following inventory model. The firm faces an uncertain demand and procures from n sources, where each source can represent a different supplier, grade, or option contract. Let $\tilde{\theta}$ be a nonnegative random variable denoting the external demand for the firm's end product. The cumulative distribution function of demand θ is $F(\theta)$. We make a critical departure from the extant literature (e.g., Martínez-de Albéniz and Simchi-Levi 2009, Fu et al. 2010) by assuming that $F(\theta)$ is *unknown*. We assume that the firm knows only the mean ($\mu > 0$) and the variance ($\sigma^2 > 0$) associated with demand $\tilde{\theta}$. Without complete knowledge of the demand distribution, the firm applies distributionally robust approach to creating her capacity plan. Let $\rho = \frac{\sigma}{\mu}$ be the coefficient of variation of demand.

The sourcing decisions proceed in two stages. In Stage 1 (which we refer to as the procurement stage before the selling season), the firm reserves q_i units of capacity from each source i ($i = 1, 2, \cdots, n$) by paying a reservation cost c_i per unit. In Stage 2 (which we refer to as the fulfillment stage), the external demand θ is realized, and the firm uses available capacities to produce the end products. When using the capacity of source i to produce one unit of end product, the firm incurs an additional processing cost $r_i \geq 0$. The end products delivered from each source have the same quality and functionality so that every external customer regards them as indistinguishable. By satisfying one unit of demand, the firm receives the same amount of revenue p from each external customer. If the reserved capacities are insufficient to satisfy all of the external demand, the excess demand is lost.[3] Any unused capacities expire without any salvage value but the processing cost is also avoided. For expositional simplicity, we define source-$(n+1)$ as an artificial source representing lost sales such that $c_{n+1} = 0$ and $r_{n+1} = p$. We use a bold letter to represent a vector. Let $\mathbf{q} = (q_i)$ be the capacity vector (where $i = 1, 2, \cdots, n$) chosen by the firm in the procurement stage.

We refer to all supply sources as a portfolio. Before introducing the assumptions on cost parameters, we define two sequences as follows.

Definition 1 *Let $\alpha_0 = 0$ and $\alpha_{n+1} = 1$. Define*

$$\begin{cases} \alpha_i = 1 - \frac{c_i - c_{i+1}}{r_{i+1} - r_i}, \text{ for } i = 1, 2, \cdots, n, \\ \beta_i = \alpha_i - \alpha_{i-1}, \quad \text{for } i = 1, 2, \cdots, n+1. \end{cases} \tag{4.1}$$

For ease of exposition and given the need to eliminate unattractive supply sources, we assume that the sequence α_i is strictly increasing in i, which is consistent with the extant literature (e.g., Martínez-de Albéniz and Simchi-Levi 2009, Fu et al. 2010) and ensures that in the robust optimal solution, the order quantity for each supply source is positive. Otherwise, some supply sources will never be used and can be eliminated from the analysis.

Remark 1 *If $\{\alpha_i\}$ is increasing in i, then the following results hold: i) $c_1 + r_1 < c_2 + r_2 < \cdots < c_n + r_n < p$ and ii) $r_1 < r_2 < \cdots < r_n$ and $c_1 > c_2 > \cdots > c_n$.*

We omit the proof of Remark 1 and refer readers to Lemma 1 of Fu et al. (2010). We can visualize the portfolio by using the reservation cost c_i as the vertical coordinate and the processing cost r_i as the horizontal coordinate. Remark 1 implies that the path of the portfolio (which starts from supply source 1 and connects all sources including the artificial source-$(n+1)$ that represents lost sales) is convex decreasing. Remark 1 also implies that if the capacity vector is fixed, the firm will use a priority rule to use the available capacity to satisfy the realized demand. Specifically, source $(i+1)$ will not be used unless the reserved capacity of source i is exhausted.

Let $Q_i = \sum_{j=1}^{i} q_i$ be the total capacity of the first i sources. By default, we let $Q_0 = 0$ such that $Q_i \geq Q_{i-1}$, which we refer to as the monotonicity constraints on Q_i. The definition of Q_i resembles the echelon inventory level that Veinott (1965) proposes. Because the mapping between vectors $\mathbf{Q} = (Q_i)$ and $\mathbf{q} = (q_i)$ is unique, hereafter, we regard Q_i as the decision variables to facilitate analysis. Let $(\cdot)^+ = \max(0, \cdot)$.

Lemma 2 *For any given capacity vector $\mathbf{Q}$, the firm's ex post profit equals*

$$Z(\theta|\mathbf{Q}) = \sum_{i=1}^{n} \left[(p - r_i)(\min(Q_i, \theta) - \min(Q_{i-1}, \theta)) - c_i(Q_i - Q_{i-1})\right], \tag{4.2}$$

which is continuous, concave, and increasing in the realized demand θ.

The $Z(\theta|\mathbf{Q})$ function in equation (4.2) is piece-wise and has $(n+1)$ different cases, depending on the value of the realized demand θ. Let $\delta_i = \frac{Z(\theta|\mathbf{Q})}{\partial Q_1}$ be the first derivative in these $(n+1)$ cases. The proof of Lemma 2 shows that the sequence $\{\delta_i\}$ satisfies the following recursive equation:

$$\delta_i = r_i - r_1 - c_1, \text{ for } i = 1, 2, \cdots, n+1. \tag{4.3}$$

The two sequences $\{\beta_i\}$ and $\{\delta_i\}$ determine the equilibrium. We refer to δ_i as the marginal impact and provide the managerial interpretation of δ_i after we derive the robust capacity plan.

4.3. Adverse Nature's Model

To apply Theorem 1, we need to simplify the relevant IC constraints. Assume that the demand distribution $F(\theta)$ is known. The firm's expected profit equals $Z(\mathbf{Q}) = \int_0^\infty Z(\theta|\mathbf{Q})\, dF(\theta)$. Let $\tilde{\mathbf{Q}}$ be the firm's distribution-dependent optimal capacity plan.

Lemma 3 *In the benchmark without ambiguity, the firm's expected profit $Z(\mathbf{Q})$ is concave in $\mathbf{Q}$, and her optimal capacity vector $\tilde{\mathbf{Q}}$ satisfies $F\left(\tilde{Q}_i\right) = \alpha_i$, implying that the firm's optimal capacity level for source-i equals*

$$\tilde{q}_i = F^{-1}(\alpha_i) - F^{-1}(\alpha_{i-1}) \text{ for any } i = 1, 2, \cdots, n, \tag{4.4}$$

where F^{-1} is the inverse function of F.

Lemma 3 indicates that the optimal capacity plan can be described by a sequence of percentiles, explaining why the cumulative probability α_i must be increasing in i to ensure that $\tilde{q}_i > 0$. In the proof of Lemma 3 (we refer readers to equation (B-7) in the Appendix), we expand the firm's expected profit as the sum of n separate identities. Each of these n identities involves only one Q_i (which is the total quantity of the first i sources) and is concave in Q_i. The definition of Q_i must imply that $Q_i \geq Q_{i-1}$. After relaxing the monotonic constraint on Q_i and solving the FOC, we obtain that $F\left(\tilde{Q}_i\right) = \alpha_i$. Thus, if α_i is monotonic in i, then the candidate solution $\tilde{Q}_i$ is increasing in i, making $\tilde{q}_i = \tilde{Q}_i - \tilde{Q}_{i-1}$ optimal and positive. If $\alpha_i \leq \alpha_{i-1}$, then the monotonicity constraint $Q_i \geq Q_{i-1}$ must be binding, making $q_i = Q_i - Q_{i-1} = 0$ (implying that source i is not used in the optimal solution). We emphasize that the binding status of the monotonicity constraint $Q_i \geq Q_{i-1}$ depends on the monotonicity of α_i rather than on the demand distribution F. Thus, we can apply the process of elimination to reduce the number of supply sources.

The concavity of $Z(\mathbf{Q})$ shown in Lemma 3 indicates that we can replace the IC constraints with the FOCs. However, for ease of analysis (i.e., $\frac{Z(\theta|\mathbf{Q})}{\partial Q_i}$ is rather complex) and the need to involve the mass probabilities $\{\beta_i\}$, we use a different form of FOCs as follows: $F(Q_i) = \alpha_i$, where α_i is uniquely determined by exogenous cost parameters according to Definition 1. The constraints $F(Q_i) = \alpha_i$ are equivalent to

$$\int_{Q_{i-1}}^{Q_i} dF(\theta) = \beta_i, \text{ for } i = 1, 2, \cdots, n, \tag{4.5}$$

which remain to be moment constraints.[4] Thus, the SIP model in equation (2.6) becomes:

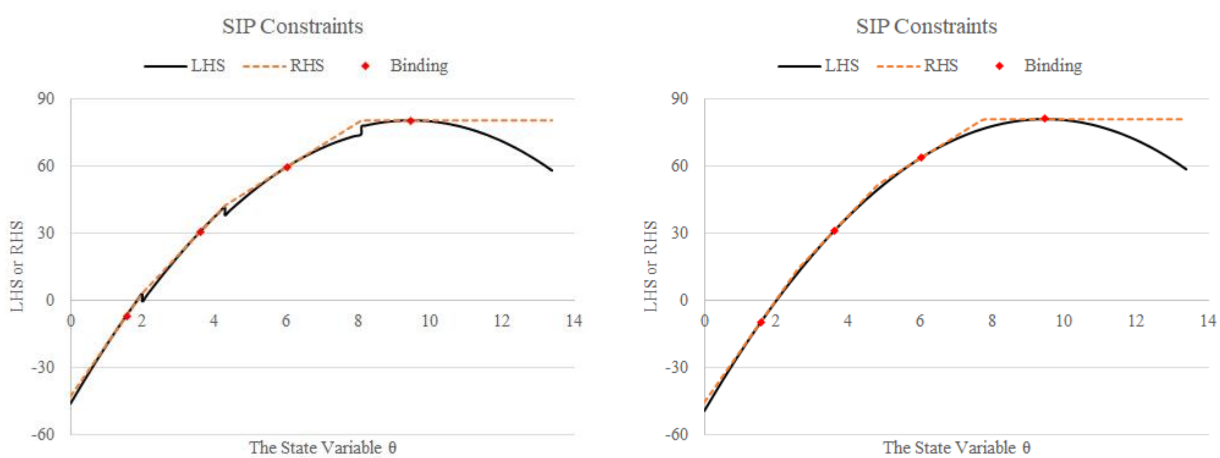
$$\begin{aligned} P_1 = & \max_{\substack{y_0, y_1, y_2, \mathbf{Q} \\ a_1, a_2, \cdots, a_n}} \left\{ a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n + y_0 + y_1\mu + y_2\left(\mu^2 + \sigma^2\right) \right\} \\ \text{s.t.} \quad & \begin{cases} a_i + y_0 + y_1\theta + y_2\theta^2 \leq Z(\theta|\mathbf{Q}), \forall \theta \in [Q_{i-1}, Q_i], i = 1, 2, \cdots, n, \\ y_0 + y_1\theta + y_2\theta^2 \leq Z(\theta|\mathbf{Q}), \forall \theta \geq Q_n. \end{cases}, \end{aligned} \tag{4.6}$$

where a_i is the shadow price for the i-th moment constraint in equation (4.5). Contrasting the two SIP models in equations (2.3) and (4.6), we observe that Theorem 1 implies that one of the optimal solutions for model P_1 satisfies $a_i = 0$ and gives the same optimal objective value for both models P and P_1.

4.4. Worst-Case Distribution

To solve the equivalent SIP model in equation (4.6), we first conjecture the binding SIP constraints and then apply a relaxation method. Figure 1a) visualizes the SIP constraints in equation (4.6) by showing that the left-hand side of the constraints is quadratic and piece-wise in θ (specifically, there exist $(n+1)$ pieces with n jumps due to $a_1, a_2, \cdots, a_n$). On the other hand, the right-hand side (RHS) of the constraints is continuous, concave, and piece-wise linear in θ as Lemma 2 suggested. We conjecture that there exist $(n+1)$ binding constraints at points labeled as θ_i $(i = 1, 2, \cdots, n+1)$ such that the first n points satisfy $\theta_i \in [Q_{i-1}, Q_i]$ and the last point satisfies $\theta_{n+1} \geq Q_n$. The condition $\theta_i \in [Q_{i-1}, Q_i]$ is consistent with Proposition 3 and incorporates the monotonicity constraints on $\mathbf{Q}$. Based on the conjectured binding constraints, we derive the relaxed solution for equation (4.6). The final step is to verify that the relaxed solution satisfies the omitted constraints and hence is optimal for equation (4.6).

Figure 1 SIP constraints



(a) When a Nonrobust $\mathbf{Q}$ Vector is Used

(b) When the Robust Vector $\mathbf{Q}^*$ is Used

Before presenting the major results, we define an important distribution as follows.

Definition 2 *Let Δ be a positive constant satisfying $\Delta = \sqrt{\sum_{i=1}^{n+1} \beta_i \delta_i^2}$, where $\{\beta_i\}$ is the mass probability sequence given by Definition 1 and $\{\delta_i\}$ is the coefficient sequence given by equation (4.3). Let*

$$\theta_i^* \stackrel{def}{=} \mu + \delta_i \frac{\sigma}{\Delta}, \text{ for } i = 1, 2, \cdots ., n+1 \tag{4.7}$$

be the i-th possible realization of θ. Using $\{\theta_i^*\}$ and $\{\beta_i\}$, we define the following distribution:

$$\Pr(\theta = \theta_i^*) = \beta_i, \text{ for } i = 1, 2, \cdots, n+1, \tag{4.8}$$

which is a discrete $(n+1)$-point distribution.

We assume $(p-r_1-c_1)\,\mu \geq \Delta\sigma$ to avoid the uninteresting case in which the firm ceases operations under mean-variance ambiguity. Because we require demand to be nonnegative, we also assume that $\theta_1^* = \mu - c_1 \frac{\sigma}{\Delta} \geq 0$.

Corollary 1 *i) It holds that $\sum_{i=1}^{n+1} \beta_i \delta_i = 0$. ii) The $(n+1)$-point discrete distribution in equation (4.8) is one of the feasible distributions in the ambiguity set Ω.*

The next proposition identifies the firm's worst-case demand distribution.

Proposition 2 *The $(n+1)$-point distribution in equation (4.8) is the firm's worst-case demand distribution F^*.*

Proposition 2 significantly advances our analysis by paving the way toward determining the firm's worst-case expected profit. Using the firm's worst-case demand distribution F^*, we can compute the firm's expected profit $Z(\mathbf{Q}|F^*)$.

Proposition 3 *There exist an infinite number of $\mathbf{Q}$ vectors that maximize the firm's expected profit $Z(\mathbf{Q}|F^*)$. However, the firm's optimal expected profit under her worst-case demand distribution F^* is unique and equals*

$$Z^* = \max_{\mathbf{Q}\geq 0} Z(\mathbf{Q}|F^*) = (p-r_1-c_1)\,\mu - \Delta\sigma. \tag{4.9}$$

Proposition 3 shows that the firm's optimal worst-case expected profit Z^* has a clean and neat form. Observe that $(p-r_1-c_1)$ is the understock cost of source-1 capacity, which is the firm's most profitable source. Equation (4.9) reveals that Z^* is increasing in the mean of demand and the understock cost of source-1 capacity and is decreasing in the constant Δ and the standard deviation of demand.

4.5. Robust Optimal Capacity Vector

Because optimizing $Z(\mathbf{Q}|F^*)$ does not result in a unique capacity vector, our remaining challenge is to solve the robust optimal capacity vector $\mathbf{Q}^*$ using Theorem 1.

Proposition 4 *The firm's robust optimal capacity vector satisfies $\mathbf{Q}^* = (Q_i^*) = \left(\frac{\theta_i^*+\theta_{i+1}^*}{2}\right)$ for $i = 1, 2, \cdots, n$, implying that the robust optimal capacity level for source-i capacity equals*

$$q_i^* = \frac{\theta_i^*+\theta_{i+1}^*}{2} - \frac{\theta_{i-1}^*+\theta_i^*}{2} = \frac{\theta_{i+1}^*-\theta_{i-1}^*}{2}. \tag{4.10}$$

Proposition 4 solves the robust optimal capacities in closed form by demonstrating that in the robust optimal solution, the total capacities of the first i sources equal the midpoint of the closed interval $\left[\theta_i^*, \theta_{i+1}^*\right]$.

Corollary 2 *The optimal solution for equation (2.3) satisfies* $y_0^* = -\frac{\Delta}{2\sigma}(\mu^2+\sigma^2)$, $y_1^* = (p-r_1-c_1)+\Delta\frac{\mu}{\sigma}$, $y_2^* = -\frac{\Delta}{2\sigma}$, *and* $Q_i^* = \left(\frac{\theta_i^*+\theta_{i+1}^*}{2}\right)$.

Corollary 2 is a direct result of Propositions 2 to 4. With all $a_i^*=0$, the left-hand side of the SIP constraints in equation (2.3) becomes one smooth piece of a quadratic curve (see Figure 1 b) for an illustration). The discrete distribution in equation (4.8) is critical for unlocking all the results. Because we can ex ante construct the sequences $\{\beta_i\}$ and $\{\delta_i\}$ by using exogenous cost parameters, the robust optimal capacity vector $\mathbf{Q}^*$ is easy to compute.

Recall that 1) the mass probabilities $\{\beta_i\}$ are based on the unambiguous solution and exogenous cost parameters and 2) the marginal impact $\delta_i = \frac{Z(\theta|\mathbf{Q})}{\partial Q_1}$ is the first derivative with respect to Q_1 in the i-th case. The cost coefficients $\{\delta_i\}$ have an interesting managerial interpretation. In our model, the firm can access $n \geq 1$ sources with source 1 being her most preferred source. If demand is deterministic, the firm can use only source 1 in her capacity plan. However, due to demand uncertainty, the firm chooses more sources with lower reservation costs even if the sum of the reservation and processing costs increases. In the i-th case, the realized demand satisfies that $Q_{i-1} \leq \theta < Q_i$ (where by default $Q_0 = 0$), source-i still has some unused capacity. If the firm increases the capacity of source-1 from q_1 to $q_1+\varepsilon$ and keeps all the other q_i's ($i \geq 2$) unchanged, then the sales quantity of source-1 increases by ε while that of source-i deceases by ε units (where $\varepsilon \in (0, Q_i-\theta)$ is a small positive number). The marginal impact includes two parts: 1) the *change* in the sales revenue due to an additional ε units of source-1 capacity and 2) the reservation cost of an additional ε units of source-1 capacity. The net impact equals

$$(p-r_1)-(p-r_i)-c_1 = r_i - r_1 - c_1 \stackrel{def}{=} \delta_i.$$

In a notable special case where $i=(n+1)$, the firm suffers lost sales when the artificial source $(n+1)$ is used. With $r_{n+1}=p$ and $c_{n+1}=0$, we observe that $\delta_{n+1} = r_{n+1}-r_1-c_1 = p-r_1-c_1 > 0$.

5. Component Commonality

The risk-pooling effect is an important topic in the supply chain literature. For example, Bimpikis and Markakis (2016) study the effect of heavy-tailed demands on risk-pooling while Govindarajan et al. (2021) study a multi-location model with transshipment. The risk-pooling effect must involve multi-dimensional random variables rather than the single-dimensional random variable that Sections 3 and 4 consider. To demonstrate that our method is capable of solving multi-dimensional problems, we revisit the inventory model that Baker et al. (1986) study.

The firm manufactures two products: product 1 and product 2. Product 1 requires one unit of product-specific component 1 and one unit of common component 0 while product 2 requires one unit of product-specific component 2 and one unit of common component 0. The production cost of component i is c_i $(i=0,1,2)$ and the selling price of product j $(j=1,2)$ is p_j. By operating an assemble-to-order system, the firm pre-stocks q_i $(i=0,1,2)$ units of component i (where $q_1+q_2\geq q_0$) and then assembles the final products only after receiving the realized demands (θ_1,θ_2), where θ_j is the realized demand for product j. The production cost of all the inventories equals $c_1q_1+c_2q_2+c_0q_0$ and is sunk at the assembly stage. When the realized demands are (θ_1,θ_2), the firm assembles s_j units of product j to maximize her total sales revenue by solving the following linear programming model:

$$\begin{aligned} & Z(\theta_1,\theta_2|q_1,q_2,q_0)=\max_{s_1,s_2\geq 0}\{p_1s_1+p_2s_3\}-c_1q_1-c_2q_2-c_0q_0 \\ \text{s.t.}\quad & s_j\leq\min(\theta_j,q_j)\ \text{ and }\ s_1+s_2\leq q_0. \end{aligned}$$

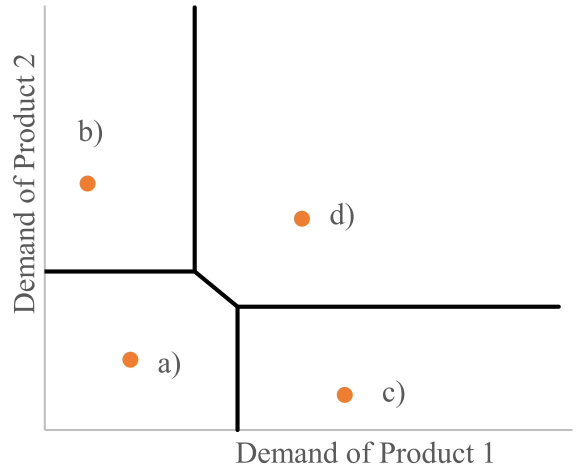
Without loss of generality, we assume that $p_1\geq p_2$ so that the firm gives product 1 a higher priority than product 2 when the supply of the common component 0 is insufficient. We obtain the marginal payoffs in the following Table 1 and illustrate the four regions associated with the realized demands (θ_1,θ_2) in Figure 2.

Table 1 The marginal payoff $\frac{\partial Z(\theta_1,\theta_2)}{\partial q_i}$

Circumstance	$\frac{\partial Z}{\partial q_1}$	$\frac{\partial Z}{\partial q_2}$	$\frac{\partial Z}{\partial q_0}$	$Z(\theta_1,\theta_2\|q_1,q_2,q_0)$
a) $\theta_1\leq q_1$, $\theta_2\leq q_2$, $\theta_1+\theta_2\leq q_0$	$-c_1$	$-c_2$	$-c_0$	$p_1\theta_1+p_2\theta_2$ $-c_1q_1-c_2q_2-c_0q_0$
b) $\theta_1\leq q_1$, $\theta_2>q_2$, $\theta_1+q_2\leq q_0$	$-c_1$	p_2-c_2	$-c_0$	$p_1\theta_1+p_2q_2$ $-c_1q_1-c_2q_2-c_0q_0$
c) $\theta_1>q_1$, $\theta_2\leq q_2$, $q_1+\theta_2\leq q_0$	p_1-c_1	$-c_2$	$-c_0$	$p_1q_1+p_2\theta_2$ $-c_1q_1-c_2q_2-c_0q_0$
d) Complementary to cases 1) to 3)	$p_1-p_2-c_1$	$-c_2$	p_2-c_0	$p_1q_1+p_2(q_0-q_1)$ $-c_1q_1-c_2q_2-c_0q_0$

The available information includes: i) $\mu_j=E\left(\tilde{\theta}_j\right)$, which is the mean of the demand for product j, ii) $\sigma_j^2=Var\left(\tilde{\theta}_j\right)$, which is the variance of the demand for product j, and iii) $\rho\in(-1,1)$, which is the correlation coefficient of $\tilde{\theta}_1$ and $\tilde{\theta}_2$. When demands are perfectly correlated (i.e., with $\rho=-1$

Figure 2 Four Circumstances Related to Table 1



or $\rho = +1$), the analysis is reduced to Scarf's model. Thus, we focus on the non-trivial case with $\rho^2 < 1$. The dual model of Adverse Nature is the following:

$$
\begin{aligned}
D = \inf_{F} \max_{q_i \geq 0} & \left\{ \int_0^\infty \int_0^\infty Z\left(\theta_1, \theta_2 | q_1, q_2, q_0\right) \right\} \\
\text{s.t.} \quad & \int_0^\infty \int_0^\infty dF\left(\theta_1, \theta_2\right) = 1 \\
& \int_0^\infty \int_0^\infty \theta_j dF\left(\theta_1, \theta_2\right) = \mu_j \\
& \int_0^\infty \int_0^\infty \theta_j^2 dF\left(\theta_1, \theta_2\right) = \mu_j^2 + \sigma_j^2 \\
& \int_0^\infty \int_0^\infty \theta_1 \theta_2 dF\left(\theta_1, \theta_2\right) = \rho \sigma_1 \sigma_2 + \mu_1 \mu_2 \\
& \int_0^\infty \int_0^\infty \frac{\partial Z\left(\theta_1, \theta_2 | q_1, q_2, q_0\right)}{\partial q_i} dF\left(\theta_1, \theta_2\right) = 0,
\end{aligned}
\tag{5.1}
$$

We formulate the objective function of the corresponding SIP model as follows:

$$
P_1 = \max \left\{ y_0 + y_{11}\mu_1 + y_{12}\mu_2 + y_{21}\left(\mu_1^2 + \sigma_1^2\right) + y_{22}\left(\mu_2^2 + \sigma_2^2\right) + y_3\left(\rho\sigma_1\sigma_2 + \mu_1\mu_2\right) \right\},
$$

where the decision variables include y_0, y_{11}, y_{12}, y_{21}, y_{22}, y_3, a_i, and q_i. When circumstance a) occurs, the SIP constraints are the following:

$$
y_0 + y_{11}\theta_1 + y_{12}\theta_2 + y_{21}\theta_1^2 + y_{22}\theta_2^2 + y_3\theta_1\theta_2 - a_0c_0 - a_1c_1 - a_2c_2 \leq p_1\theta_1 + p_2\theta_2 - c_1q_1 - c_2q_2 - c_0q_0.
$$

When circumstance a) occurs, it gives rise to the following SIP constraints:

$$
y_0 + y_{11}\theta_1 + y_{12}\theta_2 + y_{21}\theta_1^2 + y_{22}\theta_2^2 + y_3\theta_1\theta_2 - a_0c_0 - a_1c_1 + a_2\left(p_2 - c_2\right) \leq p_1\theta_1 + p_2q_2 - c_1q_1 - c_2q_2 - c_0q_0.
$$

When circumstance c) occurs, it gives rise to the following SIP constraints:

$$y_0+y_{11}\theta_1+y_{12}\theta_2+y_{21}\theta_1^2+y_{22}\theta_2^2+y_3\theta_1\theta_2-a_0c_0+a_1\left(p_1-c_1\right)-a_2c_2\leq p_1q_1+p_2\theta_2-c_1q_1-c_2q_2-c_0q_0.$$

When circumstance d) occurs, it gives rise to the following SIP constraints:

$$\begin{aligned}&y_0+y_{11}\theta_1+y_{12}\theta_2+y_{21}\theta_1^2+y_{22}\theta_2^2+y_3\theta_1\theta_2+a_0\left(p-c_0\right)+a_1\left(p_1-p_2-c_1\right)-a_2c_2\\ \leq\ & p_1q_1+p_2\left(q_0-q_1\right)-c_1q_1-c_2q_2-c_0q_0.\end{aligned}$$

Because there exist four different vectors of marginal payoffs, the number of binding constraints is four. The binding constraints are at point $(\theta_{1a},\theta_{2a})$, point $(\theta_{1b},\theta_{2b})$, point $(\theta_{1c},\theta_{2c})$, and point $(\theta_{1d},\theta_{2d})$, where point $(\theta_{1a},\theta_{2a})$ satisfies circumstance a), point $(\theta_{1b},\theta_{2b})$ satisfies circumstance b), point $(\theta_{1c},\theta_{2c})$ satisfies circumstance c), and point $(\theta_{1d},\theta_{2d})$ satisfies circumstance d) of Table 1. The four binding constraints yield eight tangent conditions (because tangent condition is also two-dimensional). Excluding a_i, the remaining unknown variables include 4 pairs of (θ_1,θ_2), 3 inventory levels, and 6 shadow prices. In summary, we encounter 17 decision variables (excluding a_i).

Let L be the Lagrangian based on four binding constraints. The FOCs with respect to a_i include:

$$\begin{aligned}\frac{\partial L}{\partial a_1}&=-c_1\left(\lambda_1+\lambda_3+\lambda_4\right)+\left(p_1-p_2\right)\lambda_4+\left(p_1-c_1\right)\lambda_2=0,\\ \frac{\partial L}{\partial a_2}&=-c_2\left(\lambda_1+\lambda_2+\lambda_4\right)+\left(p_2-c_2\right)\lambda_3=0,\\ \frac{\partial L}{\partial a_0}&=-c_0\left(\lambda_1+\lambda_2+\lambda_3\right)+\left(p_2-c_0\right)\lambda_4=0.\end{aligned}$$

Using the total probability ($\frac{\partial L}{\partial y_0}=1-\lambda_1-\lambda_2-\lambda_3-\lambda_4=0$) and the above 3 equations, we can immediately obtain that $\lambda_1=1-\frac{c_1+c_0}{p_1}-\frac{c_2}{p_2}$, $\lambda_2=\frac{c_1+c_0}{p_1}-\frac{c_0}{p_2}$, $\lambda_3=\frac{c_2}{p_2}$, and $\lambda_4=\frac{c_0}{p_2}$. These four Lagrangian multipliers determine the probability masses of Adverse Nature's equilibrium strategy F^*. It can be verified that there exist 17 equations (i.e., 5 moment conditions because the total probability is already used, 4 binding constraints and 8 tangent conditions). Solving this non-linear system with 17 unknowns and 17 equations in closed forms is possible (see the discussions in the next subsection) but the analytical solution is overly complex.

5.1. Symmetric Case

We consider a special case with the following symmetric data: i) $c_1=c_2=c$ while c_0 may not be identical to c; ii) $p_1=p_2=p$, $\mu_1=\mu_2=\mu$, and $\sigma_1=\sigma_2=\sigma$. Due to symmetry, $q_1=q_2=q$ holds and the binding constraints are at point (x_1,x_1), point (x_2,x_3), point (x_3,x_2), and point (x_4,x_4), where $x_1<q$, $x_2<q_0-q$, $x_3>q$, and $x_4>q$. We obtain the following probability masses: $\lambda_1=1-\frac{2c+c_0}{p}$,

$\lambda_2 = \frac{c}{p}$, $\lambda_3 = \frac{c}{p}$, and $\lambda_4 = \frac{c_0}{p}$. Symmetry also makes $y_{11} = y_{12} = y_1$ and $y_{21} = y_{22} = y_2$, where y_1 and y_2 are to be determined. We obtain the following symmetric tangent conditions:

$$\begin{cases} y_1 + 2y_2 x_1 + y_3 x_1 = p, \\ y_1 + 2y_2 x_2 + y_3 x_3 = p, \\ y_1 + 2y_2 x_3 + y_3 x_2 = 0, \\ y_1 + 2y_2 x_4 + y_3 x_4 = 0. \end{cases}$$

We regard y_1, y_2, and y_3 as input parameters and x_i as unknown variables. We obtain that

$$\begin{aligned} x_1 &= \frac{p - y_1}{2y_2 + y_3}, \ x_2 = \frac{2(y_1 - p)y_2 - y_1 y_3}{y_3^2 - 4y_2^2}, \\ x_3 &= \frac{(p - y_1)\, y_3 + 2y_1 y_2}{y_3^2 - 4y_2^2}, \text{ and } x_4 = \frac{-y_1}{2y_2 + y_3}. \end{aligned}$$

The above equations imply that

$$x_1 + x_4 = x_2 + x_3 = \frac{p - 2y_1}{2y_2 + y_3}.$$

The effective moment conditions include:

$$\begin{cases} \mu = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + \lambda_4 x_4, \\ \mu^2 + \sigma^2 = \lambda_1 (x_1)^2 + \lambda_2 (x_2)^2 + \lambda_3 (x_3)^2 + \lambda_4 (x_4)^2, \\ \rho\sigma^2 + \mu^2 = \lambda_1 (x_1)^2 + \lambda_2 (x_2 x_3) + \lambda_3 (x_3 x_2) + \lambda_4 (x_4)^2. \end{cases}$$

Combining these three effective moment conditions with the symmetric property $x_1 + x_4 = x_2 + x_3$, we obtain the following results:

$$\begin{aligned} x_1^* &= \mu - (c + c_0)\, \sigma \sqrt{\frac{(1+\rho)}{A}}, \\ x_2^* &= \mu + \left(\frac{p}{2} - c - c_0\right) \sigma \sqrt{\frac{(1+\rho)}{A}} - \frac{\sigma}{2} \sqrt{\frac{(1-\rho)\, p}{c}}, \\ x_3^* &= \mu + \left(\frac{p}{2} - c - c_0\right) \sigma \sqrt{\frac{(1+\rho)}{A}} + \frac{\sigma}{2} \sqrt{\frac{(1-\rho)\, p}{c}}, \\ x_4^* &= \mu + (p - c - c_0)\, \sigma \sqrt{\frac{(1+\rho)}{A}}, \end{aligned}$$

where $A = p\,(c + 2c_0) - 2\,(c + c_0)^2$. To ensure that the firm's robust inventory levels are non-zero, the values of A, x_1^* and x_2^* must be positive.

Lemma 4 *With symmetric data and $|\rho| < 1$, in the zero-sum game formulated in equation (5.1), the equilibrium strategy of Adverse Nature is the following four-point distribution:*

$$\left(\tilde{\theta}_1, \tilde{\theta}_2\right) = \begin{cases} (x_1^*, x_1^*) \textit{ with probability } 1 - \frac{2c + c_0}{p}, \\ (x_2^*, x_3^*) \textit{ with probability } \frac{c}{p}, \\ (x_3^*, x_2^*) \textit{ with probability } \frac{c}{p}, \\ (x_4^*, x_4^*) \textit{ with probability } \frac{c_0}{p}. \end{cases}$$

If parameters are asymmetric, we apply the eight tangent conditions to determine four pairs of coordinates by using the shadow prices as input parameters. Five shadow prices $(y_{11}, y_{12}, y_{21}, y_{22}, y_3)$ are involved but five moment conditions remain. We can solve the nonlinear system to obtain the shadow prices, which in turn determine the coordinates of the four binding points. This step could be too complex to be analytically tractable. However, we can easily obtain numerical solutions by using commercial software such as EXCEL. The symmetric special case gives rise to the symmetric property $x_1 + x_4 = x_2 + x_3$, enabling us to explicitly characterize the equilibrium strategy of Adverse Nature in Lemma 4.

A relevant benchmark is the case without component commonality. The firm manufacturers product j by incurring a total cost of $c + c_0$ per unit using an integrated component. Scarf's rule implies that the value of the zero-sum game without component commonality equals $Z_{nc} = 2\mu(p - c - c_0) - 2\sigma\sqrt{(p - c - c_0)(c + c_0)}$. Next, we derive the firm's equilibrium strategy and compute the value of the zero-sum game with common component.

Proposition 5 *With symmetric data and $|\rho| < 1$, the robust optimal inventory level of the product-specific component is*

$$q^* = \frac{x_1^* + x_3^*}{2} = \mu + \sigma\left[\left(\frac{p}{4} - c - c_0\right)\sqrt{\frac{1+\rho}{A}} + \frac{1}{4}\sqrt{\frac{(1-\rho)p}{c}}\right]. \tag{5.2}$$

and that of the common component is

$$q_0^* = x_1^* + x_4^* = 2\mu + \sigma(p - 2c - 2c_0)\sqrt{\frac{1+\rho}{A}}. \tag{5.3}$$

The value of the zero-sum game equals

$$Z^* = 2(p - c - c_0)\mu - \left[\sqrt{(1+\rho)A} + \sqrt{(1-\rho)pc}\right]\sigma. \tag{5.4}$$

It holds that $Z^* \geq Z_{nc}$, where the equal sign holds if $\sigma = 0$.

Proposition 5 provides several useful insights. First, we define $(p - 2c - 2c_0)\sqrt{\frac{1+\rho}{A}}$ in equation (5.3) as the safety stock factor for the common component, which is increasing in the correlation coefficient ρ. Second, equation (5.2) reveals that correlated demands create two opposite effects on the inventory levels of product-specific components. Specifically, the term $\left(\frac{p}{4} - c - c_0\right)\sqrt{\frac{1+\rho}{A}}$ in equation (5.2) captures the increasing effect. When demands are positively correlated, the inventory levels of the common and product-specific components increase, producing the increasing effect. However, the firm also benefits from component commonality and the term $\frac{1}{4}\sqrt{\frac{(1-\rho)p}{c}}$ in equation

(5.2) captures the decreasing effect. The two opposite effects also influence the value of the zero-sum game. In equation (5.4), the term $\sqrt{(1+\rho)A}$ quantifies the impact of higher inventory levels of product-specific components (i.e., the increasing effect); while the term $\sqrt{(1-\rho)pc}$ quantifies the impact of lower inventory level of common component (i.e., the decreasing effect). Third, component commonality always improves the firm's equilibrium profit unless demands are deterministic or perfect positive correlation occurs. When demands are deterministic, the symmetric profit margin of each product equals $(p-c-c_0)$ and the firm's profit equals $2(p-c-c_0)\mu$, which is the first term in equation (5.4). Finally, Proposition 5 underscores that the robust inventory levels closely relate to the binding constraints.

6. Discussions

6.1. Existence of an Equilibrium

Zero-sum games have been extensively studied by various disciplines (e.g., economics, computer science, and operations research) since Nash (1951) proves that an equilibrium exists in finite games. By definition, a finite game restricts the action space of each player to be finite and discrete. The well-known Debreu-Glicksberg-Fan (DGF) Theorem (see Fudenberg and Tirole 1991) extends the result to infinite games, where players can take an infinite number of actions. Specifically, the existence of an equilibrium in an infinite game requires three conditions: i) the action space of each player i is compact and convex; ii) the payoff function of player i is continuous with respect to other players' actions; and iii) the payoff function of player i is continuous and quasi-concave with respect to his own action.

We can extend the analysis to the case with $n \geq 1$ moments. Example 2.6 of Hettich and Kortanek (1993) indicates that we can first use a generalized finite sequence to construct the so-called "moment cones". The convex hull of the moment cones then specifies the corresponding ambiguity set Ω_n (where the subscript n indicates that the first $n \geq 1$ moments are known). It is well-known that the convex hull of the moment cones is convex and compact so that condition i) holds in our examples. In the examples that we consider, the ex post payoff functions are continuous and quasi-concave and hence, conditions ii) and iii) also hold. A minor issue is that we do require the ex post payoff function $Z(\theta|\mathbf{Q})$ and its first derivative $\frac{\partial Z(\theta|\mathbf{Q})}{\partial Q_i}$ both be finite for any $\theta \geq 0$ so that duality holds. We believe that the assumption of a bounded payoff and a bounded marginal payoff is mild, especially in supply chains.

Discontinuous payoff functions can arise in various circumstances, for example, when the (s, S) inventory policy is used. On page 1033, Reny (1999) indicates that many discontinuous games can still have an equilibrium (we refer readers to Reny 2020, for an updated literature review). A promising path for future research is to consider discontinuous payoff functions. Notably, Theorem 1 provides an alternative method to determine whether an equilibrium exists when payoff functions are discontinuous. Specifically, we solve the SIP model of Adverse Nature. If one of the optimal solutions satisfies $a_i = 0$ for all i, then an equilibrium exists; otherwise, the original zero-sum game does not have an equilibrium.

If quasi-concavity is not assumed, we either retrospectively verify the firm's second-order conditions or include them as part of the IC constraints in Adverse Nature's model. To illustrate, we use a single-action model as an example. The FOC includes $\int_0^\infty \frac{\partial Z(\theta|Q)}{\partial Q} dF(\theta) = 0$ and the second-order condition includes $\int_0^\infty \frac{\partial^2 Z(\theta|Q)}{\partial Q^2} dF(\theta) \leq 0$. Consequently, the SIP constraints in equation (2.6) change to the following:

$$y_0 + y_1\theta + y_2\theta^2 + a\frac{\partial Z(\theta|Q)}{\partial Q} + b\frac{\partial^2 Z(\theta|Q)}{\partial Q^2} \leq Z(\theta|Q), \forall \theta \geq 0,$$

where b is nonpositive. If an equilibrium exists, both a^* and b^* are zero.[5] Nonetheless, quasi-concavity not only validates the first-order approach that economists advocate but also ensures the existence of an equilibrium in the zero-sum games that we study.

6.2. Other Ambiguity Sets

The available information affects the ambiguity set and critically determines the mathematical properties of the robust solution. For example, Mulvey et al. (1995) propose a scenario-based model. Ben-Tal and Nemirovski (1999) propose an ellipsoid model to model contaminated data or parameter uncertainty. Bertsimas and Sim (2004) propose a cardinality-constrained model to control the level of conservatism. Esfahani and Kuhn (2018) consider Wasserstein balls. To extend our method to these models, the key step is to develop the IC constraints from the perspective of Adverse Nature. If the relevant IC constraints have a simple form, we believe that our method remains effective.

On the other hand, economists have considered several types of ambiguities, including unknown actions, unknown prior distributions, non-Bayesian beliefs, unknown strategic behaviors, and unknown interactions among agents (we refer readers to a recent survey done by Carroll 2019). In supply chains, an unknown prior distribution is the most relevant type of ambiguity for random demands or random yields. Again, the key step to apply our method is to simplify the IC constraints in the min-max version of the model. By following Perakis and Roels (2008), we can easily incorporate other information such as mode and range into the analysis. In future research, we plan to examine the ambiguities that economists have considered.

7. Concluding Remark

This paper proposes a new and efficient method to solve a large class of zero-sum games under moment conditions. By solving this class of zero-sum games, we solve the corresponding max-min optimization models, which have abundant theoretical and practical applications. Our method is based on the min-max inequality and reformulates the zero-sum game as a robust moral hazard model from the perspective of Adverse Nature. We show that the IC constraints of the moral hazard model become moment constraints. While the marginal payoff function (i.e., the first derivative of the ex post payoff function) appears in the SIP constraints, the number of corner points to be considered declines drastically. The key advantage of our method is that we can solve the robust solution without explicitly deriving the objective function. For example, both the $(1+t)$ model and the n-option model presented in this paper have an overly complex objective function. However, using the property of zero Lagrangian multipliers, we conveniently determine many equilibriums that the traditional method is unable to derive in closed forms.

Endnotes

1. According to a global banking survey conducted by the International Finance Corporation (see page 9 of IFC 2019), an overwhelming majority of small and medium-sized enterprises do not have formal record-keeping processes.
2. In September 2000, Sony (which also used option contracts) announced that it failed to meet customer demand for the new PlayStation console due to shortages in capacitors, LCDs, and flash memory chips (Fu et al. 2010).
3. Instead of lost sales, future research could consider other scenarios. For instance, the firm could use the spot market or persuade customers to buy alternative products. These extensions will enhance the application of our model.
4. We emphasize that the right hand side (RHS) of the IC constraints can be non-zero. For example, the IC constraint for Scarf's model is $\int_0^q (-c)dF(\theta) + \int_q^\infty (p-c)dF(\theta) = 0$, which is equivalent to $\int_0^q dF(\theta) = \frac{p-c}{p}$. Using either version to formulate the SIP model must produce the same solution. In the n-option model, the second version of the IC constraints to formulate equation (4.6) is more convenient. To ensure that the solution in model P_1 is feasible in model P, the Lagrangian multipliers of the IC constraints must be zero despite the non-zero RHS of the IC constraints.
5. With multiple actions, where $\mathbf{Q} = (Q_1, Q_2, ..., Q_n)$, the second-order conditions are complex. In the special case where the cross derivative $\frac{\partial^2 Z(\theta|\mathbf{Q})}{\partial Q_i \partial Q_j}$ is zero for any $i \neq j$, the second-order conditions create a summation of $b_i \frac{\partial^2 Z(\theta|\mathbf{Q})}{\partial Q_i^2}$ at the left-hand-side of the SIP constraints.

Acknowledgments

We thank the seminar audience at The University of Sydney. Please address all correspondence to Dr. Qi Fu at gracefu@um.edu.mo. We thank the language editing service provided by American Journal Experts (`www.aje.com`)

.

References

Baker, K. R., Magazine, M. J., and Nuttle, H. L. W. (1986). The effect of commonality on safety stock in a simple inventory model. *Management Science*, 32(8):982–988.

Baron, D. P. and Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica*, 50(4):911–930.

Ben-Tal, A. and Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, 25(1):1–13.

Bertsimas, D. and Sim, M. (2004). The price of robustness. *Operations Research*, 52(1):35–53.

Bhatia, R. and Davis, C. (2000). A better bound on the variance. *American Mathematical Monthly*, 107(4):353–357.

Bimpikis, K. and Markakis, M. G. (2016). Inventory pooling under heavy-tailed demand. *Management Science*, 62(6):1800–1813.

Carrasco, V., Luz, V. F., Kos, N., Messner, M., Monteiro, P., and Moreira, H. (2018). Optimal selling mechanisms under moment conditions. *Journal of Economic Theory*, 177:245–279.

Carrasco, V., Luz, V. F., Monteiro, P. K., and Moreira, H. (2019). Robust mechanisms: the curvature case. *Economic Theory*, 68:203–222.

Carroll, G. (2015). Robustness and linear contracts. *American Economics Review*, 105(2):536–563.

Carroll, G. (2019). Robustness in mechanism design and contracting. *Annual Review of Economics*, 11:139–166.

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

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

Esfahani, M. P. and Kuhn, D. (2018). Data-driven distributionally robust optimization using the wasserstein metric: performance guarantees and tractable reformulations. *Mathematical Programming*, 171(1-2):115–166.

Fu, Q., Lee, C.-Y., and Teo, C.-P. (2010). Procurement management using option contracts: random spot price and the portfolio effect. *IIE Transactions*, 42(11):793–811.

Fu, Q., Sim, C.-K., and Teo, C.-P. (2018). Profit sharing agreements in decentralized supply chains: A distributionally robust approach. *Operations Research*, 66(2):500–513.

Fudenberg, D. and Tirole, J. (1991). *Game Theory.* The MIT Press, Cambridge, Massachusetts.

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

Govindarajan, A., Sinha, A., and Uichanco, J. (2021). Distribution-free inventory risk pooling in a multilocation newsvendor. *Management Science*, 67(4):2272–2291.

Guo, J., He, S., Jiang, B., and Wang, Z. (2022). A unified framework for generalized moment problems: a novel primal-dual approach. Technical report, `arXiv:2201.01445`, dated 11 January.

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

He, R. and Lu, Y. (2021). A robust price-setting newsvendor problem. *Production and Operations Management*, 30(1):276–292.

Hettich, R. and Kortanek, K. O. (1993). Semi-infinite programming: Theory, methods, and applications. *SIAM Review*, 35(3):380–429.

IFC (2019). Banking on SMEs: Trends and challenges. `https://www.ifc.org/wps/wcm/connect/dd06b824-c38b-4933-9108-0c834f182fee/IFC+on+Banking+SMEs+Publication+June+2019.pdf?MOD=AJPERES&CVID=mSdrGtA`.

Kerr, W. and Ryan, C. (2001). Eco-efficiency gains from remanufacturing: A case study of photocopier remanufacturing at fuji xerox australia. *Journal of Cleaner Production*, 9(1):75–81.

Koçyiğit, Ç., Iyengar, G., Kuhn, D., and Wiesemann, W. (2020). Distributionally robust mechanism design. *Management Science*, 66(1):159–189.

Kouvelis, P., Xiao, G., and Yang, N. (2021). Role of risk aversion in price postponement under supply random yield. *Management Science*, 67(8):4826–4844.

Li, Z. and Kirshner, S. N. (2021). Salesforce compensation and two-sided ambiguity: Robust moral hazard with moment information. *Production and Operations Management*, 30(9):2944–2961.

Lilien, G. L., Rangaswamy, A., and Bruyn, A. D. (2017). *Principles of Marketing Engineering and Analytics, Third Edition.* DecisionPro, State College, Pennsylvania, USA.

Lu, M. and Shen, Z.-J. M. (2021). A review of robust operations management under model uncertainty. *Production and Operations Management*, 30(6):1927–1943.

Martínez-de Albéniz, V. and Simchi-Levi, D. (2009). Competition in the supply option market. *Operations Research*, 57(5):1082–1097.

Mulvey, J. M., Vanderbei, R. J., and Zenios, S. A. (1995). Robust optimization of large-scale systems. *Operations Research*, 43(2):264–281.

Nash, J. (1951). Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295.

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

Natarajan, K. and Zhou, L. (2007). A mean–variance bound for a three-piece linear function. *Probability in the Engineering and Informational Sciences*, 21(4):611–621.

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

Pınar, M. Ç. and Kızılkale, C. (2017). Robust screening under ambiguity. *Mathematical Programming*, 163(1-2):273–299.

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

Reny, P. J. (2020). Nash equilibrium in discontinuous games. *Annual Review of Economics*, 12(1):439–470.

Scarf, H. (1958). A min-max solution of an inventory problem. *Studies in the Mathematical Theory of Inventory and Production*, 10(2):201–209.

Sion, M. (1958). On general minimax theorems. *Pacific Journal of Mathemati*, 8(1):171–176.

Tang, C. S. (2006). Perspectives in supply chain risk management. *International Journal of Production Economics*, 103(2):451–488.

Veinott, J. A. F. (1965). The optimal inventory policy for batch ordering. *Operations Research*, 13(2):424–432.

Yang, C., Hu, Z., and Zhou, S. X. (2021). Multilocation newsvendor problem: Centralization and inventory pooling. *Management Science*, 67(1):185–200.

Yang, L., Cai, G. G., and Chen, J. (2018). Push, pull, and supply chain risk-averse attitude. *Production and Operations Management*, 27(8):1534–1552.

Yue, J., Chen, B., and Wang, M.-C. (2006). Expected value of distribution information for the newsvendor problem. *Operations Research*, 54(6):1128–1136.

Online Appendix: Technical Proofs

Part A: Proofs for Sections 2 and 3

Proof of Theorem 1

The first step is to establish strong duality (i.e., $P = D$ and $P_1 = D_1$). This step is rather routine and we omit the details. The second step has been fully explained in the paragraph preceding Theorem 1. Q.E.D.

Proof of Lemma 1

a) We consider the following nonlinear system:

$$\begin{cases} \frac{p-r-c}{p-r}x_1 + \frac{c}{p-r}x_2 = \mu, \\ \frac{p-r-c}{p-r}x_1^t + \frac{c}{p-r}x_2^t = m_t. \end{cases}, \tag{A-1}$$

which has two unknowns and two equations with the power of t. Using the equation $\frac{c}{p-r}x_2 = \mu - \frac{p-r-c}{p-r}x_1$, we can rewrite equation (A-1) as the following one-variable equation:

$$H\left(x_1\right) = \frac{p-r-c}{p-r}x_1^t + \frac{c}{p-r}\left[x_1 + \frac{p-r}{c}\left(\mu - x_1\right)\right]^t = m_t.$$

It is easy to verify that for $x_1 \in [0, \mu]$, $H\left(x_1\right)$ is decreasing. When $x_1 = 0$, $H\left(x_1\right) = \frac{\mu^t}{\left(\frac{c}{p-r}\right)^{t-1}}$ and when $x_1 = \mu$, $H\left(\mu\right) = \mu^t$. Hence, when condition (3.4) holds, $H\left(x_1\right) = m_t$ has a nonnegative solution θ_1 satisfying $0 \leq \theta_1 \leq \mu$. We then obtain that

$$\theta_2 = \theta_1 + \left(\frac{p-r}{c}\right)\left(\mu - \theta_1\right) \geq \theta_1.$$

Because (θ_1, θ_2) satisfy the nonlinear system in (A-1), the two-point distribution in equation (3.5) satisfies the conditions on the mean and t-th moment.

b) It is readily verified that the distribution in equation (3.7) satisfies the conditions on the mean and t-th moment. Q.E.D.

Proof of Proposition 1

a) We conjecture that binding constraints occur at θ_1 and θ_2, where $\theta_1 \leq q \leq \theta_2$ such that overstock (lost sales) occurs when the realized demand is θ_1 (θ_2). If the conjectured constraints satisfy $\theta_1 \leq \theta_2 \leq q$, then the firm can lower the inventory level to avoid overstocking. If the conjectured constraints satisfy $q \leq \theta_1 \leq \theta_2$, then the firm can increase the inventory level to reduce lost sales. The Lagrangian equals

$$\begin{aligned} L = {} & y_0 + y_1\mu + y_2 m_t - \lambda_1\left[y_0 + y_1\theta_1 + y_2\theta_1^t - ac - (p-r)\,\theta_1 + cq\right] \\ & - \lambda_2\left[y_0 + y_1\theta_2 + y_2\theta_2^t + a\left(p-r-c\right) - \left(p-r-c\right)q\right]. \end{aligned}$$

The FOCs with respect to y_0 and a are the following:

$$\frac{\partial L}{\partial y_0} = 1 - \lambda_1 - \lambda_2 = 0 \text{ and } \frac{\partial L}{\partial a} = \lambda_1 c - \lambda_2 (p - r - c) = 0.$$

These two equations yield that $\lambda_1 = \frac{p-r-c}{p-r}$ and $\lambda_2 = \frac{c}{p-r}$. Observe that the multiplier $\lambda_1 = \frac{p-r-c}{p-r}$ corresponds to the newsvendor ratio. The FOCs with respect to y_1 and y_2 are the following:

$$\frac{\partial L}{\partial y_1} = \mu - \lambda_1 \theta_1 - \lambda_2 \theta_2 = 0 \text{ and } \frac{\partial L}{\partial y_2} = m_t - \lambda_1 \theta_1^t - \lambda_2 \theta_2^t = 0.$$

Because $\lambda_1 = \frac{p-r-c}{p-r}$ and $\lambda_2 = \frac{c}{p-r}$ are known, the above two equations yield the following nonlinear system:

$$\begin{cases} \frac{p-r-c}{p-r}\theta_1 + \frac{c}{p-r}\theta_2 = \mu, \\ \frac{p-r-c}{p-r}\theta_1^t + \frac{c}{p-r}\theta_2^t = m_t. \end{cases}, \tag{A-2}$$

which involves two equations and two unknowns. Lemma 1a) solves this nonlinear system and identifies the two-point distribution in equation (3.5). We label this two-point distribution as F^*, which is independent of q and is the strategy that Adverse Nature plays in the equilibrium.

Without the shadow prices, the analysis remains incomplete. To solve for the shadow prices, we use the binding constraints and the tangent conditions as follows:

$$y_0 + y_1\theta_1 + y_2\theta_1^t - ac - (p-r)\theta_1 + cq = 0,$$
$$y_0 + y_1\theta_2 + y_2\theta_2^t + a(p-r-c) - (p-r)q + cq = 0,$$
$$y_1 + ty_2\theta_1^{t-1} = p - r \text{ and } y_1 + ty_2\theta_2^{t-1} = 0.$$

We emphasize that in model P_1, the tangent condition satisfies that $y_1 + ty_2\theta^{t-1} + a\frac{\partial^2 Z(\theta|q)}{\partial q \partial \theta} = \frac{\partial Z(\theta|q)}{\partial \theta}$ while in (the traditional) model P, the tangent condition satisfies that $y_1 + ty_2\theta^{t-1} = \frac{\partial Z(\theta|q)}{\partial \theta}$. Due to the characteristics of the newsvendor model (in which the ex post profit function is piece-wise linear), the cross derivative $\frac{\partial^2 Z(\theta|q)}{\partial q \partial \theta}$ is zero.

Solving this system of 4 unknowns and 4 equations, we obtain that

$$a = q - \theta_1 - \frac{\theta_2^{t-1}(\theta_2 - \theta_1)}{\theta_2^{t-1} - \theta_1^{t-1}} + \frac{\theta_2^t - \theta_1^t}{t\left(\theta_2^{t-1} - \theta_1^{t-1}\right)}.$$

According to Theorem 1, we let $a = 0$ to obtain the firm's equilibrium strategy q^*, which is shown in equation (3.8).

The shadow prices y_1 and y_2 are the following:

$$y_1 = \theta_2^{t-1}\left(\frac{p-r}{\theta_2 - \theta_1}\right) \text{ and } y_2 = -\frac{p-r}{t\left(\theta_2^{t-1} - \theta_1^{t-1}\right)},$$

and the fourth shadow price y_0 equals

$$\begin{aligned}
y_0 &= -y_1\theta_1 - y_2\theta_1^t + ac + (p-r)\,\theta_1 - cq \\
&= -\theta_2^{t-1}\theta_1\left(\frac{p-r}{\theta_2-\theta_1}\right) + \frac{p-r}{t\left(\theta_2^{t-1}-\theta_1^{t-1}\right)}\theta_1^t + c\left(q-\theta_1-\frac{\theta_2^{t-1}\left(\theta_2-\theta_1\right)}{\theta_2^{t-1}-\theta_1^{t-1}}+\frac{\theta_2^t-\theta_1^t}{t\left(\theta_2^{t-1}-\theta_1^{t-1}\right)}\right) \\
&\quad + (p-r)\,\theta_1 - cq \\
&= (p-r)\left[\theta_1+\frac{\theta_1^t}{t\left(\theta_2^{t-1}-\theta_1^{t-1}\right)}-\frac{\theta_2^{t-1}\theta_1}{\theta_2-\theta_1}\right] - c\left(\theta_1+\frac{\theta_2^{t-1}\left(\theta_2-\theta_1\right)}{\theta_2^{t-1}-\theta_1^{t-1}}-\frac{\theta_2^t-\theta_1^t}{t\left(\theta_2^{t-1}-\theta_1^{t-1}\right)}\right).
\end{aligned}$$

There are two alternatives to compute P_1^*, which is the value of the zero-sum game. The first alternative is to compute P_1^* from the firm's perspective. When Adverse Nature plays the two-point distribution in equation (3.5), we can easily verify that for any $q \in [\theta_1, \theta_2]$,

$$\begin{aligned}
Z\left(q, F^*\right) &= \Pr(\tilde{\theta}=\theta_1)Z(\theta_1|q) + \Pr(\tilde{\theta}=\theta_2)Z(\theta_2|q) = \frac{p-r-c}{p-r}[(p-r)\theta_1] + \frac{c}{p-r}[(p-r)q] - cq \\
&= (p-r-c)\theta_1 + cq - cq = (p-r-c)\theta_1.
\end{aligned}$$

The second alternative is to substitute the optimal shadow prices that we have developed into the objective function of model P_1. Note that (θ_1, θ_2) satisfy the nonlinear system in equation (A-2). We find that

$$\begin{aligned}
P_1^* &= y_0 + y_1\mu + y_2 m_t = y_0 + y_1\left(\frac{p-r-c}{p-r}\theta_1 + \frac{c}{p-r}\theta_2\right) + y_2\left(\frac{p-r-c}{p-r}\theta_1^t + \frac{c}{p-r}\theta_2^t\right) \\
&= \frac{p-r-c}{p-r}\left(y_0 + y_1\theta_1 + y_2\theta_1^t\right) + \frac{c}{p-r}\left(y_0 + y_1\theta_2 + y_2\theta_2^t\right).
\end{aligned}$$

The binding SIP constraints imply that $y_0 + y_1\theta_1 + y_2\theta_1^t = ac + (p-r)\,\theta_1 - cq$ and $y_0 + y_1\theta_2 + y_2\theta_2^t = -a\,(p-r-c) + (p-r)\,q - cq$. Thus, we obtain that

$$\begin{aligned}
P_1^* &= \frac{p-r-c}{p-r}\left[ac + (p-r)\,\theta_1 - cq\right] + \frac{c}{p-r}\left[-a\,(p-r-c) + (p-r)\,q - cq\right] \\
&= a\left[\frac{(p-r-c)\,c}{p-r} - \frac{c\,(p-r-c)}{p-r}\right] + (p-r-c)\,\theta_1 + cq - cq = (p-r-c)\,\theta_1.
\end{aligned}$$

Both alternatives lead to the same conclusion. Note that the values of θ_1 and θ_2 satisfy Lemma 1a) and condition (3.4) ensures that θ_1 is positive.

b) Suppose that condition (3.6) holds and Adverse Nature chooses the distribution in equation (3.7) as the equilibrium strategy. We can easily verify that the firm's expected profit is decreasing with respect to q. Hence, the firm's best response is $q^* = 0$, which results in zero profit. Because Adverse Nature's objective is to minimize the value of the zero-sum game, $P_1 \geq 0$ must hold. Because the distribution in equation (3.7) makes $P_1 = 0$, we conclude that $q^* = 0$ is the firm's equilibrium strategy and the distribution in equation (3.7) is Adverse Nature's equilibrium strategy. Q.E.D.

Scarf's Model

In a special case where $t=2$ (such that $m_2=\mu^2+\sigma^2$), the analysis reduces to Scarf's model and condition (3.4) requires that $\mu^2 \geq \frac{c}{p-r}\left(\mu^2+\sigma^2\right)$, which is equivalent to $(p-r-c) \geq c\rho^2$. It is well-known that when $(p-r-c)<c\rho^2$, the firm's robust optimal solution is $q^*=0$. Based on Lemma 1, we re-write the nonlinear equation of (3.3) as follows:

$$H(x)=\frac{p-r-c}{p-r}x^2+\frac{c}{p-r}\left[x+\frac{p-r}{c}(\mu-x)\right]^2=\mu^2+\sigma^2,$$

which readily yields that $\theta_1=\mu-\sigma\sqrt{\frac{c}{p-r-c}}$ and $\theta_2=\mu+\sigma\sqrt{\frac{p-r-c}{c}}$. As a result, we obtain the following distribution:

$$\begin{cases} \Pr\left(\tilde{\theta}=\mu-\sigma\sqrt{\frac{c}{p-r-c}} \overset{def}{=} \theta_1^*\right)=\frac{p-r-c}{p-r}, \\ \Pr\left(\tilde{\theta}=\mu+\sigma\sqrt{\frac{p-r-c}{c}} \overset{def}{=} \theta_2^*\right)=\frac{c}{p-r}. \end{cases} \tag{A-3}$$

Corollary 3 *In Scarf's model, when $(p-r-c) \geq c\rho^2$ holds, the equilibrium strategy F^* chosen by Adverse Nature is the two-point distribution shown in equation (A-3). The SIP model in equation (3.2) encompasses multiple solutions satisfying*

$$\begin{cases} a=q-\frac{\theta_1^*+\theta_2^*}{2}, \\ y_0=-\frac{\Delta}{2\sigma}\left(\mu^2+\sigma^2\right), \\ y_1=(p-r-c)+\Delta\frac{\mu}{\sigma}, \\ y_2=-\frac{\Delta}{2\sigma}, \end{cases} \tag{A-4}$$

where $q \in [\theta_1^*, \theta_2^*]$ and $\Delta=\sqrt{(p-r-c)\,c}$ such that the value of the zero-sum game equals

$$P_1^*=(p-r-c)\theta_1^*=(p-r-c)\,\mu-\Delta\sigma.$$

Proof. Let $\Delta=\sqrt{(p-r-c)\,c}$. Based on the proof of Proposition 1, we obtain that

$$a=q-\theta_1-\frac{\theta_2\left(\theta_2-\theta_1\right)}{\theta_2-\theta_1}+\frac{\theta_2^2-\theta_1^2}{2\left(\theta_2-\theta_1\right)}=q-\frac{\theta_1+\theta_2}{2}.$$

On the other hand, the remaining shadow prices include: $y_0=-\frac{\Delta}{2\sigma}\left(\mu^2+\sigma^2\right)$, $y_1=(p-r-c)+\Delta\frac{\mu}{\sigma}$, $y_2=\frac{-(p-r)}{2(\theta_2-\theta_1)}=-\frac{\Delta}{2\sigma}$, confirming the results in equation (A-4). Notice that when q is fixed, the SIP model P_1 is also linear with respect to decision variables y_i and a, making KKT conditions sufficient and necessary. Using Theorem 1, we let $a^*=0$ to obtain $q^*=\frac{\theta_1+\theta_2}{2}$. Q.E.D. ■

From the firm's perspective, if the distribution in equation (A-3) arises, any $q \in [\theta_1^*, \theta_2^*]$ gives her the same expected profit, which equals $P_1^*=(p-r-c)\,\theta_1^*$. Although any $q \in [\theta_1^*, \theta_2^*]$ does not affect the optimal value of model P_1, the firm cannot arbitrarily choose any q from the interval $[\theta_1^*, \theta_2^*]$.

To construct a solution that is feasible for both models P and P_1, Theorem 1 suggests that we let $a=0$ such that

$$q^* = \frac{1}{2}\left(\theta_1^* + \theta_2^*\right) = \mu + \frac{1}{2}\sigma\left(\sqrt{\frac{p-r-c}{c}} - \sqrt{\frac{c}{p-r-c}}\right),$$

which is the well-known Scarf's rule and happens to be the midpoint of the closed interval $[\theta_1^*, \theta_2^*]$. A few useful observations can be made.

- When a nonrobust solution is used, the shadow price a associated with the IC constraint is non-zero. The shadow price a could be positive or negative depending on which direction that the nonrobust q deviates from the equilibrium. Whenever $a \neq 0$, the solution (a, y_i, q), which is feasible for model P_1, cannot be implemented in model P, making $P < P_1$ and q a nonrobust solution.

- In terms of computational complexity, we bypass the middle step of solving the inner SIP model in equation (2.3) and directly attack the robust solution. Because the first derivative is a step function, we encounter multiple solutions in model P_1. We then apply Theorem 1 to solve the robust solution.

- If the firm deviates from the equilibrium, it is well-known that Adverse Nature plays the following strategy:

$$\int_0^q dF\left(\theta|q\right) = \frac{1}{2} + \frac{q-\mu}{2\sqrt{\left(q-\mu\right)^2 + \sigma^2}},$$

which represents a credible threat to the firm such that she has no incentive to deviate from her equilibrium strategy. Interestingly, by letting $q = \frac{1}{2}\left(\theta_1^* + \theta_2^*\right)$ in the above equation, we obtain the two-point distribution equation (A-3), which is labeled as F^*. If distribution F^* realizes, the firm's expected profit is monotonically increasing in q if $q < \theta_1^*$ and is monotonically decreasing in q if $q > \theta_2^*$. We can conclude that the firm never plays a q outside of the closed interval $[\theta_1^*, \theta_2^*]$. In other words, to satisfy the IC constraints in Adverse Nature's model, the capacity level must come from the interval $[\theta_1^*, \theta_2^*]$.

- Let $k = \frac{p-r-c}{p-c}$. In a special case with $k = 0.5$, equation (A-3) yields that $\theta_1^* = \mu - \sigma$ and $\theta_2^* = \mu + \sigma$, which are the theoretical bounds on the median. The bounds on the median are well-known in statistics. However, equation (A-3) generalizes the bounds to any $k \times 100\%$ percentile, where $0 < k < 1$. Specifically, the upper bound on the $k \times 100\%$ percentile is $\mu + \sigma\sqrt{\frac{k}{1-k}}$ and the lower bound is $\mu - \sigma\sqrt{\frac{1-k}{k}}$. In relation to supply chains, if the newsvendor ratio is known to be k, the firm should not order less than $\mu - \sigma\sqrt{\frac{1-k}{k}}$ or order more than $\mu + \sigma\sqrt{\frac{k}{1-k}}$; otherwise, she behaves off the equilibrium. Certainly, using Lemma 1, we can generalize the bounds on percentile by using the mean and the t-th moment.

Part B: Proofs for Section 4

Proof of Lemma 2:

In the fulfillment stage, the capacity vector $\mathbf{q}$ is already chosen. After observing the realized demand θ, the firm solves the following linear programming model to maximize her ex post profit:

$$Z(\theta|\mathbf{q}) = -\sum_{i=1}^{n} c_i q_i + \max_{x_i \geq 0} \left\{ \sum_{i=1}^{n} (p - r_i) x_i \right\}, \tag{B-1}$$

subject to the capacity availability constraints:

$$x_i \leq q_i, \ \forall i \in \{1, 2, \cdots, n\}, \tag{B-2}$$

and the total demand constraint:

$$x_1 + x_2 + \cdots + x_n \leq \theta. \tag{B-3}$$

The nonnegative decision variable x_i represents the quantity of the end products delivered by source-i. The capacity availability constraints (B-2) ensure that the quantity of the end products delivered by source-i does not exceed the available capacity q_i, and the total demand constraint (B-3) ensures that the total delivered quantities do not exceed the realized demand θ. In equation (B-1), the total cost $\sum_{i=1}^{n} c_i q_i$ is sunk, and the coefficient $(p - r_i)$ is decreasing in i due to Definition 1. Hence, the firm must prefer source i to source $(i+1)$, implying that the firm's optimal fulfillment plan follows a priority rule such that source $(i+1)$ is not used unless the reserved capacity of source i is exhausted.

In equation (B-1), the coefficient of decision variable x_i equals $(p - r_i)$, which is decreasing in i. The firm must exhaust all of source-i capacities before using any source-$(i+1)$ capacity. Thus, it holds that

$$x_i^* = \min \left(q_i, \left(\theta - \sum_{j=1}^{i-1} q_i \right)^+ \right),$$

which is optimal for equation (B-1). Using the definition of Q_i, we obtain that

$$x_i^* = \min \left(Q_i - Q_{i-1}, (\theta - Q_{i-1})^+ \right) = \min(Q_i, \theta) - \min(Q_{i-1}, \theta),$$

indicating that the sales quantity contributed by source-i capacity equals the sales of the first i sources minus the sales of the first $(i-1)$ sources. Thus, we obtain that

$$\begin{aligned} Z(\theta|\mathbf{q}) &= Z(\theta|\mathbf{Q}) = -\sum_{i=1}^{n} c_i (Q_i - Q_{i-1}) + \sum_{i=1}^{n} (p - r_i) x_i^* \\ &= \sum_{i=1}^{n} \left[(p - r_i) (\min(Q_i, \theta) - \min(Q_{i-1}, \theta)) - c_i (Q_i - Q_{i-1}) \right], \end{aligned}$$

which yields equation (4.2). It is easy to verify that $Z(\theta|\mathbf{Q})$ is continuous in θ.

To understand the concave and increasing properties of $Z(\theta|\mathbf{Q})$, we expand equation (4.2) into $(n+1)$ cases. i) When $\theta \in [0, Q_1]$:

$$Z(\theta|\mathbf{Q}) = (p - r_1)\,\theta - \sum_{i=1}^{n} c_i q_i, \tag{B-4}$$

in which $\frac{\partial Z(\theta|\mathbf{Q})}{\partial \theta} = p - r_1$ and $\frac{\partial Z(\theta|\mathbf{Q})}{\partial q_1} = -c_1$. ii) When $\theta \in [Q_1, Q_2]$:

$$Z(\theta|\mathbf{Q}) = (p - r_2)(\theta - q_1) + (p - r_1)\,q_1 - \sum_{i=1}^{n} c_i q_i = (p - r_2)\,\theta + (r_2 - r_1 - c_1)\,q_1 - \sum_{j=2}^{n} c_j q_j,$$

in which $\frac{\partial Z(\theta|\mathbf{Q})}{\partial \theta} = p - r_2$ and $\frac{\partial Z(\theta|\mathbf{Q})}{\partial q_1} = r_2 - r_1 - c_1$. iii) In general, when $\theta \in [Q_{i-1}, Q_i]$ and $2 \le i \le n$:

$$Z(\theta|\mathbf{Q}) = (p - r_i)\,\theta + \sum_{j=1}^{i-1} (r_i - r_j - c_j)\,q_j - \sum_{j=i}^{n} c_j q_j, \tag{B-5}$$

in which $\frac{\partial Z(\theta|\mathbf{Q})}{\partial \theta} = p - r_i$ and $\frac{\partial Z(\theta|\mathbf{Q})}{\partial q_1} = r_i - r_1 - c_1$. iv) Finally, when $\theta \ge Q_n$:

$$Z(\theta|\mathbf{Q}) = \sum_{i=1}^{n} (p - r_i - c_i)\,q_i, \tag{B-6}$$

in which $\frac{\partial Z(\theta|\mathbf{Q})}{\partial \theta} = 0$ and $\frac{\partial Z(\theta|\mathbf{Q})}{\partial q_1} = p - r_1 - c_1$.

The first derivative $\frac{\partial Z(\theta|\mathbf{Q})}{\partial \theta}$ is nonnegative and decreasing in θ, making $Z(\theta|\mathbf{Q})$ concave and increasing in θ. The first derivative $\frac{\partial Z(\theta|\mathbf{Q})}{\partial q_1} = \delta_i$ represents the marginal impact of source-i. Q.E.D.

Proof of Lemma 3:

Using equation (4.2), we obtain that the firm's expected profit equals:

$$\begin{aligned}
Z(\mathbf{Q}) &= \sum_{i=1}^{n} (p - r_i)\left[E\min(Q_i, \theta) - E\min(Q_{i-1}, \theta)\right] - c_i\,(Q_i - Q_{i-1}) \\
&= \sum_{i=1}^{n-1} \left\{\left[(p - r_i) - (p - r_{i+1})\right] E\left[\min(Q_i, \theta)\right] - (c_i - c_{i+1})\,Q_i\right\} \\
&\quad + (p - r_n)\,E\min(Q_n, \theta) - c_n Q_n \\
&= \sum_{i=1}^{n-1} \left\{(r_{i+1} - r_i)\,E\left[\min(Q_i, \theta)\right] - (c_i - c_{i+1})\,Q_i\right\} \\
&\quad + (r_{n+1} - r_n)\,E\min(Q_n, \theta) - c_n Q_n.
\end{aligned} \tag{B-7}$$

In equation (B-7), we regard Q_i as the decision variable. Based on the assumptions on the cost parameters, we observe that $(r_{i+1} - r_i) \ge 0$ and $(c_i - c_{i+1}) \ge 0$. The expected sales quantity $E[\min(Q_i, \theta)]$ is continuous and concave with respect to Q_i. Thus, the term $(r_{i+1} - r_i) E[\min(Q_i, \theta)] - (c_i - c_{i+1}) Q_i$ is concave in Q_i. The last term of equation (B-7) is also concave because $(r_{n+1} - r_n) > 0$. We conclude that equation (B-7) is concave in Q_i.

By relaxing the monotonicity constraints $Q_i \geq Q_{i-1}$, we can separately optimize each term in equation (B-7). Observe that $\frac{\partial}{\partial Q_i} E[\min(Q_i, \theta)] = 1 - F(Q_i)$. The FOC yields a candidate solution $\tilde{Q}_i$ satisfying

$$F\left(\tilde{Q}_i\right) = 1 - \frac{c_i - c_{i+1}}{r_{i+1} - r_i} = \alpha_i,$$

for $i = 1, 2, \cdots, n-1$. For $i = n$, the FOC yields that

$$F\left(\tilde{Q}_n\right) = 1 - \frac{c_n - c_{n+1}}{r_{n+1} - r_n} = 1 - \frac{c_n}{p - c_n} = \alpha_n.$$

Because α_i is increasing in i, the candidate solution $\tilde{Q}_i$ satisfies the monotonicity constraints and hence is optimal. We obtain that

$$\tilde{q}_i = F^{-1}\left(\tilde{Q}_i\right) - F^{-1}\left(\tilde{Q}_{i-1}\right) = F^{-1}(\alpha_i) - F^{-1}(\alpha_{i-1}) \geq 0,$$

where F^{-1} is the inverse function of F. Q.E.D.

Proof of Corollary 1:

i) Equation (4.3) indicates that $\delta_i = r_i - r_1 - c_1$ for $i = 1, 2, \cdots n+1$.

$$\sum_{i=1}^{n+1} \beta_i \delta_i = \left(\sum_{i=1}^{n+1} \beta_i r_i\right) - \left(\sum_{i=1}^{n+1} \beta_i (r_1 + c_1)\right) = \left(\sum_{i=1}^{n+1} \beta_i r_i\right) - c_1 - r_1.$$

Observe that

$$\begin{aligned}
\sum_{i=1}^{n+1} \beta_i r_i &= \left(1 - \frac{c_1 - c_2}{r_2 - r_1}\right) r_1 + \left(\frac{c_1 - c_2}{r_2 - r_1} - \frac{c_2 - c_3}{r_3 - r_2}\right) r_2 + \cdots + \left(\frac{c_{n-1} - c_n}{r_n - r_{n-1}} - \frac{c_n}{p - r_n}\right) r_n + \frac{p c_n}{p - r_n} \\
&= r_1 - \frac{c_1 - c_2}{r_2 - r_1} r_1 + \frac{c_1 - c_2}{r_2 - r_1} r_2 - \frac{c_2 - c_3}{r_3 - r_2} r_2 + \cdots + \frac{c_{n-1} - c_n}{r_n - r_{n-1}} r_n - \frac{c_n r_n}{p - r_n} + \frac{p c_n}{p - r_n} \\
&= r_1 + \frac{c_1 - c_2}{r_2 - r_1}(r_2 - r_1) + \cdots + \frac{c_{n-1} - c_n}{r_n - r_{n-1}}(r_n - r_{n-1}) - \frac{c_n r_n}{p - r_n} + \frac{p c_n}{p - r_n} \\
&= r_1 + c_1 - c_2 + \cdots + c_{n-1} - c_n - \frac{c_n r_n}{p - r_n} + \frac{p c_n}{p - r_n} \\
&= r_1 + c_1 - c_2 + \cdots + c_{n-1} - c_n + c_n = r_1 + c_1.
\end{aligned}$$

Thus, we conclude that $\sum_{i=1}^{n+1} \beta_i \delta_i = 0$.

ii) We shall verify that the discrete distribution shown in equation (4.8) satisfies the mean and variance constraints. Equation (4.7) indicates that θ_i^* is a linear transformation of δ_i. Thus, we immediately obtain that

$$\sum_{i=1}^{n+1} \beta_i \theta_i^* = \sum_{i=1}^{n+1} \beta_i \left(\mu + \frac{\sigma}{\Delta} \delta_i\right) = \mu + \frac{\sigma}{\Delta} \sum_{i=1}^{n+1} \beta_i \delta_i = \mu$$

and
$$\sum_{i=1}^{n+1}\beta_i\left(\theta_i^*-\mu\right)^2=\sum_{i=1}^{n+1}\beta_i\left(\frac{\sigma}{\Delta}\delta_i\right)^2=\frac{\sigma^2}{\Delta^2}\sum_{i=1}^{n+1}\beta_i\delta_i^2=\frac{\sigma^2}{\Delta^2}\Delta^2=\sigma^2,$$
where the third equality is based on the definition of Δ. We conclude that the discrete distribution shown in equation (4.8) satisfies the mean and variance constraints, becoming an element of the ambiguity set Ω. Q.E.D.

Proof of Proposition 2:

Step 1): Conjecture binding constraints. We conjecture that the SIP model in equation (4.6) has $(n+1)$ binding constraints at points θ_i (where $i=1,2,\cdots,n+1$) such that $\theta_i \in [Q_{i-1}, Q_i]$. At the first segment with $\theta \in [0, Q_1]$, using equation (B-4), we find that the SIP constraint is
$$a_1+y_0+y_1\theta+y_2\theta^2 \le (p-r_1)\,\theta-\sum_{i=1}^{n}c_iq_i.$$
By assuming that $\theta=\theta_1$ is the binding constraint at the first segment, we obtain
$$a_1+y_0+y_1\theta_1+y_2\theta_1^2=(p-r_1)\,\theta_1-\sum_{i=1}^{n}c_iq_i=Z\left(\theta_1|\mathbf{Q}\right).$$
The above condition alone is insufficient to ensure that θ_1 is a locally binding constraint for the first segment. Similar to in Scarf (1958), the *tangent* condition $y_1+2y_2\theta_1=p-r_1$ must also hold. By repeating the same procedure on equations (B-5) and (B-6), we obtain all of the binding and tangent conditions. We summarize the $(n+1)$ binding conditions as follows:
$$\begin{cases} a_1+y_0+y_1\theta_1+y_2\theta_1^2=Z\left(\theta_1|\mathbf{Q}\right)\\ \vdots\\ a_i+y_0+y_1\theta_i+y_2\theta_i^2=Z\left(\theta_i|\mathbf{Q}\right)\\ \vdots\\ a_n+y_0+y_1\theta_n+y_2\theta_n^2=Z\left(\theta_n|\mathbf{Q}\right)\\ y_0+y_1\theta_{n+1}+y_2\theta_{n+1}^2=Z\left(\theta_{n+1}|\mathbf{Q}\right)\end{cases}\tag{B-8}$$
and $(n+1)$ tangent conditions as follows:
$$\begin{cases} y_1+2y_2\theta_1=p-r_1\\ \vdots\\ y_1+2y_2\theta_i=p-r_i\\ \vdots\\ y_1+2y_2\theta_n=p-r_n\\ y_1+2y_2\theta_{n+1}=p-r_{n+1}=0\end{cases}\tag{B-9}$$
Step 2: Solve the KKT conditions. The Lagrangian of equation (4.6) equals:
$$\begin{aligned}L=\;&a_1\beta_1+a_2\beta_2+...+a_n\beta_n+y_0+y_1\mu+y_2\left(\mu^2+\sigma^2\right)\\ &-\sum_{i=1}^{n}\lambda_i\left[a_i+y_0+y_1\theta_i+y_2\theta_i^2-Z\left(\theta_i|\mathbf{Q}\right)\right]-\lambda_{n+1}\left[y_0+y_1\theta_{n+1}+y_2\theta_{n+1}^2-Z\left(\theta_{n+1}|\mathbf{Q}\right)\right],\end{aligned}$$

where $\lambda_i \geq 0$ is the Lagrangian multiplier.

After solving the FOC with respect to a_i, we obtain that $\frac{\partial L}{\partial a_i} = \beta_i - \lambda_i = 0$, implying that the first n Lagrangian multipliers satisfy that $\lambda_i = \beta_i$ for $i = 1, 2, \cdots, n$. Solving the FOC with respect to y_0, we obtain that $1 = \sum_{i=1}^{n+1} \lambda_i$. Because we have just shown that $\lambda_i = \beta_i$ for $i = 1, 2, \cdots, n$, we obtain that $\lambda_{n+1} = 1 - \sum_{i=1}^{n} \beta_i = \beta_{n+1}$. We conclude that the mass probabilities β_i shown in Definition 1 happen to be the Lagrangian multipliers.

We derive the conjectured binding points θ_i. We solve the FOCs with respect to y_1 and y_2 to obtain

$$\mu = \sum_{i=1}^{n+1} \beta_i \theta_i \text{ and } \mu^2 + \sigma^2 = \sum_{i=1}^{n+1} \beta_i \theta_i^2. \tag{B-10}$$

Equations (B-9) and (B-10) form a system with $(n+3)$ unknown variables (i.e., θ_i, y_1, y_2) and $(n+3)$ equations. To streamline the expressions, let $x_i = \theta_i - \mu$ for $i = 1, 2, \cdots, n+1$. We also define $\{v_i\}$ and $\{h_i\}$ for $i = 1, 2, \cdots, n$ in Table 2 below.

Table 2 The $\{v_i\}$ and $\{h_i\}$ sequences and their relationship to sequences $\{\beta_i\}$ and $\{\alpha_i\}$.

i	v_i	h_i	β_i	α_i
1	$c_1 - c_2$	$r_2 - r_1$	$1 - \frac{v_1}{h_1}$	$1 - \frac{v_1}{h_1}$
2	$c_2 - c_3$	$r_3 - r_2$	$\frac{v_1}{h_1} - \frac{v_2}{h_2}$	$1 - \frac{v_2}{h_2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$n-1$	$c_{n-1} - c_n$	$r_n - r_{n-1}$	$\frac{v_{i-1}}{h_{i-1}} - \frac{v_i}{h_i}$	$1 - \frac{v_{i-1}}{h_{i-1}}$
n	c_n	$p - r_n$	$\frac{v_{n-1}}{h_{n-1}} - \frac{v_n}{h_n}$	$1 - \frac{c_n}{p - r_n}$
$n+1$	N/A	N/A	$\frac{c_n}{p - r_n}$	1

Using the sequences $\{v_i\}$ and $\{h_i\}$ given in Table 2, we rewrite the mean condition shown in equation (B-10) as:

$$\left(1 - \frac{v_1}{h_1}\right) x_1 + \cdots + \left(\frac{v_{i-1}}{h_{i-1}} - \frac{v_i}{h_2}\right) x_i + \cdots + \left(\frac{v_{n-1}}{h_{n-1}} - \frac{v_n}{h_n}\right) x_n + \frac{v_n}{h_n} x_{n+1} = 0.$$

We also simplify the tangent conditions (B-9) to obtain that $2y_2(x_i - x_{i+1}) = h_i$ for $i = 1, 2, \cdots, n$. After substituting the recursive equation $2y_2(x_i - x_{i+1}) = h_i$ into the mean condition and performing some algebra, we obtain that

$$\begin{cases} x_1 = \frac{\sum_{j=1}^{n} v_j}{2y_2} = \frac{c_1}{2y_2}, \\ x_i = \frac{\sum_{k=1}^{i} v_k - \sum_{k=1}^{i-1} h_k}{2y_2} = \frac{c_1 + r_1 - r_i}{2y_2}, \text{ for } i = 2, ..., n, \\ x_{n+1} = \frac{\sum_{i=1}^{n} v_i - \sum_{i=1}^{n} h_i}{2y_2} = \frac{c_1 + r_1 - p}{2y_2}. \end{cases}$$

We can simplify x_i by using equation (4.3) to obtain $x_i = -\frac{\delta_i}{2y_2}$, where y_2 is the remaining unknown variable. Substituting $x_i = -\frac{\delta_i}{2y_2}$ into the variance condition, we obtain that

$$\sigma^2 = \sum_{i=1}^{n+1} \beta_i x_i^2 = \sum_{i=1}^{n+1} \beta_i \left(-\frac{\delta_i}{2y_2}\right)^2 = \frac{1}{(2y_2)^2} \sum_{i=1}^{n+1} \beta_i \delta_i^2 = \frac{1}{(2y_2)^2} \Delta^2,$$

where the last equality follows the definition of Δ in Definition 2. We obtain that $(y_2)^2 = \left(\frac{\Delta}{2\sigma}\right)^2$. Recall that for $\theta \geq Q_n$, the SIP constraints of equation (4.6) include:

$$y_0 + y_1\theta + y_2\theta^2 \leq \sum_{i=1}^{n} (p - r_i - c_i)\, q_i,$$

where the RHS is a positive constant, indicating that $y_2 \leq 0$ (otherwise, the constraint cannot hold when $\theta \to \infty$). We obtain that $y_2^* = -\frac{\Delta}{2\sigma}$. We find that each conjecture binding point θ_i satisfies

$$\theta_i = x_i + \mu = -\frac{\delta_i}{2y_2^*} + \mu = -\frac{\delta_i}{-\frac{\Delta}{\sigma}} + \mu = \mu + \delta_i \frac{\sigma}{\Delta} = \theta_i^*.$$

Complementary slackness indicates that the discrete distribution satisfying $\Pr(\theta = \theta_i) = \lambda_i$ (for $i = 1, 2, \cdots, n+1$) represents the firm's worst-case distribution. Because we have demonstrated that the Lagrangian multipliers satisfy $\lambda_i = \beta_i$ and the conjectured binding points satisfy $\theta_i = \theta_i^*$, we conclude that the discrete distribution shown in equation (4.8) is the firm's worst-case distribution.

Step 3: We derive the remaining shadow prices. The final tangent condition in equation (B-9) indicates that:

$$y_1^* = -2y_2^*\theta_{n+1}^* = \frac{\Delta}{\sigma}\left[\mu + \frac{\sigma}{\Delta}(p - r_1 - c_1)\right] = (p - r_1 - c_1) + \Delta\frac{\mu}{\sigma}.$$

With y_1^* and y_2^* being known, the number of unknown variables in equation (B-8) decreases from $(n+3)$ to $(n+1)$. By solving the remaining $(n+1)$ equations in (B-8), we obtain

$$\begin{cases} y_0^* = y_2^*\left(\theta_{n+1}^*\right)^2 + Z\left(\theta_{n+1}^*|\mathbf{Q}\right) \\ a_n^* = -y_0^* - y_1^*\theta_n^* - y_2^*\left(\theta_n^*\right)^2 + Z\left(\theta_n^*|\mathbf{Q}\right) \\ \vdots \\ a_2^* = -y_0^* - y_1^*\theta_2^* - y_2^*\left(\theta_2^*\right)^2 + Z\left(\theta_2^*|\mathbf{Q}\right) \\ a_1^* = -y_0^* - y_1^*\theta_1^* - y_2^*\left(\theta_1^*\right)^2 + Z\left(\theta_1^*|\mathbf{Q}\right) \end{cases}. \tag{B-11}$$

Because we shall apply Theorem 1, it is unimportant to simplify equation (B-11) at this stage.

With $y_2^* < 0$, we can verify that the proposed shadow prices satisfy all of the SIP constraints in equation (4.6) because the tangent conditions (B-9) are the FOCs and $y_2 < 0$ is the second-order condition for guaranteeing a locally binding constraint. Q.E.D.

Proof of Proposition 3:

As the proof of Proposition 1 indicated, we can compute the value of the zero-sum game from the perspective of the firm or Adverse Nature. When both players play their equilibrium strategy, either perspective will lead to the same value of the zero-sum game. Because we defer the characterization of the firm's equilibrium strategy in the subsequent Proposition 4, we compute the value of the zero-sum game from the firm's perspective. Suppose that the firm's worst-case distribution F^* given by equation (4.8) is realized. For simplicity of exposition, we suppress the superscript $*$ in θ_i^* (i.e., we write θ_i^* in equation (4.8) as θ_i in this proof) but retain the superscript $*$ in F^*. Let $\mathbf{Q}$ be a capacity vector satisfying $Q_i \in [\theta_i, \theta_{i+1}]$.

Recall that $r_{n+1} = p$ and hence, in Table 2, $h_n = r_{n+1} - r_n = p - r_n$. Using r_i for all $i = 1, 2, \cdots, n+1$, we find that equation (B-5) becomes valid for all i for all $i = 1, 2, \cdots, n+1$. We obtain that when the realized demand is $\theta = \theta_i$, the ex post profit equals

$$Z(\theta_i|\mathbf{Q}) = (p - r_i)\,\theta_i + \sum_{j=1}^{i-1}(r_i - r_j - c_j)\,q_j - \sum_{j=i}^{n} c_j q_j = (p - r_i)\,\theta_i + \sum_{j=1}^{i-1}(r_i - r_j)\,q_j - \sum_{j=1}^{n} c_j q_j,$$

where $\sum_{j=1}^{n} c_j q_j = C$ is the total cost associated with the given capacity vector $\mathbf{Q}$. Thus, the firm's expected profit equals:

$$\begin{aligned} Z(\mathbf{Q}|F^*) &= \sum_{i=1}^{n+1}\beta_i Z(\theta_i|\mathbf{Q}) = \sum_{i=1}^{n+1}\beta_i\,(p - r_i)\,\theta_i + \sum_{i=1}^{n+1}\beta_i\left(\sum_{j=1}^{i-1}(r_i - r_j)\,q_j\right) - \sum_{i=1}^{n+1}\beta_i C \\ &= \sum_{i=1}^{n+1}\beta_i\,(p - r_i)\,\theta_i + \sum_{i=1}^{n+1}\beta_i\left(\sum_{j=1}^{i-1}(r_i - r_j)\,q_j\right) - C. \end{aligned} \tag{B-12}$$

We simplify equation (B-12) in the next two steps.

First, we find that the first summation in equation (B-12) equals

$$\sum_{i=1}^{n+1}\beta_i(p - r_i)\theta_i = p\sum_{i=1}^{n+1}\beta_i\theta_i - \sum_{i=1}^{n+1}\beta_i\theta_i r_i. \tag{B-13}$$

According to Corollary 1, the first term in equation (B-13) is $p\mu$. Using equation (4.3), we find that for $i = 1, 2, \cdots, n+1$, $\delta_i - r_i = -r_1 - c_1$, implying that

$$\theta_i r_i = \left(\mu + \delta_i\frac{\sigma}{\Delta}\right)(\delta_i + r_1 + c_1) = \mu\delta_i + \mu\,(r_1 + c_1) + \delta_i^2\frac{\sigma}{\Delta} + (r_1 + c_1)\,\delta_i\frac{\sigma}{\Delta}.$$

Applying Corollary 1, we simplify the second term in equation (B-13) as:

$$\begin{aligned} \sum_{i=1}^{n+1}\beta_i\theta_i r_i &= \sum_{i=1}^{n+1}\beta_i\left[\mu\delta_i + \mu\,(r_1 + c_1) + \delta_i^2\frac{\sigma}{\Delta} + (r_1 + c_1)\,\delta_i\frac{\sigma}{\Delta}\right] \\ &= 0\cdot\mu + \mu\,(r_1 + c_1) + \Delta^2\frac{\sigma}{\Delta} + (r_1 + c_1)\,\frac{\sigma}{\Delta}\cdot 0 = \mu\,(r_1 + c_1) + \Delta\sigma. \end{aligned}$$

Thus, equation (B-13) becomes

$$\sum_{i=1}^{n+1} \beta_i (p-r_i)\theta_i = (p-c_1-r_1)\,\mu - \Delta\sigma.$$

Next, we find that the second summation in equation (B-12) equals

$$\begin{aligned} S &= \sum_{i=1}^{n+1} \beta_i \left(\sum_{j=1}^{i-1} (r_i - r_j)\, q_j \right) = \beta_1 \cdot 0 + \beta_2 \left[(r_2 - r_1)\, q_1 \right] + \beta_3 \left[(r_2 - r_1) q_1 + (r_3 - r_2)(q_1 + q_2) \right] \\ &\quad + \beta_4 \left[(r_2 - r_1) q_1 + (r_3 - r_2)(q_1 + q_2) + (r_4 - r_3)(q_1 + q_2 + q_3) \right] \\ &\quad + \cdots + \beta_n \begin{bmatrix} (r_2 - r_1) q_1 + (r_3 - r_2)(q_1 + q_2) \\ + \cdots + (r_n - r_{n-1})(q_1 + q_2 + \cdots + q_{n-1}) \end{bmatrix} \\ &\quad + \beta_{n+1} \begin{bmatrix} (r_2 - r_1) q_1 + (r_3 - r_2)(q_1 + q_2) \\ + \cdots + (r_n - r_{n-1})(q_1 + q_2 + \cdots + q_{n-1}) \\ + (r_{n+1} - r_n) \sum_{i=1}^{n} q_i \end{bmatrix}. \end{aligned}$$

By reorganizing the terms, we obtain

$$\begin{aligned} S &= (\beta_2 + \beta_3 + \cdots + \beta_{n+1})\,(r_2 - r_1) q_1 + (\beta_3 + \cdots + \beta_{n+1})(r_3 - r_2)(q_1 + q_2) \\ &\quad + \cdots + (\beta_n + \beta_{n+1})\,(r_n - r_{n-1}) \left(\sum_{i=1}^{n-1} q_i \right) + \beta_{n+1} (r_{n+1} - r_n) \left(\sum_{i=1}^{n} q_i \right) \\ &= (1 - \alpha_1)\,(r_2 - r_1) q_1 + (1 - \alpha_2)\,(r_3 - r_2)(q_1 + q_2) \\ &\quad + \cdots + (1 - \alpha_{n-1})\,(r_n - r_{n-1}) \left(\sum_{i=1}^{n-1} q_i \right) + (1 - \alpha_n)\,(r_{n+1} - r_n) \left(\sum_{i=1}^{n} q_i \right) \\ &= \left(\frac{c_1 - c_2}{r_2 - r_1} \right) (r_2 - r_1) q_1 + \left(\frac{c_2 - c_3}{r_3 - r_2} \right) (r_3 - r_2)(q_1 + q_2) \\ &\quad + \cdots + \left(\frac{c_{n-1} - c_n}{r_n - r_{n-1}} \right) (r_n - r_{n-1}) \left(\sum_{i=1}^{n-1} q_i \right) + \left(\frac{c_n}{r_{n+1} - r_n} \right) (r_{n+1} - r_n) \left(\sum_{i=1}^{n} q_i \right) \\ &= (c_1 - c_2) q_1 + (c_2 - c_3)(q_1 + q_2) + \cdots + (c_{n-1} - c_n)(q_1 + q_2 + \cdots + q_{n-1}) \\ &\quad + c_n (q_1 + q_2 + \cdots + q_n) \\ &= q_1 (c_1 - c_2 + c_2 - c_3 + \cdots + c_{n-1} - c_n + c_n) + q_2\,(c_2 - c_3 + c_3 - c_4 + \cdots + c_{n-1} - c_n + c_n) \\ &\quad + \cdots + q_{n-1}\,(c_{n-1} - c_n + c_n) + c_n q_n = \sum_{i=1}^{n} c_i q_i = C. \end{aligned}$$

Hence, equation (B-12) becomes $Z(\mathbf{Q}|F^*) = (p - r_1 - c_1)\mu - \Delta\sigma$, which is a constant whenever $\theta_i \in [Q_{i-1}, Q_i]$ holds. We conclude that under the firm's worst-case distribution F^*, the firm is indifferent among an infinite number of capacity vectors, and her optimal expected profit is a constant that equals $(p - r_1 - c_1)\mu - \Delta\sigma$. Q.E.D.

Proof of Proposition 4:

For simplicity of exposition, we suppress the superscript $*$ (i.e., we write θ_i^* and y_i^* as θ_i and y_i in this proof). In the proof of Proposition 2, we find that the shadow prices satisfy $y_2 = -\frac{\Delta}{2\sigma}$,

$y_1=(p-r_1-c_1)+\Delta\frac{\mu}{\sigma}$, and the remaining shadow prices y_0 and a_i in equations (B-11). By forcing $a_i=0$ for all $i=1,2,\cdots n$, we can solve the firm's optimal capacity levels. Note that we never force $y_0=0$. To illustrate the recursive procedure, we consider $a_n=0$. Equations (B-11) show that

$$\begin{cases} y_0=y_2\left(\theta_{n+1}\right)^2+Z\left(\theta_{n+1}|\mathbf{Q}\right) \\ a_n=-y_0-y_1\theta_n-y_2\left(\theta_n\right)^2+Z\left(\theta_n|\mathbf{Q}\right)=0 \end{cases}$$

We obtain that

$$y_2\left(\theta_{n+1}\right)^2+Z\left(\theta_{n+1}|\mathbf{Q}\right)=-y_1\theta_n-y_2\left(\theta_n\right)^2+Z\left(\theta_n|\mathbf{Q}\right).$$

Using the tangent condition $y_1=-2y_2\theta_{n+1}$, we can rewrite the above equation as:

$$Z\left(\theta_{n+1}|\mathbf{Q}\right)-Z\left(\theta_n|\mathbf{Q}\right)=2y_2\theta_{n+1}\theta_n-y_2\left(\theta_n\right)^2-y_2\left(\theta_{n+1}\right)^2=-y_2\left(\theta_{n+1}-\theta_n\right)^2. \tag{B-14}$$

Because the total cost is sunk, by using equation (4.2), we find that:

$$Z\left(\theta_{n+1}|\mathbf{Q}\right)-Z\left(\theta_n|\mathbf{Q}\right)=\left(p-r_n\right)\left(Q_n-\theta_n\right).$$

Applying the two tangent conditions related to $i=n$ and $i=n+1$, we obtain that $y_1+2y_2\theta_n=p-r_n$ and $y_1=-2y_2\theta_{n+1}$. We find that $2y_2\left(\theta_{n+1}-\theta_n\right)=-\left(p-r_n\right)$ and equation (B-14) becomes:

$$\left(p-r_n\right)\left(Q_n-\theta_n\right)=\frac{1}{2}\left(p-r_n\right)\left(\theta_{n+1}-\theta_n\right),$$

which results in $Q_n=\frac{1}{2}\left(\theta_{n+1}+\theta_n\right)$.

Applying the same method, we can generalize equation (B-14) for any $i=1,2,\cdots,n-1$ by establishing that:

$$Z\left(\theta_{i+1}|\mathbf{Q}\right)-Z\left(\theta_i|\mathbf{Q}\right)=\left(r_{i+1}-r_i\right)\left(Q_i-\theta_i\right)=-y_2\left(\theta_{i+1}-\theta_i\right)^2=\frac{1}{2}\left(r_{i+1}-r_i\right)\left(\theta_{i+1}-\theta_i\right),$$

resulting in $Q_i=\frac{1}{2}\left(\theta_{i+1}+\theta_i\right)$. Q.E.D.

Proof of Corollary 2:

We reintroduce the superscript $*$ in y_i^*. While Proposition 4 solves for the robust optimal capacity vector, the proof of Proposition 2 shows that $y_2^*=-\frac{\Delta}{2\sigma}$ and $y_1^*=(p-r_1-c_1)+\Delta\frac{\mu}{\sigma}$, leaving y_0^* as the only unknown variable. Using the equivalence $P^*=P_1^*=(p-r_1-c_1)\mu-\sigma\Delta$, we obtain that

$$y_0^*+y_1^*\mu+y_2^*\left(\mu^2+\sigma^2\right)=\left(p-r_1-c_1\right)\mu-\sigma\Delta.$$

By reorganizing the terms in the above equation, we obtain

$$y_0^*=\left(p-r_1-c_1\right)\mu-\sigma\Delta-\left[\left(p-r_1-c_1\right)+\Delta\frac{\mu}{\sigma}\right]\mu+\frac{\Delta}{2\sigma}\left(\mu^2+\sigma^2\right)=-\frac{\Delta}{2\sigma}\left(\mu^2+\sigma^2\right).$$

The final remark is that equation (B-11) is more complex than the counterpart equation (A-4), prompting us to derive q_i^* by forcing $a_i = 0$. Readers might concern that this step could weakly reduce the value of P_1^*. Because we compute the value of the zero-sum game based on Adverse Nature's equilibrium strategy F^* in the proof of Proposition 3 and find that the solution in Corollary 2 yields the same value of the zero-sum game, we conclude that the optimal solution shown in Corollary solves P_1 and satisfies Theorem 1. Thus, the robust capacity vector in Proposition 4 is the firm's equilibrium strategy. Q.E.D.

Part C: Proofs for Section 5

Proof of Lemma 4

Notice that when $\rho = \pm 1$, there exist only two binding constraints and we cannot apply the same set of λ_i to determine the probability masses. When $\rho \in (-1, 1)$, it is readily verified that x_i^* shown in Lemma 4 satisfies the moment conditions and the symmetric condition. Q.E.D.

Proof of Proposition 5

For exposition simplicity, we omit the superscript $*$ when involving x_i^*. Using the first and fourth tangent conditions shown in Section 5, we obtain that

$$(y_1 + 2y_2x_1 + y_3x_1)\,x_4 - (y_1 + 2y_2x_4 + y_3x_4)\,x_1 = px_4 - 0 \cdot x_1,$$

which yields that $y_1 = \frac{px_4}{x_4 - x_1}$. With $a_i = 0$, we simplify the binding constraints as follows:

$$\begin{aligned}
&y_0 + y_1x_1 + y_1x_1 + y_2\,(x_1)^2 + y_2\,(x_1)^2 + y_3x_1x_1 = p\,(x_1 + x_1) - 2cq - c_0q_0 \\
&y_0 + y_1x_2 + y_1x_3 + y_2\,(x_2)^2 + y_2\,(x_3)^2 + y_3x_2x_3 = p\,(q + x_2) - 2cq - c_0q_0 \\
&y_0 + y_1x_3 + y_1x_2 + y_2\,(x_3)^2 + y_2\,(x_2)^2 + y_3x_3x_2 = p\,(x_2 + q) - 2cq - c_0q_0 \\
&y_0 + y_1x_4 + y_1x_4 + y_2\,(x_4)^2 + y_2\,(x_4)^2 + y_3x_4x_4 = pq_0 - 2cq - c_0q_0
\end{aligned}$$

Using the first and fourth binding constraints, we obtain

$$\begin{aligned}
p\,(q_0 - x_1 - x_1) &= 2y_1\,(x_4 - x_1) + 2y_2\left(x_4^2 - x_1^2\right) + y_3\left(x_4^2 - x_1^2\right) \\
&= y_1\,(x_4 - x_1) - px_1 = \frac{px_4}{x_4 - x_1}\,(x_4 - x_1) - px_1 = p\,(x_4 - x_1)\,,
\end{aligned}$$

which yields that $q_0^* = x_1 + x_4$, which proves equation (5.3).

The sum of the first and fourth binding constraints minus that of the second and third binding constraints yields that

$$p\,(q_0 - 2q - 2x_2 + 2x_1) = 2y_2\left(x_1^2 + x_4^2 - x_2^2 - x_3^2\right) + y_3\left(x_1^2 + x_4^2 - 2x_2x_3\right).$$

Using tangent conditions, we obtain

$$\begin{aligned}
y_1x_1+2y_2x_1^2+y_3x_1^2 &= px_1\\
y_1x_2+2y_2x_2^2+y_3x_2x_3 &= px_2\\
y_1x_3+2y_2x_3^2+y_3x_2x_3 &= 0\\
y_1x_4+2y_2x_4^2+y_3x_4^2 &= 0.
\end{aligned}$$

Hence, we find that

$$\begin{aligned}
px_1-px_2 &= y_1\left(x_1+x_4-x_2-x_3\right)+2y_2\left(x_1^2+x_4^2-x_2^2-x_3^2\right)+y_3\left(x_1^2+x_4^2-2x_2x_3\right)\\
&= 2y_2\left(x_1^2+x_4^2-x_2^2-x_3^2\right)+y_3\left(x_1^2+x_4^2-2x_2x_3\right).
\end{aligned}$$

We find that $p\left(q_0-2q-2x_2+2x_1\right)=p\left(x_1-x_2\right)$, which is equivalent to

$$q^*=\frac{q_0-x_2+x_1}{2}=\frac{x_1+x_4-x_2+x_1}{2}=\frac{x_2+x_3-x_2+x_1}{2}=\frac{x_1+x_3}{2},$$

which proves equation (5.2).

Under the robust optimal production plan, the total cost equals

$$\begin{aligned}
TC &= 2cq^*+c_0q_0^*=c\left(x_1+x_3\right)+c_0\left(x_1+x_4\right)\\
&= 2(c+c_0)\mu+\sigma\left[A-\frac{1}{2}p(c+2c_0)\right]\sqrt{\frac{1+\rho}{A}}+\frac{\sigma}{2}\sqrt{(1-\rho)pc}.
\end{aligned}$$

The value of the zero-sum game equals

$$\begin{aligned}
Z^* &= \left(1-\frac{2c+c_0}{p}\right)\cdot 2px_1+\frac{c}{p}\cdot 2p\left(x_2+q\right)+\frac{c_0}{p}\cdot p\left(x_1+x_4\right)-TC\\
&= \left(1-\frac{2c+c_0}{p}\right)2px_1+c\left(2x_2+x_1+x_3\right)+c_0\left(x_1+x_4\right)-c\left(x_1+x_3\right)-c_0\left(x_1+x_4\right)\\
&= 2\left[\left(p-2c-c_0\right)x_1+cx_2\right]=2(p-c-c_0)\mu-\left[\sqrt{(1+\rho)A}+\sqrt{(1-\rho)pc}\right]\sigma.
\end{aligned}$$

We obtain that the deference between Z^* with Z_{nc} equals

$$Z^*-Z_{nc}=\sigma\left[2\sqrt{\left(p-c-c_0\right)\left(c+c_0\right)}-\sqrt{(1+\rho)A}-\sqrt{(1-\rho)pc}\right].$$

Observe that

$$A+pc=p\left(c+2c_0\right)-2\left(c+c_0\right)^2+pc=2\left(p-c-c_0\right)\left(c+c_0\right).$$

We find that

$$Z^*-Z_{nc}=\sigma\left[\sqrt{2A+2pc}-\sqrt{(1+\rho)A}-\sqrt{(1-\rho)pc}\right].$$

We can verify that

$$
\begin{aligned}
&2A+2pc-\left(\sqrt{(1+\rho)A}+\sqrt{(1-\rho)pc}\right)^2\\
=\ &2A+2pc-(1+\rho)A-(1-\rho)pc-2\sqrt{(1+\rho)A(1-\rho)pc}\\
=\ &(1-\rho)\,A+(1+\rho)\,pc-2\sqrt{(1+\rho)A(1-\rho)pc}\geq 0,
\end{aligned}
$$

where the last inequality is due to the well-known geometric inequality. Thus, $Z^*\geq Z_{nc}$. Q.E.D.

Part D: Asymmetric Distribution

In the newsvendor model that Natarajan et al. (2018) study, the asymmetric demand $\tilde{\theta}$ gives rise to the following moment constraint:

$$
\int_0^{\mu}-(\theta-\mu)^2\,dF(\theta)+\int_{\mu}^{\infty}(\theta-\mu)^2\,dF(\theta)=s\sigma^2, \tag{D-1}
$$

where s is a known constant satisfying $\frac{\sigma^2-\mu^2}{\sigma^2+\mu^2}\leq s<1$. The selling price of the product is p and the procurement cost is c (where $p>c$). We formulate Adverse Nature's model as follows:

$$
\begin{aligned}
Z_1=&\inf_{q\geq 0,F\in\Omega}\left\{\int_0^{\infty}\left[p\min(\theta,q)-cq\right]dF(\theta)\right\},\\
&\text{s.t.}\ \begin{cases}
\int_0^{\infty}dF(\theta)=1,\\
\int_0^{\infty}\theta dF(\theta)=\mu,\\
\int_0^{\infty}\theta^2 dF(\theta)=\mu^2+\sigma^2,\\
\int_0^{\mu}-(\theta-\mu)^2\,dF(\theta)+\int_{\mu}^{\infty}(\theta-\mu)^2\,dF(\theta)=s\sigma^2,\\
\int_0^{q}(-c)dF(\theta)+\int_q^{\infty}(p-c)dF(\theta)=0.
\end{cases}
\end{aligned} \tag{D-2}
$$

Let $b=1-\frac{(1-s)\sigma^2}{2\mu^2}$ be a useful constant.

Corollary 4 *(Natarajan et al. 2018) In the zero-sum game formulated in equation (D-2), DM's equilibrium strategy q^* is one of the following four cases:*

a) If $b\leq\frac{c}{p}$, then $q^*=0$.

b) If $\frac{1}{2}(1-s)\leq\frac{c}{p}<b$, then $q^*=\mu-\frac{\sigma}{2}\sqrt{\frac{p(1-s)}{2(p-c)}}$.

c) If $\frac{1}{2}\frac{(1-s)^2\sigma^2}{(1+s)\mu^2}\leq\frac{c}{p}<\frac{1}{2}(1-s)$, then $q^*=\mu+\frac{\sigma}{2}\sqrt{\frac{p(1+s)}{2c}}$.

d) If $\frac{c}{p}<\frac{1}{2}\frac{(1-s)^2\sigma^2}{(1+s)\mu^2}$, then $q^*=\frac{\mu}{b}+\frac{(pb-2c)}{2b}\sqrt{\frac{b\sigma^2-(1-b)\mu^2}{c(pb-c)}}$.

The advantages of our new method include the following. First, we bypass the intermediate step of deriving the objective function $Z_{wst}(q)$ because the FOCs are integrated into the new SIP model. Second, the number or corner points to be considered drastically decreases. Specifically, we consider only three candidate corner points while Natarajan et al. (2018) consider six candidate corner points to derive the objective function $Z_{wst}(q)$. We demonstrate the close connection between the robust inventory level and binding constraints.

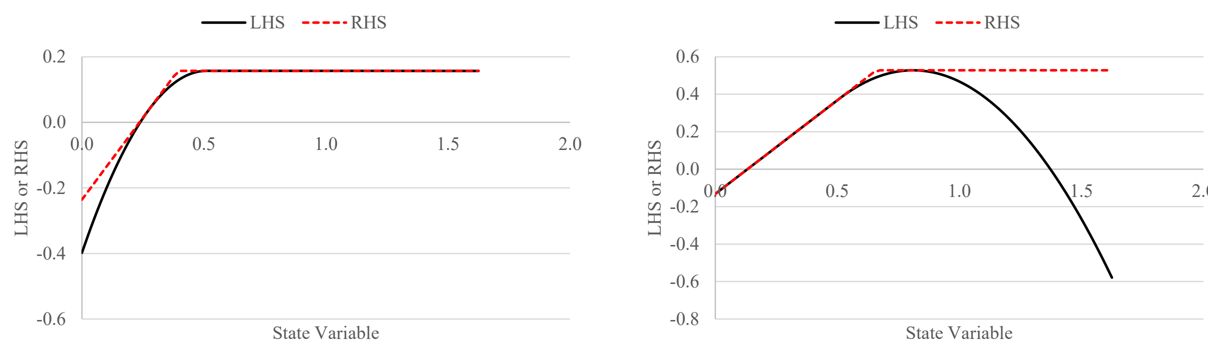
Proof of Corollary 4:

Because the pair of (q,μ) affects the SIP constraints, we first consider case i) $q \le \mu$. Based on Theorem 1, we formulate the SIP model as follows:

$$
\begin{aligned}
& P_1 = \max_{q,y_i,a} \left\{ y_0 + y_1\mu + y_2\left(\mu^2+\sigma^2\right) + y_3 s\sigma^2 \right\} \\
\text{s.t.}\quad & y_0 + y_1\theta + y_2\theta^2 - y_3\left(\theta-\mu\right)^2 - ac \le p\theta - cq, \forall\theta\in[0,q]\,, \\
& y_0 + y_1\theta + y_2\theta^2 - y_3\left(\theta-\mu\right)^2 + a\left(p-c\right) \le \left(p-c\right)q, \forall\theta\in[q,\mu]\,, \\
& y_0 + y_1\theta + y_2\theta^2 + y_3\left(\theta-\mu\right)^2 + a(p-c) \le (p-c)q, \forall\theta\ge\mu.
\end{aligned}
\tag{D-3}
$$

The conjectured binding constraints include: $\theta_1 < q$, and $\theta \ge \mu$ (implying that there exist an infinite number of binding constraints). Figure 3a) illustrates the circumstance using the following parameters $p=1$, $c=0.6$, $\mu=0.51$, $\sigma=0.2$, and $s=-0.1$.

Figure 3 SIP constraints for the Asymmetric Model



(a) Case b) of Corollary 4

(b) Case c) of Corollary 4

If for any $\theta \ge \mu$, $y_0 + y_1\theta + y_2\theta^2 + y_3\left(\theta-\mu\right)^2 + a(p-c) = (p-c)q$ holds, then we find that

$$
y_0 + \left(y_1 - 2y_3\mu\right)\theta + \left(y_2+y_3\right)\theta^2 + y_3\mu^2 + a(p-c) = (p-c)q,
$$

indicating that $y_1 = 2y_3\mu$, $y_2 = -y_3$, and $y_0 + y_3\mu^2 = (p-c)\left(q-a\right)$. Using these three intermediate identities, we simplify the the objective function in equation (D-3) as follows:

$$
\begin{aligned}
& y_0 + y_1\mu + y_2\left(\mu^2+\sigma^2\right) + y_3 s\sigma^2 \\
= & \left(p-c\right)\left(q-a\right) - y_3\mu^2 + 2y_3\mu\cdot\mu - y_3\left(\mu^2+\sigma^2\right) + y_3 s\sigma^2 = \left(p-c\right)\left(q-a\right) - y_3\sigma^2\left(1-s\right).
\end{aligned}
$$

Similarly, when $\theta \in [0,q]$, the SIP constraints in equation (D-3) are simplified as follows:

$$\begin{aligned}
& y_0 + y_1\theta + y_2\theta^2 - y_3(\theta-\mu)^2 - ac \\
= {} & (p-c)(q-a) - y_3\mu^2 + 2y_3\mu\theta - y_3\theta^2 - y_3(\theta-\mu)^2 - ac \\
= {} & (p-c)(q-a) - 2y_3(\theta-\mu)^2 - ac \le p\theta - cq.
\end{aligned}$$

When $\theta \in [q,\mu]$, the SIP constraints in equation (D-3) are simplified as follows:

$$\begin{aligned}
& y_0 + y_1\theta + y_2\theta^2 - y_3(\theta-\mu)^2 + a(p-c) \\
= {} & (p-c)(q-a) - y_3\mu^2 + 2y_3\mu\theta - y_3\theta^2 - y_3(\theta-\mu)^2 + a(p-c) \\
= {} & (p-c)(q-a) - 2y_3(\theta-\mu)^2 + a(p-c) = (p-c)q - 2y_3(\theta-\mu)^2 \le (p-c)\,q,
\end{aligned}$$

which is equivalent to $y_3 \ge 0$. When $\theta \ge q$, the SIP constraints are replaced by $y_0 + y_3\mu^2 = (p-c)(q-a)$.

Therefore, the simplified SIP model is the following:

$$\begin{aligned}
P_1 = {} & \max_{q\ge0, y_3\ge0, a} \left\{(p-c)(q-a) - y_3\sigma^2(1-s)\right\} \\
\text{s.t.} \quad & (p-c)(q-a) - 2y_3(\theta-\mu)^2 - ac \le p\theta - cq, \forall\theta\in[0,q]\,.
\end{aligned} \tag{D-4}$$

The Lagrangian becomes

$$L = (p-c)(q-a) - y_3\sigma^2(1-s) - \lambda_1\Big((p-c)(q-a) - 2y_3(\theta_1-\mu)^2 - ac - p\theta_1 + cq\Big).$$

The FOCs include:

$$\frac{\partial L}{\partial q} = p - c - \lambda_1(p-c+c) = 0 \text{ and } \frac{\partial L}{\partial a} = -(p-c) - \lambda_1(-(p-c)-c) = 0,$$

yielding $\lambda_1 = \frac{p-c}{p}$ (which is the newsvendor ratio). Additionally,

$$\frac{\partial L}{\partial y_3} = -\sigma^2(1-s) + 2\lambda_1(\theta_1-\mu)^2 = 0,$$

which yields that $\theta_1 = \mu - \sigma\sqrt{\frac{p(1-s)}{2(p-c)}}$, which is the binding constraint. The tangent and binding conditions for the point $\theta = \theta_1$ include:

$$(p-c)(q-a) - 2y_3(\theta_1-\mu)^2 - ac = p\theta_1 - cq \text{ and } -4y_3(\theta_1-\mu) = p.$$

We obtain

$$y_3 = \frac{p}{4\sqrt{\frac{p\sigma^2(1-s)}{2(p-c)}}} = \sqrt{\frac{p(p-c)}{8(1-s)\sigma^2}} \text{ and } a = q - \frac{1}{2}(\theta_1+\mu)\,.$$

Using Theorem 1, we let $a = 0$ to obtain that

$$q^* = \frac{1}{2}(\theta_1 + \mu) = \mu - \frac{\sigma}{2}\sqrt{\frac{p(1-s)}{2(p-c)}},$$

which is case b) of Corollary 4. The value of the zero-sum game equals

$$\begin{aligned} P_1^* &= (p-c)(q-a) - y_3\sigma^2(1-s) = (p-c)\frac{1}{2}(\theta_1+\mu) - \sqrt{\frac{p(p-c)}{8(1-s)\sigma^2}}\sigma^2(1-s) \\ &= (p-c)\theta_1 = (p-c)\mu - \frac{\sigma}{2}\sqrt{2p(p-c)(1-s)} = Z_1^*, \end{aligned}$$

which is consistent with Theorem 2.2 of Natarajan et al. (2018). It must hold that $\theta_1 = \mu - \sigma\sqrt{\frac{p(1-s)}{2(p-c)}} > 0$, which yields that $\frac{p-c}{p} > \frac{(1-s)\sigma^2}{2\mu^2} = 1-b$ (equivalent to $b > \frac{c}{p}$). Otherwise, if $b \le \frac{c}{p}$, $\theta_1 \le 0$ and hence, $q^* = 0$, which proves case a) of Corollary 4.

Next, we consider case ii) $q \ge \mu$. Based on Theorem 1, we formulate the SIP model as follows:

$$\begin{aligned} P_1 = \max_{q, y_i, a} &\left\{y_0 + y_1\mu + y_2\left(\mu^2+\sigma^2\right) + y_3 s\sigma^2\right\} \\ \text{s.t.} \quad & y_0 + y_1\theta + y_2\theta^2 - y_3(\theta-\mu)^2 - ac \le p\theta - cq, \forall\theta \in [0,\mu], \\ & y_0 + y_1\theta + y_2\theta^2 + y_3(\theta-\mu)^2 - ac \le p\theta - cq, \forall\theta \in [\mu,q], \\ & y_0 + y_1\theta + y_2\theta^2 + y_3(\theta-\mu)^2 + a(p-c) \le (p-c)q, \forall\theta \ge q. \end{aligned}$$

The main difference occurs when $\theta \in [\mu, q]$.

Case ii-A) The conjectured binding constraints include any $\theta \in [0,\mu]$ and $\theta_1 > q > \mu$. Figure 3b) illustrates the circumstance using the same parameters except $c = 0.2$. To ensure that the constraints are binding for any $\theta \in [0,\mu]$, it must hold that

$$y_0 + y_1\theta + y_2\theta^2 - y_3(\theta-\mu)^2 - ac = p\theta - cq, \forall\theta \in [0,\mu].$$

We obtain that $y_1 + 2y_3\mu = p$, $y_2 = y_3$, and $y_0 = y_3\mu^2 + c(a-q)$. Using these three intermediate identities, we obtain the following equation:

$$\begin{aligned} y_0 + y_1\mu + y_2\left(\mu^2+\sigma^2\right) + y_3 s\sigma^2 &= c(a-q) + y_3\mu^2 + (p - 2y_3\mu)\mu + y_3\left(\mu^2+\sigma^2\right) + y_3 s\sigma^2 \\ &= c(a-q) + y_3(1+s)\sigma^2 + p\mu. \end{aligned}$$

The constraints for $\theta \in [\mu, q]$ become $2y_3(\theta-\mu)^2 \le 0$, and those for $\theta \ge q$ become

$$\begin{aligned} & y_3\mu^2 + c(a-q) + (p-2y_3\mu)\theta + y_3\theta^2 + y_3(\theta-\mu)^2 + a(p-c) \\ =\ & 2y_3(\theta-\mu)^2 + p\theta + ap - cq \le pq - cq. \end{aligned}$$

As such, we simplify the SIP model as follows:

$$\begin{aligned} P_1 = & \max_{q\geq 0, y_3\leq 0, a} \left\{c\left(a-q\right)+y_3\left(1+s\right)\sigma^2+p\mu\right\} \\ \text{s.t.} \quad & 2y_3\left(\theta-\mu\right)^2+p\theta+ap\leq pq, \forall\theta\geq q. \end{aligned}$$

The Lagrangian equals

$$L = c\left(a-q\right)+y_3\left(1+s\right)\sigma^2+p\mu-\lambda_1\left(2y_3\left(\theta-\mu\right)^2+p\theta+ap-pq\right).$$

The FOCs include:

$$\frac{\partial L}{\partial q}=-c+\lambda_1 p=0,\ \frac{\partial L}{\partial a}=c-\lambda_1 p=0,\ \text{and}\ \frac{\partial L}{\partial y_3}=\sigma^2\left(1+s\right)-2\lambda_1\left(\theta_1-\mu\right)^2=0.$$

We obtain that $\lambda_1=\frac{c}{p}$ (which also relates to the newsvendor ratio) and $\theta_1=\mu+\sigma\sqrt{\frac{p(1+s)}{2c}}$ (which is the binding constraint). The tangent and binding conditions for the point $\theta=\theta_1$ include:

$$2y_3\left(\theta_1-\mu\right)^2+p\theta_1+ap-pq=0 \text{ and } 4y_3\left(\theta_1-\mu\right)+p=0.$$

We obtain

$$y_3=-\frac{p}{4\sqrt{\frac{p\sigma^2(1+s)}{2c}}}=-\sqrt{\frac{pc}{8\left(1+s\right)\sigma^2}} \text{ and } a=q-\frac{1}{2}\left(\theta_1+\mu\right).$$

Using Theorem 1, we let $a=0$ to obtain that

$$q^*=\frac{1}{2}\left(\theta_1+\mu\right)=\mu+\frac{\sigma}{2}\sqrt{\frac{p\left(1+s\right)}{2c}}.$$

The value of the zero-sum game equals

$$\begin{aligned} P_1^* &= Z_1^*=c\left(a-q\right)+y_3\left(1+s\right)\sigma^2+p\mu \\ &= -\frac{c}{2}\left(\theta_1+\mu\right)-\sqrt{\frac{pc}{8\left(1+s\right)\sigma^2}}\left(1+s\right)\sigma^2+p\mu=\left(p-c\right)\mu-\frac{\sigma}{2}\sqrt{2pc\left(1+s\right)}, \end{aligned}$$

which proves case c) of Corollary 4.

Case ii-B) While cases b) and c) of Corollary 4 involve mixed distributions (not mixed strategy) as the equilibrium strategy of Adverse Mature, the remaining case d) of Corollary 4 involves a three-point distribution. The conjectured binding constraints include: $\theta=0$, $\theta=\theta_2\in[\mu,q]$, and $\theta=\theta_3>q$. The Lagrangian equals

$$\begin{aligned} L = & y_0+y_1\mu+y_2\left(\mu^2+\sigma^2\right)+y_3 s\sigma^2-\lambda_1\left(y_0-y_3\mu^2-ac+cq\right) \\ & -\lambda_2\left(y_0+y_1\theta_2+y_2\theta_2^2+y_3\left(\theta_2-\mu\right)^2-ac-p\theta_2+cq\right) \\ & -\lambda_3\left(y_0+y_1\theta_3+y_2\theta_3^2+y_3\left(\theta_3-\mu\right)^2+a(p-c)-(p-c)q\right). \end{aligned}$$

The FOCs include:

$$\begin{cases} \frac{\partial L}{\partial a} = \lambda_1 c + \lambda_2 c + \lambda_3 (p-c) = 0 \\ \frac{\partial L}{\partial y_0} = 1 - \lambda_1 - \lambda_2 - \lambda_3 = 0 \\ \frac{\partial L}{\partial y_1} = \mu - \lambda_2 \theta_2 - \lambda_3 \theta_3 = 0 \\ \frac{\partial L}{\partial y_2} = (\mu^2 + \sigma^2) - \lambda_2 \theta_2^2 - \lambda_3 \theta_3^2 = 0 \\ \frac{\partial L}{\partial y_3} = s\sigma^2 + \lambda_1 \mu^2 - \lambda_2 (\theta_2 - \mu)^2 - \lambda_3 (\theta_3 - \mu)^2 . \end{cases}$$

Solving this system with 5 unknowns and 5 equations, we obtain that $\lambda_1 = \frac{(1-s)\sigma^2}{2\mu^2}$, $\lambda_2 = 1 - \frac{c}{p} - \frac{(1-s)\sigma^2}{2\mu^2}$, $\lambda_3 = \frac{c}{p}$, $\theta_2 = \frac{\mu}{b} - \frac{\sqrt{c(pb-c)(b\sigma^2-(1-b)\mu^2)}}{b(pb-c)}$, and $\theta_3 = \frac{\mu}{b} + \frac{\sqrt{c(pb-c)(b\sigma^2-(1-b)\mu^2)}}{bc}$. Case d) of Corollary 4 requires that the binding point satisfies $\theta_2 > \mu$, which implies that $\frac{c}{p} < \frac{1}{2}\frac{(1-s)^2\sigma^2}{(1+s)\mu^2}$.

The binding and tangent conditions include:

$$\begin{cases} y_0 - y_3\mu^2 - ac + cq = 0, \\ y_0 + y_1\theta_2 + y_2\theta_2^2 + y_3 (\theta_2 - \mu)^2 - ac - p\theta_2 + cq = 0, \\ y_0 + y_1\theta_3 + y_2\theta_3^2 + y_3 (\theta_3 - \mu)^2 + a(p-c) - (p-c)q = 0, \\ y_1 + 2y_2\theta_2 + 2y_3 (\theta_2 - \mu) = p, \\ y_1 + 2y_2\theta_3 + 2y_3 (\theta_3 - \mu) = 0. \end{cases}$$

Solving the above system of equations, we obtain the following results:

$$\begin{aligned} a &= q - \frac{1}{2}(\theta_2 + \theta_3) \\ y_3 &= \frac{c\theta_2[\mu^2 + \sigma^2 - \mu\theta_3]}{4\mu^2(b\sigma^2 - (1-b)\mu^2)} \\ y_2 &= -y_3 - \frac{p}{2(\theta_3 - \theta_2)} \\ y_1 &= \frac{p}{2} - (y_2 + y_3)(\theta_2 + \theta_3) + 2y_3\mu \\ y_0 &= y_3\mu^2 - cq \end{aligned}$$

Setting $a = 0$, we obtain

$$\begin{aligned} q^* &= \frac{1}{2}(\theta_2 + \theta_3) = \frac{\mu}{b} + \frac{1}{2}\left(\frac{\sqrt{c(pb-c)(b\sigma^2 - (1-b)\mu^2)}}{bc} - \frac{\sqrt{c(pb-c)(b\sigma^2 - (1-b)\mu^2)}}{b(pb-c)} \right) \\ &= \frac{\mu}{b} + \frac{(pb-2c)}{2b}\sqrt{\frac{b\sigma^2 - (1-b)\mu^2}{c\,(pb-c)}}. \end{aligned}$$

To reconcile with Theorem 2.2 in Natarajan et al. (2018), we apply the identity $1 - b = \frac{(1-s)\sigma^2}{2\mu^2}$ to show that

$$\begin{aligned} \frac{b\sigma^2 - (1-b)\mu^2}{c\,(pb-c)} &= \frac{2b\sigma^2 - 2(1-b)\mu^2 + 2(1-b)^2\mu^2 - 2(1-b)^2\mu^2}{2c\,(pb-c)} \\ &= \frac{2b\sigma^2 - (1-s)\,\sigma^2 + (1-b)\,(1-s)\,\sigma^2 - 2(1-b)^2\mu^2}{2c\,(pb-c)} = \frac{(1+s)b\sigma^2 - 2(1-b)^2\mu^2}{2c\,(pb-c)}. \end{aligned}$$

Hence,

$$q^* = \frac{\mu}{b} + \frac{(pb-2c)}{2b}\sqrt{\frac{b\sigma^2 - (1-b)\mu^2}{c\,(pb-c)}} = \frac{\mu}{b} + \frac{(pb-2c)}{2b}\sqrt{\frac{(1+s)b\sigma^2 - 2(1-b)^2\mu^2}{2c\,(pb-c)}}.$$

Using the shadow prices, we can compute the value of the zero-sum game case d) of Corollary 4. Q.E.D.

Adverse Nature's Strategy for Cases b) and c) of Corollary 4

The proof of Corollary 4 solves DM's equilibrium strategy q^* without deriving $Z_{wst}(q)$. The remaining task is to determine when cases b) or c) could occur. Because case a) occurs when $b \le \frac{c}{p}$ and case d) occurs when $\frac{c}{p} < \frac{1}{2}\frac{(1-s)^2\sigma^2}{(1+s)\mu^2}$, we assume that $b > \frac{c}{p} \ge \frac{1}{2}\frac{(1-s)^2\sigma^2}{(1+s)\mu^2}$.

When case b) occurs, we show that the SIP constraints are binding at point θ_1 (where $\theta_1 < q^* < \mu$) and any point $\theta \ge \mu$. We conclude that Adverse Nature's equilibrium strategy F^* is a mixed distribution. The mixed distribution F^* exhibits the following characteristics:

- It allocates a probability mass that $\lambda_1 = \frac{p-c}{p}$ at the point $\theta = \theta_1 = \mu - \sqrt{\frac{p\sigma^2(1-s)}{2(p-c)}}$.
- It allocates nonnegative density for any $\theta \ge \mu$.
- It allocates zero density for point $\theta \in [0, \theta_1) \cup (\theta_1, \mu)$.
- It satisfies the moment conditions on the mean, variance, and asymmetry.

Based on the above characteristics of F^*, we find that

$$0 = \int_0^\mu (\theta-\mu)dF(\theta) + \int_\mu^\infty (\theta-\mu)dF(\theta) = \lambda_1(\theta_1-\mu) + \int_\mu^\infty (\theta-\mu)dF(\theta),$$

which implies that the upper semi-mean equals

$$\int_\mu^\infty (\theta-\mu)dF(\theta) = \lambda_1(\mu-\theta_1) = \sigma\sqrt{\frac{(1-s)\,(p-c)}{2p}}.$$

Before verifying whether the moment constraint (D-1) can hold, we develop the following intermediate result.

Lemma 5 *Suppose that a random variable $\tilde{\theta}$ satisfies that $\int_\mu^\infty dF(\theta) = \frac{c}{p} < 1$, $\int_\mu^\infty (\theta-\mu)\,dF(\theta) = \sigma\sqrt{\frac{(1-s)(p-c)}{2p}}$. It holds that $\int_\mu^\infty (\theta-\mu)^2\,dF(\theta) \ge \frac{(1-s)(p-c)\sigma^2}{2c}$.*

Proof. Let $\varepsilon = \theta - \mu$ as a transformed variable. Using duality, we formulate the following SIP model:

$$\begin{aligned} P = \max_{y_0,y_1} & \left\{ y_0\frac{c}{p} + y_1\sigma\sqrt{\frac{(1-s)\,(p-c)}{2p}} \right\} \\ \text{s.t.} \quad & y_0 + y_1\varepsilon \le \varepsilon^2, \forall \varepsilon \ge 0. \end{aligned}$$

It is readily verified that the optimal solution satisfies $y_0^* = -\frac{\sigma^2(p-c)p}{2c}$ and $y_1^* = \frac{\sigma}{c}\sqrt{2p\,(1-s)\,(p-c)}$ (with a unique binding constraint at point $\varepsilon_1 = \frac{\sigma}{c}\sqrt{\frac{p(1-s)(p-c)}{2}}$). Hence,

$$P^* = -\frac{\sigma^2\,(p-c)\,p}{2c}\frac{c}{p} + \frac{\sigma}{c}\sqrt{2p\,(1-s)\,(p-c)}\sigma\sqrt{\frac{(1-s)\,(p-c)}{2p}} = \frac{(1-s)\,(p-c)\,\sigma^2}{2c}.$$

We obtain that $\int_\mu^\infty (\theta-\mu)^2\,dF(\theta) \geq \frac{(1-s)(p-c)\sigma^2}{2c}$. ■

Notice that the moment constraint (D-1) requires that

$$\begin{aligned} s\sigma^2 &= -\int_0^\mu (\theta-\mu)^2 dF(\theta) + \int_\mu^\infty (\theta-\mu)^2 dF(\theta) = -\lambda_1(\theta_1-\mu)^2 + \int_\mu^\infty (\theta-\mu)^2 dF(\theta) \\ &= -\frac{(p-c)}{p}\frac{p\sigma^2\,(1-s)}{2\,(p-c)} + \int_\mu^\infty (\theta-\mu)^2 dF(\theta) = -\frac{(1-s)\,\sigma^2}{2} + \int_\mu^\infty (\theta-\mu)^2 dF(\theta). \end{aligned}$$

We obtain that the upper semi-variance equals $\int_\mu^\infty (\theta-\mu)^2 dF(\theta) = \frac{(1+s)\sigma^2}{2}$. However, Lemma 5 indicates that

$$\int_\mu^\infty (\theta-\mu)^2 dF(\theta) \geq \frac{(1-s)\,(p-c)\,\sigma^2}{2c}.$$

To avoid a contraction between the moment constraint (D-1) with Lemma 5, it must hold that

$$\frac{(1+s)\,\sigma^2}{2} \geq \frac{(1-s)\,(p-c)\,\sigma^2}{2c},$$

which implies that $\frac{c}{p} \geq \frac{1}{2}\,(1-s)$. We conclude that case b) of Corollary 4 occurs when $\frac{1}{2}\,(1-s) \leq \frac{c}{p} < b$. When $\frac{1}{2}\,(1-s) = \frac{c}{p}$, Lemma 5 implies that Adverse Nature's strategy F^* is the following two-point distribution:

$$\begin{cases} \Pr\left(\tilde{\theta} = \mu - \sigma\sqrt{\frac{1-s}{1+s}} = \mu - \sigma\sqrt{\frac{c}{p-c}}\right) = \frac{p-c}{p} = \frac{1+s}{2}, \\ \Pr\left(\tilde{\theta} = \mu + \sigma\sqrt{\frac{1+s}{1-s}} = \mu + \sigma\sqrt{\frac{p-c}{c}}\right) = \frac{c}{p} = \frac{1-s}{2}. \end{cases}$$

However, Adverse Nature's equilibrium strategy is not unique when $\frac{1}{2}\,(1-s) < \frac{c}{p} < b$. Because the remaining space on cost parameters is $\frac{1}{2}\frac{(1-s)^2\sigma^2}{(1+s)\mu^2} \leq \frac{c}{p} < \frac{1}{2}\,(1-s)$, case c) of Corollary 4 must occur and Adverse Nature's equilibrium strategy F^* is also a mixed distribution.

Part E: Robust Adverse Selection

Robust adverse selection has been studied by Pınar and Kızılkale (2017), Carrasco et al. (2018), and Carrasco et al. (2019) in the context of selling a single unit of good to a buyer whose valuation of the good is private. The seller knows only the mean or variance of the buyer's private valuation but not the exact distribution. Koçyiğit et al. (2020) study a circumstance where the auctioneer sells one unit of a good to multiple buyers. Unlike the case with a single buyer, the auctioneer faces a challenge in inducing a truth-revealing equilibrium with multiple buyers. The IC constraints can follow the notion of ex post incentive compatibility or dominating strategy incentive compatibility. Bayesian incentive compatibility (which must involve a prior distribution) cannot be used because the joint distribution is unknown to all players.

To demonstrate our method, we choose the mean-only model that Carrasco et al. (2018, 2019) study but make some minor modifications as follows. DM knows that the buyer's private value of the good is a random variable $\tilde{\theta}$ satisfying the mean condition $E\left(\tilde{\theta}\right) = \mu$ and the boundary condition $0 \leq \theta \leq r$, where $r > 0$ is a known constant representing the buyer's highest valuation. By applying the revelation principle (Baron and Myerson 1982), it suffices to focus on the class of direct and incentive compatible mechanisms. Let $(R(\cdot), q(\cdot))$ be DM's mechanism where $R(\cdot)$ represents the payment function and $q(\cdot)$ characterizes the allocation policy. To induce a truth-revealing equilibrium, the following IC constraints must hold:

$$\theta q(\theta) - R(\theta) \geq \theta q(m) - R(m), \text{ for } \forall \theta \in [0, r], \ \forall m \in [0, r], \text{ and } \theta \neq m.$$

First, consider the following two IC constraints associated with type θ and m (where $\theta \neq m$):

$$\begin{cases} \theta q(\theta) - R(\theta) \geq \theta q(m) - R(m) \\ m q(m) - R(m) \geq m q(\theta) - R(\theta) \end{cases}.$$

By summing these two constraints, we find that

$$\theta q(\theta) - \theta q(m) + m q(m) - m q(\theta) = (\theta - m)(q(\theta) - q(m)) \geq 0,$$

indicating that $q(\theta)$ is weakly increasing in θ. Additionally, DM has only 1 unit of the good and hence, $0 \leq q(\theta) \leq 1$ for any $\theta \in [0, r]$. It must also hold that $q(r) = 1$ (i.e., no distortion at the top such that the buyer with the highest valuation always receives the good). Second, we apply the envelope theorem to simplify the IC constraints to obtain the following equation:

$$R(\theta) = \theta q(\theta) - \int_0^\theta q(t)\, dt. \tag{E-1}$$

The second step is a variable reduction process, in which the payment function $R(\theta)$ is represented by the allocation policy $q(\theta)$. We omit the relevant details on deriving equation (E-1) because this procedure is now a textbook procedure. We emphasize that equation (E-1) is distribution-free, meaning that any non-decreasing function $q(\theta)$ satisfying $0 \leq q(\theta) \leq 1$ would be feasible even under ambiguity. This observation explains why the issue of infeasibility does not arise in adverse selection. Thus, DM's task is to determine the allocation policy $q(\theta)$ to maximize her expected sales revenue, which equals:

$$Z = \int_0^r R(\theta)\, dF(\theta) = \int_0^r \left[\theta q(\theta) - \int_0^\theta q(t)\, dt\right] dF(\theta) = \int_0^r \left(\theta - \frac{1 - F(\theta)}{f(\theta)}\right) q(\theta)\, dF(\theta),$$

where the term $\theta - \frac{1-F(\theta)}{f(\theta)}$ is the so-called the virtual valuation, which is weakly increasing in θ if the log-concave property holds. However, we do not impose any assumption on $F(\theta)$. Without knowing the underline distribution, we cannot determine the virtual value. We obtain that the ex post payoff function equals

$$Z(\theta|q(\theta)) = \theta q(\theta) - \int_0^\theta q(t)\,dt = \left(\theta - \frac{1-F(\theta)}{f(\theta)}\right) q(\theta).$$

Because $\theta \in [0, r]$, we observe that the allocation policy $q(\theta)$ is a function (or an infinite vector as opposed to the case with a finite vector). The first derivative with respect to $q(\theta)$ then equals

$$\frac{\partial Z(\theta|q(\theta))}{\partial q(\theta)} = \left(\theta - \frac{1-F(\theta)}{f(\theta)}\right).$$

We formulate the min-max version of the model as follows:

$$\begin{aligned} Z_1 &= \inf_{F\in\Omega} \max_{q(\cdot)} \int_0^r \left(\theta - \frac{1-F(\theta)}{f(\theta)}\right) q(\theta)\,dF(\theta) \\ \text{s.t.} &\quad \theta - \frac{1-F(\theta)}{f(\theta)} = 0, \text{ for } \theta \in [0, r) \text{ and } q(r) = 1, \end{aligned} \tag{E-2}$$

where the boundary condition $q(r) = 1$ is due to the classic result of no distortion at the top.

Corollary 5 *Let h^* be the solution satisfying the following equation:*

$$\mu = h^* + h^* \ln\left(\frac{r}{h^*}\right). \tag{E-3}$$

Additionally let

$$y_1^* = \frac{1}{\ln\left(\frac{r}{h^*}\right)}.$$

DM's robust optimal allocation policy satisfies the following equation:

$$q^*(\theta) = \begin{cases} 0 & \text{if } 0 \le \theta \le h^*, \\ 1 - y_1^* \ln\left(\frac{r}{\theta}\right) & \text{if } h^* \le \theta \le r, \end{cases} \tag{E-4}$$

On the other hand, DM's worst-case distribution satisfies

$$F^*(\theta) = \begin{cases} 0 & \text{if } 0 \le \theta \le h^*, \\ 1 - \frac{h^*}{\theta} & \text{if } h^* \le \theta < r, \\ 1 & \text{if } \theta = r. \end{cases} \tag{E-5}$$

DM's worst-case expected profit is $Z^* = h^*$.

Both $q^*(\theta)$ and $F^*(\theta)$ in Corollary 5 are consistent with Carrasco et al. (2019) but our method is quicker. There are several noteworthy observations.

- The worst-case distribution $F^*(\theta)$ in equation (E-5) is a mixed distribution with a jump at point $\theta = r$. The intuition is that the buyer with the highest valuation always obtains the good with probability 1 and always secures a positive information rent. In terms of reducing DM's expected payoff, the buyer with the highest valuation is more important than any other buyers, making Adverse Nature allocate a positive mass probability to the point $\theta = r$.
- Under the worst-case distribution $F^*(\theta)$, the virtual value is negative for any $\theta < h^*$, is zero for any $\theta \in [h^*, r)$, and is positive only when $\theta = r$, due to the jump in the cumulative distribution function. It can be easily verified that every feasible mechanism yields the same expected profit under distribution F^*. However, when the worst-case distribution does not realize, only the robust mechanism in equation (E-3) can guarantee a weakly higher expected profit, while any nonrobust mechanism could experience a lower expected profit. Therefore, the robust mechanism can serve as a useful benchmark for practitioners and researchers to evaluate the performance of their proposed mechanism.
- Let σ_0^2 be the variance of the extreme distribution in equation (E-5). Suppose that the variance is also known to be σ^2. Bhatia and Davis (2000) state that $0 \le \sigma^2 \le \mu(r-\mu)$. We can claim that when $\sigma > \sigma_1$, the optimal shadow price y_2^* for the variance resource in the mean-variance model is strictly positive. In addition, the point $\theta = 0$ could have a positive probability mass (i.e., the worst-case distribution has two jumps at $\theta = 0$ and $\theta = r$) when σ^2 approaches $\mu(r-\mu)$. When $\sigma > \sigma_0$, $y_2^* < 0$; and when $\sigma = \sigma_0$, $y_2^* = 0$. The information about the variance strictly improves DM's worst-case expected profit unless $\sigma = \sigma_0$.

Proof of Corollary 5

The IC constraint in equation (E-2) immediately yields the following mixed distribution:

$$F(\theta|h) = \begin{cases} 0 & \text{if } \theta \in [0, h], \\ 1 - \frac{h}{\theta} & \text{if } h \le \theta < r, \\ 1 & \text{if } \theta = r. \end{cases}$$

which has a jump at $\theta = r$ and characterizes the best response of Adverse Nature. Using the condition on the mean, we obtain that

$$\mu = \int_h^r \theta d\left(1 - \frac{h}{\theta}\right) + \Pr\left(\tilde{\theta} = r\right) r = \int_h^r \theta \frac{h}{\theta^2} d\theta + \left(1 - 1 + \frac{h}{r}\right) r = h + h \ln\left(\frac{r}{h}\right).$$

which has only one unknown variable h and yields equation (E-3). We also find that the equilibrium payoff of DM (or the value of the zero-sum game) is $Z^* = h^*$ and the worst-case distribution is $F^*(\theta) = F(\theta|h^*)$, confirming equation (E-5). We observe that our new method quickly identifies the value of the zero-sum game and the best response of Adverse Nature.

The final task is to determine the robust procurement policy. We consider the SIP model

$$P_1 = \max\{y_0 + y_1\mu\}$$
$$\text{s.t. } y_0 + y_1\theta \le \theta q(\theta) - \int_0^\theta q(t)\,dt, \forall\theta \in [0, r].$$

This SIP model is binding for $\theta \in [h^*, r]$ because Adverse Nature allocates positive density over this interval. We obtain that $y_0 + y_1\theta = \theta q(\theta) - \int_0^\theta q(t)\,dt$. Taking the first derivative with respect to θ, we obtain that

$$y_1 = q(\theta) - \theta q'(\theta) - q(\theta) = -\theta q'(\theta).$$

The above differential equation characterizes DM's allocation policy as

$$q(\theta|y_1) = \begin{cases} 0 & \text{if } \theta \in [0, h^*], \\ 1 - y_1 \ln\left(\frac{r}{\theta}\right) & \text{if } h^* \le \theta \le r. \end{cases}$$

Because we already know that h^* from equation (E-3), we can then determine y_1^* by solving

$$q(h^*|y_1) = 1 - y_1 \ln\left(\frac{r}{h^*}\right) = 0.$$

We find that $\frac{1}{y_1^*} = \ln\left(\frac{r}{h^*}\right)$ based on $q(h^*) = 0$, confirming equation (E-4). Q.E.D.