

Profit Sharing Agreements in Decentralized Supply Chains: A Distributionally Robust Approach

Qi Fu*, Chee-Khian Sim[†], Chung-Piaw Teo[‡]

Abstract

How should decentralized supply chains set the profit sharing terms using minimal information on demand and selling price? We develop a distributionally robust Stackelberg game model to address this question. Our framework uses only the first and second moments of the price and demand attributes, and thus can be implemented using only a parsimonious set of parameters. More specifically, we derive the relationships among the optimal wholesale price set by the supplier, the order decision of the retailer, and the corresponding profit shares of each supply chain partner, based on the information available. Interestingly, in the distributionally robust setting, the correlation between demand and selling price has no bearing on the order decision of the retailer. This allows us to simplify the solution structure of the profit sharing agreement problem dramatically. Moreover, the result can be used to recover the optimal selling price when the mean demand is a linear function of the selling price (cf. Raza (2014)).

Keywords. profit sharing agreements; decentralized supply chains; distributionally robust planning; revealed preference; completely positive program

*University of Macau, gracefu@umac.mo

[†]University of Portsmouth, chee-khian.sim@port.ac.uk

[‡]NUS Business School & Institute of Operations Research and Analytics, National University of Singapore, bizteocp@nus.edu.sg

1 Introduction

Profit sharing agreements are among the most common types of contractual arrangements for firms in a supply chain. Under such agreements, the retailer, as a provider to a consumer market, engages another firm (the supplier) as a contractor to provide essential supplies for the operation. In addition to paying for the essential supplies, the retailer pays the supplier a certain portion of the profit generated from the operation. In this way, profit sharing agreements serve to align incentives between the retailer and the supplier in a supply chain.

A variant of this concept, known as sharecropping, has been used in agriculture where the landlord allows the tenant to use his land in exchange for a specified share of production. However, the terms of payment can vary widely. For instance, the landlord can bear a portion of the production cost, and thus demand a higher share of the production, or may bear the entire risk and employ the farmers as fixed wage workers. Consider, for example, B-BOVID, an agricultural firm operating in Takoradi, Ghana. In 2014, this firm introduced a unique profit sharing model to alleviate the plight of oil palm farmers in the region. The farmers would sell their palm fruits to B-BOVID for on-the-spot money. B-BOVID would then refine the palm fruits for sale and return a portion of the profits generated back to the farmers, based on the quantity supplied by the farmers to the company. Thus, in addition to reducing poverty, the company incentivises and sustains the interest of the farmers in oil palm plantations. Stiglitz (1989) presents early economic analysis of these contractual forms, and confirms the benefits of sharecropping in streamlining the allocation of risk and rewards among the different parties involved.

The profit sharing contract is also prevalent in the entertainment industry. See for instance, Weinstein (1998) and the references therein for a discussion of the evolution of profit sharing contracts in Hollywood, and an interesting discussion on the difference between “net profit” sharing and “gross profit” sharing (i.e., revenue sharing) contract. Filson et al. (2005) study the profit sharing contracts written between certain movie distributors and a theatre chain in St. Louis, Missouri, and conclude through empirical analysis that the profit sharing contracts “evolved to help distributors and exhibitors share risks and overcome measurement problems and not to overcome asymmetric information problems.”

In the natural resource industry, the profit sharing agreement (or in the form of production

sharing agreement, (PSA)) has been practiced for a long period of time. Under a PSA, the government engages a foreign oil company to operate petroleum exploration. After deducting the royalty and operating costs, the remainder (called profit oil) will be split between the government and company.

For many commodity products, such as oil and agricultural commodities, sellers face not only uncertain demands but also extreme price fluctuations. For example, according to the United Nations Food and Agriculture Organization (FAO), food prices have fluctuated wildly over the last few years. The index rose from 122 in 2006 to 214 in June 2008 as the 2007-2008 food price crisis unfolded. The index then fell rapidly in the second half of 2008, sinking to 140 in March 2009. In the latter half of 2010, it increased considerably, and reaching 215 in December¹. The prices for agricultural products are often largely influenced by supply and demand factors, such as weather conditions, population changes, and food stocks, which lead to considerable risk and uncertainty in the agricultural commodity price. Similarly, the price of oil has also fluctuated significantly over the last few years. Therefore, facing such high price and demand risks, profit sharing contracts allow the supply chain partners to share risks.

An additional challenge is to establish a reasonable method to share profits. One approach is to use the Nash bargaining framework, where the total rents of the supply chain operation are split between the supplier and retailer according to the bargaining strength of the parties involved, together with the value of outside options for the partners. The approach has been applied in the natural resource industry to provide a benchmark to disentangle the effect of other competing theories on the bargaining outcome (cf. McMillan and Waxman (2007)). However, the distribution of profits in the Nash bargaining model often depends on subjective input of the relative bargaining power of the parties. More importantly, these models ignore the fact that contractual elements in the profit sharing agreements are often based on full knowledge of the selling price, demand, and/or cost of operations, which are often not assured due to the difficulty in fully characterizing the distributions of random elements. Accordingly, in many cases, we may only have confidence in estimating some distribution parameters, such as mean and variance. This problem is compounded by the provision of the retailer for the operational costs in the computation of net profit. The evaluation of this cost component is often the bone of contention between the two parties involved in the profit sharing agreements, as often seen in the mining

and entertainment industries.

Motivated by these concerns, in this paper, we develop a distributionally robust Stackelberg game model to study the profit sharing contract design problem with limited information about demand and price distributions. Aghassi and Bertsimas (2006) present a distribution-free, robust optimization model for incomplete information games, in which the players commonly know an uncertainty set of possible game parameter values without knowledge of exact distributional information and each player seeks to maximize the worst-case expected payoff. Their proposed robust game model presumes that each player uses a robust optimization, and therefore a worst-case approach to the uncertainty to enable mutual prediction of the game outcomes. We adopt the same assumption that the partial distribution information on price and demand is common knowledge between the two players in our base model analysis (the case of demand information asymmetry is discussed in the online appendix). That is, both the retailer and supplier have access to the same demand information, such as POS data, and market price information. As the exact distribution information on demand and price is not available, both parties apply a distributionally robust approach to minimize their worst-case expected profit.

As a direct application, we show that this approach can be used to derive testable restrictions on the outcomes in the supply chain setting. More specifically, given the decisions of the supplier and retailer, the unit cost of production (incurred by the supplier) and the selling price information (mean and variance), but without the demand and selling price correlation information, we develop testable conditions on the outcomes that are consistent with the prediction of the distributionally robust Stackelberg game based on some random demand model. In the traditional Stackelberg game model, this is a difficult problem because finding the optimal wholesale price is already a complicated problem, and a well known result by Lariviere and Porteus (2001) indicates that the problem is solvable when the demand distribution satisfies the “increasing-generalized failure-rate” property. A closed-form expression for the optimal wholesale price is known only for specific demand distribution (for instance, when demand is uniform with support in a bounded interval). Otherwise, a closed-form expression is generally not available. Thus, it is difficult to find testable conditions on the optimal wholesale price and order quantity. Interestingly, as shown in this paper, this problem becomes tractable in a distributionally robust Stackelberg game model. By building a prescriptive model on the optimal behavior in the choice of the

contract parameters and operational decisions, we hope to find testable relationships between these parameters if the parties act in a rational manner as prescribed by the Stackelberg game.

The remainder of this paper is organized as follows. In Section 2, we review the relevant literature. Section 3 introduces the distributionally robust model. We derive the robust ordering decision of the last stage problem in Section 4, and address the profit sharing contract design problem in Section 5. We then extend our model to the price dependent demand case in Section 6. We also discuss other potential applications of this model in Appendix A. Section 7 reports numerical studies. Concluding thoughts are provided in Section 8. All proofs to the technical results are relegated to Appendix B.

2 Literature Review

Our research is related to two streams of literature. The first stream of literature is robust inventory management, where decisions have to be made with incomplete distribution information. One common approach to this situation, according to the literature, is the max-min approach, which was pioneered by Scarf (1958) to model the newsvendor problem with only mean and variance of the demand distribution known, aiming to find an order quantity to maximize the worst-case expected profit over all possible distributions with the given mean and variance. Scarf derives a closed-form solution for the robust order quantity. Gallego and Moon (1993) provide a more concise proof of Scarf's ordering rule and apply this approach to several variants of the newsvendor problem, for example, the recourse case and the fixed ordering cost case. Moon and Choi (1995) apply the max-min approach to the newsvendor problem with customer balking. A wide variety of extensions to the basic robust newsvendor problem have been addressed in the literature, including multi-period (e.g., Gallego 1992, Moon and Gallego 1994), multi-echelon (e.g., Moon and Choi 1997), and multi-item (e.g., Moon and Silver 2000) extensions. Readers may refer to Gallego et al. (2001) for a more detailed review of the early literature. Mamani et al. (2016) obtain a closed-form solution to a multi-period inventory management problem, using a class of uncertainty sets motivated by central limit theorem. More recently, Ben-Tal et al. (2005) apply the min-max approach to a multi-period stochastic inventory management problem with a flexible commitment contract. Raza (2014) applies the max-min approach to

the newsvendor problem with joint pricing and inventory decisions. Han et al. (2014) enrich the analysis by combining both risk aversion and ambiguity aversion to obtain a closed-form solution. Another approach to making decisions with limited distribution information is the minimax regret principle, which involves minimizing the maximum opportunity cost of not making the optimal decision. For example, Yue et al. (2006) apply this regret approach to a newsvendor problem when only the mean and variance of demand distribution are known. Perakis and Roels (2008) further investigate the minimax regret newsvendor problem by considering more general partial demand information scenarios. Hanasusanto et al. (2015) use a distributionally robust approach to study the multi-item newsvendor problem and obtain a related SDP relaxation for this NP-hard problem.

Whereas most papers investigate the robust decisions from the perspective of a single decision maker, there are several papers examining the robust decisions in a game setting. Jiang et al. (2011) study a robust inventory competition problem with stock-out substitution and asymmetric information on the support of demand distribution, under the criteria of absolute regret minimization. Wagner (2015) employs the robust optimization approach to a supply chain with a price-only contract, where the supplier may not correctly assess the informational state of the retailer. Wagner compares the order quantity and profit of individual firms and the supply chain under different informational states of the two parties and concludes that information advantage does not necessarily benefit a firm. A common feature in the extant literature on robust inventory models is that randomness comes only from demand. This study extends the literature by investigating the robust max-min inventory decision when both demand and price are random.

Another related literature stream is the use of profit or revenue sharing contract in supply chains, where the retailer pays the supplier a unit price, plus a percentage of the profit/revenue generated from sales. Pasternack (2002) investigates a newsvendor problem with both wholesale price and revenue sharing contracts. Cachon and Lariviere (2005) provide a comprehensive study on the use of revenue sharing contract to achieve channel coordination and compare revenue sharing with a number of other supply chain contracts. Tang and Kouvelis (2014) study supply chain coordination in the presence of supply uncertainty, and propose a pay-back-revenue sharing contract to coordinate the supply chain. These papers focus on the design of the contract parameters from a central planner perspective to coordinate the supply chain. Other papers

study revenue sharing contracts in decentralized supply chains. For example, Wang et al. (2004) investigate the channel performance under a consignment contract with revenue sharing, where the retailer, acting as the leader, decides the share of sales revenue to be remitted to the supplier, and the supplier sets the retail price and delivery quantity. Gerchak and Wang (2003) consider an assembly problem in which the firm assembling the final product decides the allocation of sales revenue between the firm and several suppliers producing the components.

Revenue sharing contract has been shown to allow the supply chain profit to be arbitrarily divided between the supplier and retailer. If the two contract parameters, wholesale price and profit share, are both set by one party, then this party can extract most of the supply chain profit, leaving the other party only the reservation profit. Therefore, in this paper, we let each supply chain partner determine one of the two contract parameters. As the supplier incurs the costs of making products, he is in a better position to set the wholesale price. The retailer, as a price-taker, determines the profit share. We model this as a three-stage Stackelberg game. The retailer, acting as the leader, decides a profit share that benefits herself, and then the supplier manipulates his share of the supply chain profit by choosing the wholesale price. Finally, given the agreed-upon contract, the retailer decides the order quantity.

3 The Model

There are two common scenarios to model the market price. One models the case of a monopoly market, where a dominant seller has the power to endogenously set the retail price. Thus, product price is an internal decision, and demand is typically modelled as a price dependent function. Another considers an oligopoly market with a few key competing sellers or a market with a number of small sellers, whose actions jointly determine a competitive market price. In this case, product price is typically exogenous and random in advance, subject to some factors that cannot be observed at the time strategic decisions are made, such as competitors' prices, total market supply and demand. We consider the latter case with both product demand and price being random in our base model, and analyze the monopolistic case in Section 6.

We modify the setup in McMillan and Waxman (2007) to find the optimal split of profit between the retailer and supplier. In this setting, with only the first and second moments

information on the demand and selling price of the product, the retailer would like to devise a profit sharing formula with a key supplier to maximize her own profit. Through sharing, the retailer hopes to procure from the supplier at a lower wholesale price.

To maximize supply chain profit, the retailer would also need to determine the capacity/quantity investment decision, Q , in advance. Let $D(\geq 0)$ denote the hitherto unknown demand and $P_s(\geq 0)$ denote the uncertain selling price. Both P_s and D are unknown when the contractual decisions are made, and the retailer cannot sell more than the installed capacity Q . The capacity is sourced from an external supplier with a unit cost of f . The supplier charges a wholesale price $c(Q) = wQ$, for Q units of capacity. The wholesale price w is a decision by the supplier that the retailer would like to influence through the profit sharing arrangement.

Let Π_C denote the expected profit made by the supply chain in a centralized setting.

$$\Pi_C = E_{P_s, D} \left[\underbrace{P_s}_{\text{Price}} \underbrace{\min(Q, D)}_{\text{Quantity Sold}} \right] - \underbrace{f \times Q}_{\text{Cost}}. \quad (1)$$

In the traditional Nash bargaining approach, the profits allocated to the retailer (Π_R) and supplier (Π_S) are obtained by solving

$$\max_{\alpha} \left\{ (\Pi_R - g)^\delta (\Pi_S - b)^{1-\delta} \mid \Pi_R = \alpha \Pi_C, \Pi_S = (1 - \alpha) \Pi_C \right\},$$

assuming all parameters are known, where g and b are the outside option value for the retailer and supplier respectively. The parameter δ is the relative bargaining power of the retailer. The profit Π_R of the retailer is thus obtained by solving the Nash bargaining solution.

This analysis however requires qualitative judgment of the relative bargaining power of the retailer and the supplier, and ignores the other economic and strategic considerations of the parties involved (e.g., the setting of the wholesale price w by the supplier). It assumes that the bargaining outcome has no effect on the profit Π in the decentralized setting, which is often violated in practice, as the profit share decision affects the capacity investment decision Q made by the firm, and thus affects the profit Π . Furthermore, the supplier can also extract rents through judicious choice of wholesale [price](#) decision w .

In the decentralized setting, consider the net profit of the retailer, denoted by

$$\Pi = E_{P_s, D} \left[P_s \min(Q, D) \right] - w \times Q. \quad (2)$$

The retailer allocates a share γ of the total profit Π to the supplier. In this paper, we model directly the impact of profit share $1 - \gamma$ accrued by the retailer, on the investment decision Q made by the retailer in response to the wholesale price decision w made by the supplier. A very common form of the wholesale price in practice is a linear wholesale price. The popularity of linear contract has been addressed by many papers in the economics literature. For example, Holmstrom and Milgrom (1987) state that the popularity of linear contracts probably lies in its great robustness. A recent paper, Carroll (2015), using a principal-agent framework with worst-case performance objective, shows that out of all possible contracts, a linear contract offers the principal the best guarantee. Therefore, we restrict our analysis to a linear wholesale price. The problem is solved in three stages.

- In Stage 1, the retailer determines the profit share $1 - \gamma$ she would retain on the profit made. The goal is to maximize her share of the profit pie $\Pi_R := (1 - \gamma)\Pi$, allocating $\gamma\Pi$ to the supplier.
- In Stage 2, the supplier determines the wholesale price $w(\geq 0)$ per unit, for the items supplied to the retailer. Note that the actual cost per unit is $f(\geq 0)$ for the supplier. The firm is thus interested in maximizing the total profit

$$\Pi_S := \gamma\Pi + (w - f)Q.$$

- In Stage 3, the retailer chooses the optimal capacity decision Q , to maximize the profit Π , which is given by Equation (2). Any unused capacity will be lost with zero salvage value.

In this way, the protracted bargaining process between the retailer and the supplier on the profit sharing agreement is modelled using a multi-stage Stackelberg game. Our goal in this paper is to develop verifiable relationships between the model parameters and the optimal decisions obtained under the multi-stage Stackelberg game, in a distributionally robust setting.

We are interested in the situation when demand and price distributions are not known precisely at the time the contract is being negotiated. Instead, we only have information on the means, variances and covariance of price P_s and demand D . These parameters are given by a

matrix Σ , called the moments matrix:

$$\Sigma = \begin{pmatrix} E(P_s^2) & E(P_s D) & E(P_s) \\ E(P_s D) & E(D^2) & E(D) \\ E(P_s) & E(D) & 1 \end{pmatrix}.$$

We assume that both parties in the Stackelberg game are ambiguity averse and plan for the worst case. There is some distinction between risk aversion and ambiguity aversion. Risk refers to the situation that the probabilities of possible outcomes are known, i.e., known risks, while ambiguity refers to the situation that the probability distribution of possible outcomes is unknown, i.e., unknown risks. An ambiguity averse individual prefers known risk over unknown risk. In our problem setting, without knowing the price and demand distributions, the retailer and supplier face an ambiguity situation and both plan against the worst-case distribution, which could be characterized.

It is easy to see that Σ is positive semi-definite with nonnegative entries for any random $P_s, D \geq 0$. Throughout this paper, capital letters in bold font denote matrices. Let

$$\Phi(\Sigma_0) = \{(P_s, D) ; P_s, D \geq 0, P_s, D \text{ have moments matrix } \Sigma = \Sigma_0\}$$

denote the space of distributions with the prescribed moments conditions, where Σ_0 is given. All decisions are evaluated against the worst possible distribution in $\Phi(\Sigma_0)$. That is, the first and second moments of the demand and price, $E(P_s), E(D), E(P_s^2), E(D^2), E(P_s D)$, should take values from the entries of Σ_0 .

The distributionally robust Stackelberg game model reduces to the following problem: Let

$$Q(w) := \operatorname{argmax}_{Q \geq 0} \left\{ -wQ + \inf_{(P_s, D) \in \Phi(\Sigma_0)} E_{P_s, D} \left(P_s \min(Q, D) \right) \right\}, \quad (3)$$

$$\Pi(w) := \max_{Q \geq 0} \left\{ -wQ + \inf_{(P_s, D) \in \Phi(\Sigma_0)} E_{P_s, D} \left(P_s \min(Q, D) \right) \right\}. \quad (4)$$

Our goal is to choose the profit sharing formula γ based on solving the following model:

$$(\text{Model } M) : \quad \max_{\gamma \geq 0} (1 - \gamma) \Pi(w(\gamma)), \quad (5)$$

where $w(\gamma)$, for given γ , is chosen to maximize the supplier's profit function:

$$w(\gamma) := \operatorname{argmax}_{w \geq 0} \left[(w - f)Q(w) + \gamma \underbrace{\left\{ \max_{Q \geq 0} \left\{ -wQ + \inf_{(P_s, D) \in \Phi(\Sigma_0)} E_{P_s, D} \left(P_s \min(Q, D) \right) \right\} \right\}}_{\text{Retailer's Stage 3 Problem}} \right] \quad (6)$$

This allows us to study the profit sharing model using a very parsimonious setup that uses only moments information on demand and selling price. Surprisingly, this model permits a very simple solution.

4 The Robust Newsvendor Problem: Closed-Form Solution

In the rest of the paper, we derive formally the main results obtained in this paper. Given the wholesale price w per unit of capacity from the supplier, the retailer's problem in the last stage is to maximize the worst-case profit of the supply chain over the space of demand and price distributions with the moment constraint, by choosing the optimal capacity decision Q . The robust newsvendor problem (PR) can be formulated as:

$$(PR) : \quad \max_{Q \geq 0} \left\{ \Pi = -wQ + \inf_{(P_s, D) \in \Phi(\Sigma_0)} E_{P_s, D} \left(P_s \min(Q, D) \right) \right\}.$$

In the literature (Moon and Gallego (1994), Scarf (1958)), closed-form solution is derived for the robust newsboy problem, when only the demand is random and we have partial information regarding its mean and variance. However, the problem here is more challenging as there are two random variables with only information about their means, variances, and covariance. To tackle the problem, we use an indirect approach by first transforming it into positive semi-definite program.

Let $I(1)$ denote the event $\{D \leq Q\}$ and $I(2)$ denote the event $\{D > Q\}$. Define the variables

$$\begin{pmatrix} y(i) \\ P(i) \\ D(i) \\ PP(i) \\ DD(i) \\ PD(i) \end{pmatrix} = \begin{pmatrix} Prob(I(i)) \\ E(P_s|I(i))Prob(I(i)) \\ E(D|I(i))Prob(I(i)) \\ E(P_s^2|I(i))Prob(I(i)) \\ E(D^2|I(i))Prob(I(i)) \\ E(P_s D|I(i))Prob(I(i)) \end{pmatrix}, \text{ for } i = 1, 2. \quad (7)$$

Using conditional expectation, we can decompose the second stage profit function into

$$\begin{aligned} & E_{P_s, D} \left(\min \{ P_s D, P_s Q \} \right) \\ &= E_{P_s, D} \left(P_s D \mid D \leq Q \right) Prob(D \leq Q) + E_{P_s, D} \left(P_s Q \mid D > Q \right) Prob(D > Q) \\ &= PD(1) + QP(2). \end{aligned} \quad (8)$$

Proposition 1 *The inner problem in the robust newsvendor problem (PR) can be formulated as the following completely positive program² (CPP):*

$$\min_{\mathbf{M}_i, i=1,2} \left\{ PD(1) + QP(2) \right\}$$

subject to

$$\begin{aligned} (A) : \quad & \mathbf{M}_i = \begin{pmatrix} PP(i) & PD(i) & P(i) \\ PD(i) & DD(i) & D(i) \\ P(i) & D(i) & y(i) \end{pmatrix} \succeq 0, i = 1, 2, \\ (B) : \quad & \mathbf{M}_i \geq 0, i = 1, 2, \\ (C) : \quad & \mathbf{M}_1 + \mathbf{M}_2 = \mathbf{\Sigma}_0. \end{aligned}$$

Remark 1 *A symmetric matrix \mathbf{A} of size n is completely positive (denoted $A \succeq_{CP} 0$) if it can be written as $\mathbf{B}\mathbf{B}^T$, where $\mathbf{B} \in \mathbb{R}^{n \times m}$ has nonnegative entries. It is known that if $n \leq 4$, then \mathbf{A} is completely positive if and only if it is positive semi-definite and has all entries nonnegative (Berman and Shaked-Monderer (2003)). Hence, the minimization problem in the above proposition is a completely positive program.*

Proposition 1 implies that the inner problem in the robust problem (PR) can be rewritten as:

$$\min_{\mathbf{M}_i, i=1,2} \text{Tr} \left[\begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{M}_1 + \begin{pmatrix} 0 & 0 & Q/2 \\ 0 & 0 & 0 \\ Q/2 & 0 & 0 \end{pmatrix} \mathbf{M}_2 \right] \quad (9)$$

subject to constraints (A), (B) and (C).

We can find the dual of this completely positive program. Let $\mathbf{Y}_1, \mathbf{Y}_2$ denote the dual variables associated with constraint (B) and \mathbf{Y} denote the dual variable associated with constraint (C). By taking the dual of Problem (9), we have

$$\begin{aligned} \max_{\mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}} \quad & -\text{Tr}(\Sigma_0 \mathbf{Y}) \\ \text{s.t.} \quad & \\ & \mathbf{Y} \succeq - \begin{pmatrix} 0 & 0 & Q/2 \\ 0 & 0 & 0 \\ Q/2 & 0 & 0 \end{pmatrix} + \mathbf{Y}_2, \\ & \mathbf{Y} \succeq - \begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mathbf{Y}_1, \quad \mathbf{Y}_1 \geq 0, \mathbf{Y}_2 \geq 0. \end{aligned} \quad (10)$$

Note that since Problem (10) is convex and strictly feasible, strong duality holds (Boyd and Vandenberghe (2004)), with its optimal value equal to that of Problem (9).

Adding the outer problem, we can obtain the dual CPP formulation of the robust problem (PR) as shown in the following theorem. The dual is again a positive semi-definite programming problem.

Theorem 1 *The robust problem (PR) can be reformulated as a dual CPP as follows:*

$$\begin{aligned}
& \max_{Q \geq 0, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}} && -wQ - \text{Tr}(\boldsymbol{\Sigma}_0 \mathbf{Y}) && (11) \\
& \text{s.t.} && && \\
& \mathbf{Y} \succeq - \begin{pmatrix} 0 & 0 & Q/2 \\ 0 & 0 & 0 \\ Q/2 & 0 & 0 \end{pmatrix} + \mathbf{Y}_2, && && \\
& \mathbf{Y} \succeq - \begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mathbf{Y}_1, && \mathbf{Y}_1 \geq 0, \mathbf{Y}_2 \geq 0. &&
\end{aligned}$$

To simplify the formulation, we perform a linear transformation on \mathbf{Y} in Problem (11), by adding $\begin{pmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ to it, to obtain the following equivalent problem, which we will consider from this point onwards:

$$\begin{aligned}
& \max_{Q \geq 0, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Y}} && -wQ - \text{Tr}(\boldsymbol{\Sigma}_0 \mathbf{Y}) + E(P_s D) && (12) \\
& \text{s.t.} && && \\
& \mathbf{Y} \succeq \frac{1}{2} \begin{pmatrix} 0 & 1 & -Q \\ 1 & 0 & 0 \\ -Q & 0 & 0 \end{pmatrix} + \mathbf{Y}_2, && && \\
& \mathbf{Y} \succeq \mathbf{Y}_1, && \mathbf{Y}_1 \geq 0, \mathbf{Y}_2 \geq 0. &&
\end{aligned}$$

Let $(Q^*, \mathbf{Y}_1^*, \mathbf{Y}_2^*, \mathbf{Y}^*)$ be an optimal solution to Problem (12). Note that Problem (12) is a dual CPP with matrices of size three. Let $\Pi(w)$ denote the optimal objective value to Problem (12).

In the following proposition, we observe an important property of Problem (12).

Proposition 2 *$\Pi(w)$ is convex decreasing in w .*

To solve Problem (12), and thus the robust problem (PR), we first solve a simpler version of the problem, and then use the optimal solution constructed to recover the solution to the

original robust problem (PR). By setting $\mathbf{Y}_1, \mathbf{Y}_2$ in Problem (12) to zero, the simpler problem is obtained as follows:

$$\begin{aligned} \max_{Q \geq 0, \mathbf{Y}} \quad & -wQ - \text{Tr}(\boldsymbol{\Sigma}_0 \mathbf{Y}) + E(P_s D) \\ \text{s.t.} \quad & \\ & \mathbf{Y} \succeq \frac{1}{2} \begin{pmatrix} 0 & 1 & -Q \\ 1 & 0 & 0 \\ -Q & 0 & 0 \end{pmatrix}, \quad \mathbf{Y} \succeq 0, \end{aligned} \quad (13)$$

Let (Q_r^*, \mathbf{Y}_r^*) be an optimal solution to Problem (13). For ease of exposition, we define the following terms:

Definition 1 Define $\alpha(w) = E(P_s)/2 - w$, $\beta = E(P_s^2)/4$.

Theorem 2 For Problem (13), the optimal solution Q_r^* is greater than zero if and only if

$$w < \frac{1}{2} \left[\sqrt{\frac{E(P_s^2)}{E(D^2)}} E(D) + E(P_s) \right], \quad (14)$$

in which case,

$$Q_r^* = E(D) + \frac{\alpha(w)}{\sqrt{\beta - \alpha(w)^2}} \sigma(D), \quad (15)$$

and the worst-case profit is

$$\begin{aligned} \Pi_{wst} &= -wQ_r^* + \inf_{(P_s, D) \in \Phi(\boldsymbol{\Sigma}_0)} E(P_s \min(D, Q_r^*)) \\ &= \alpha(w)E(D) - \sigma(D)\sqrt{(\beta - \alpha(w)^2)} + \frac{1}{2}E(P_s D), \end{aligned} \quad (16)$$

where $\sigma(D)$ is the standard deviation of demand D .

Remark 2 Theorem 2 in fact presents a closed-form expression for the robust newsvendor problem when P_s and D are not restricted to be nonnegative. In this case, when the condition (14) is not met, namely, w is larger or equal to the given threshold, no investment will be made. This places a bound on the wholesale price set by the supplier. Otherwise, we can obtain the optimal capacity, Q_r^* , and the resulting worst-case profit, Π_{wst} , in closed-form.

In the following theorem, we recover the solution to Problem (12) (and the original robust capacity investment problem (PR)), using the dual constructed from this simpler problem, Problem (13).

Theorem 3 *For the robust problem (PR), if*

$$w \leq \frac{1}{2} \left[E(P_s) + \frac{E(P_s D)E(D) - \sigma(D)\sqrt{E(P_s^2)E(D^2) - E(P_s D)^2}}{E(D^2)} \right], \quad (17)$$

the optimal solution Q^ is given as:*

$$Q^* = E(D) + \frac{\alpha(w)}{\sqrt{\beta - \alpha(w)^2}} \sigma(D), \quad (18)$$

with the worst-case profit $\Pi(w) = \Pi_{wst}$ being the same as that presented in Theorem 2 by replacing Q_r^ with Q^* . Otherwise $Q^* = 0$.*

Definition 2 *Define $w^{UB} = \frac{1}{2} \left[E(P_s) + \frac{E(P_s D)E(D) - \sigma(D)\sqrt{E(P_s^2)E(D^2) - E(P_s D)^2}}{E(D^2)} \right]$.*

We assume the parameters are chosen such that $w^{UB} > f$. From Theorem 3, the optimal solution Q^* for the original problem is the same as that of the simpler problem, Problem (13), but with a different *threshold* condition on w . We can check that the upper bound w^{UB} here is lower than the upper bound given in Theorem 2, so the non-negative constraints on P_s, D reduce the feasible region of w . The reason is that for the simpler problem, $\mathbf{Y}_1, \mathbf{Y}_2$ are set to zero, which reduces the feasible region. An interesting observation is that the optimal robust solution Q^* only depends on the first and second moments of the demand and selling price distributions. *The correlation term $E(P_s D)$ does not affect the optimal choice of Q in the distributionally robust problem.* If the threshold condition (17) is not satisfied, the retailer is better off not ordering anything.

Scarf (1958) obtains the min-max solution for a newsvendor problem with unit cost c , and unit selling price p . Note that the newsvendor problem is a special case of our problem, with P_s being constant. In this case, let $P_s = p$ and $w = c$. It can be verified that our result in Theorem 3 reduces to Scarf's formula, with exactly the same Q^* and the corresponding condition. Thus our result generalizes Scarf's formula to a more general setting when both demand and price are random.

The following useful inequality (Scarf's bound) is well known:

$$E\left[(D - Q)^+\right] \leq \frac{1}{2}\left[E(D) - Q + \sqrt{(E(D) - Q)^2 + (E(D^2) - E(D)^2)}\right].$$

This bound has found applications in a variety of domains from finance to supply chain management. As a by-product, our analysis actually yields an extension of this classical inequality to the case when the price P_s is also random.

Corollary 1 *If price P_s and demand D are both random, the following bound holds:*

$$E\left[P_s(D - Q)^+\right] \leq \frac{1}{2}\left[E(P_s D) - QE(P_s) + \sqrt{E(P_s^2)}\sqrt{(E(D) - Q)^2 + (E(D^2) - E(D)^2)}\right].$$

The above bound follows from Proposition 1 and the proof of Theorem 2.

It is well known in the literature that for the min-max newsvendor problem, the worst possible distribution of demand only has positive mass at two points. We find that the worst-case distribution for the robust problem (PR) is a three-point distribution as characterized in Proposition 6 in the Appendix. We verify that under the special case when P_s becomes a constant, our three-point worst-case distribution reduces to the two-point distribution of Gallego and Moon (1993) (Lemma 2, Remark 2). Hence, our three-point distribution generalizes the literature on the min-max newsvendor problem to a more challenging setting when the random elements are two dimensional (demand D and price P_s).

The min-max order quantity obtained by Scarf (1958) has been pointed out to be too conservative, because the order quantity would be zero under certain condition. However, as shown in Perakis and Roels (2008) and Wagner (2015), when we focus on the range of contract prices that induce a positive order quantity, the min-max order quantity is not much different than the risk-neutral order quantities under well-known distributions. For our robust newsvendor problem with both demand and price being random, we also make the comparison. Consider the following problem setting: $E(D) = 100$, $\sigma(D) = 30$, $E(P_s) = 40$, $\sigma(P_s) = 15$, $\rho = 0.5$, $f = 5$. Figure 1 compares the robust order quantity given by Equation (18) with the risk-neutral order quantity obtained assuming bivariate normal distribution with the above given moments, by changing the wholesale price w . In this setting, the wholesale price should be in the range of production cost f and the expected selling price $E(P_s)$. However, as the upper bound on the wholesale price to guarantee a positive robust order quantity is $w^{UB} = 36.79$, we will restrict w

in this range. In fact, for many problem cases, w^{UB} is found to be quite close to $E[P_s]$, so only when the wholesale price is close to the selling price, it will result in not ordering in the robust setting. We can observe from the comparison that the two order quantities are quite close, with the robust quantity slightly lower. In the following sections, we will consider the case that the supplier chooses a wholesale price that induces a positive order quantity from the retailer.

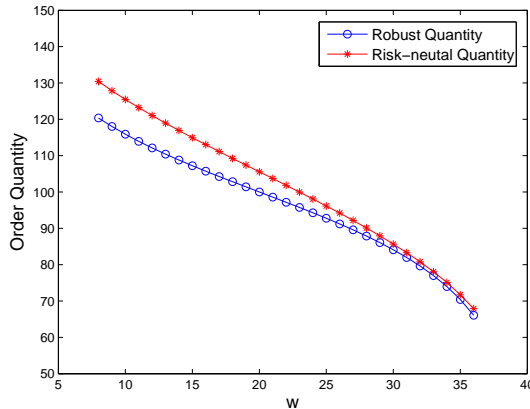


Figure 1: Comparison between the robust and the risk-neutral order quantities

5 The Profit Sharing Contract Design Problem

For the supplier, in anticipation of the retailer's robust decision $Q(w)$ for given wholesale price w , his ambiguity-averse solution is to choose an optimal wholesale price $w \geq f$, for the given profit share γ , such that his worst-case profit Π_S is maximized, where

$$\begin{aligned} \Pi_S &= (w - f)Q(w) + \gamma\Pi(w) \\ &= \begin{cases} (w - f)Q(w) + \gamma \left[-wQ(w) + \inf_{(P_s, D) \in \Phi(\Sigma_0)} E_{P_s, D} \left(P_s \min\{D, Q(w)\} \right) \right], & \text{if } w \leq w^{UB} \\ 0, & \text{if } w > w^{UB} \end{cases} \end{aligned}$$

Note that

$$\begin{aligned} \Pi(w) &= -wQ(w) + \inf_{(P_s, D) \in \Phi(\Sigma_0)} E_{P_s, D} \left(P_s \min\{D, Q(w)\} \right) \\ &= \alpha(w)E(D) - \sigma(D)\sqrt{\beta - \alpha(w)^2} + \frac{1}{2}E(P_s D), \end{aligned}$$

for $w \leq w^{UB}$, by Theorem 3.

Let $\Pi_1(w) = (w - f)Q(w)$. The supplier's worst-case profit, under the profit sharing agreement, is therefore given by

$$\Pi_S(w, \gamma) = \Pi_1(w) + \gamma\Pi(w), \quad \text{for } w \leq w^{UB}.$$

To determine the optimal wholesale price w , the supplier needs to solve

$$w(\gamma) := \operatorname{argmax}_{f \leq w \leq w^{UB}} \left\{ \Pi_1(w) + \gamma\Pi(w) \right\}. \quad (19)$$

Figure 2 presents the general shape of the supplier's revenue (when $f = 0$) as a function of the wholesale price w and profit share parameter γ . For each γ , finding the corresponding optimal wholesale price for the supplier is a complex non-linear problem.

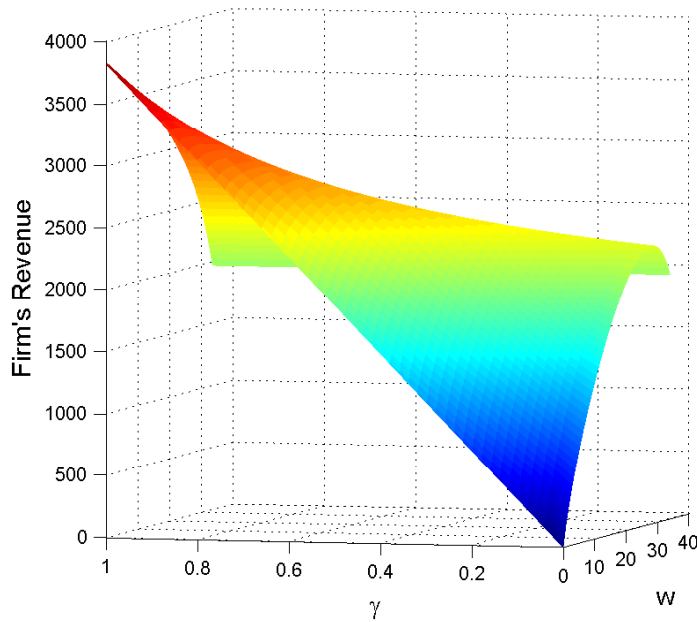


Figure 2: The worst-case revenue of the supplier as a function of (γ, w)

5.1 Relationship between w and Q , γ

We now examine how the optimal wholesale price w changes with γ . Note that when $\gamma = 1$, the supply chain is vertically integrated and the supplier chooses $w = f$ to maximize the supply

chain profit³. On the other hand, when $\gamma = 0$, the problem reduces to the classical wholesale price optimization problem. For intermediate value of γ , we can project out the inner retailer's optimization problem in the supplier's problem, to solve a single variable optimization problem as shown in Equation (19) to obtain $w(\gamma)$ as a function of the profit sharing formula γ .

We present the following propositions.

Proposition 3 $Q(w)$ is strictly decreasing in $w \in [f, w^{UB}]$.

This proposition is in line with our intuition that the retailer orders less, if the wholesale price increases.

Proposition 4 (a) $w(\gamma) > f$ for all $\gamma \in [0, 1)$.

(b) The supplier's optimal wholesale price $w(\gamma)$ is decreasing in $\gamma \in [0, 1)$.

Part (a) of Proposition 4 indicates that in the distributionally robust setting under a profit sharing contract, the wholesale price charged by the supplier is strictly larger than his production cost. Part (b) states that the two contract parameters, w and γ move in opposite direction. From Proposition 2, we know that the total profit, $\Pi(w)$, decreases in the wholesale price, so the intuition here is that if the supplier gets a larger profit share γ , he will charge a lower wholesale price w to induce the retailer to order more to increase the overall profit.

5.2 Relationship between γ and Q

In anticipation of the supplier's optimal choice of the wholesale price $w^*(\gamma)$, the retailer's problem is to determine the optimal profit share parameter γ^* to maximize her worst-case profit given by:

$$\Pi_R = (1 - \gamma)\Pi(w(\gamma)).$$

By Proposition 2, $\Pi(w)$ decreases with w . On the other hand, $w(\gamma)$ is decreasing in γ by Proposition 4. Hence $\Pi(w(\gamma))$ is increasing in γ . More interestingly, the profit $(1 - \gamma)\Pi(w(\gamma))$ is the product of two functions, one decreasing in γ , and another increasing in γ . Intuitively, the optimal profit share parameter γ^* for the retailer should be an interior solution. However, as there is no closed-form expression for $w^*(\gamma)$, we cannot write the retailer's profit Π_R as an

explicit function of the profit share γ , so finding the optimal γ^* turns out to be a challenging problem and can only be obtained through numerical methods.

We solve the problem indirectly by addressing the following question next: *How should the retailer choose the right profit share parameter γ to ensure that the supplier will respond with a wholesale price $w(\gamma)$ such that $Q = Q(w(\gamma))$?*

Let $\alpha(w) = E(P_s)/2 - w$. For given wholesale price $w \leq w^{UB}$, the optimal order quantity is

$$Q(w) = E(D) + \sigma(D) \frac{\alpha(w)}{\sqrt{\beta - \alpha(w)^2}}.$$

We can invert the above expression to obtain

$$\alpha(Q) = \begin{cases} \sqrt{\beta} / \sqrt{1 + \left(\frac{\sigma(D)}{Q-E(D)}\right)^2}, & \text{if } Q \geq E(D), \\ -\sqrt{\beta} / \sqrt{1 + \left(\frac{\sigma(D)}{Q-E(D)}\right)^2}, & \text{if } Q < E(D). \end{cases}$$

We know that the retailer will choose capacity investment level of Q if the supplier sets the wholesale price at

$$w(Q) = E(P_s)/2 - \alpha(Q) = \begin{cases} \frac{E(P_s)}{2} - \sqrt{\beta} / \sqrt{1 + \left(\frac{\sigma(D)}{Q-E(D)}\right)^2}, & \text{if } Q \geq E(D), \\ \frac{E(P_s)}{2} + \sqrt{\beta} / \sqrt{1 + \left(\frac{\sigma(D)}{Q-E(D)}\right)^2}, & \text{if } Q < E(D). \end{cases}$$

On the other hand, the profit seeking supplier will set the wholesale price to be w^* if it optimizes its profit Π_S . Note that the first order condition of optimization problem (19) is necessary (but may not be sufficient) for optimality, and it has to hold at w^* . We have $\gamma(Q) = 0$ if and only if $w^* = w_0$, where w_0 ($f < w_0 \leq w^{UB}$) is defined in Definition 3 below, while $\gamma(Q) = 1$ if and only if $w^* = f$. For $\gamma(Q) \in (0, 1)$, by Proposition 4, w^* is in the interior of f and w^{UB} . In this case, we have w^* satisfies the first order optimality condition:

$$\frac{\partial \Pi_S}{\partial w}(w^*, \gamma) = \frac{d\Pi_1}{dw}(w^*) + \gamma \frac{d\Pi}{dw}(w^*) = 0.$$

Hence, to induce the profit seeking supplier to set a wholesale price of $w^* = w(Q)$, the retailer only needs to choose

$$\gamma(Q) = -\frac{\frac{d\Pi_1}{dw}(w(Q))}{\frac{d\Pi}{dw}(w(Q))}. \quad (20)$$

This set of parameters will ensure that the retailer's order quantity in the Stackelberg game is precisely Q . An explicit expression for $\gamma(Q)$ is given in Theorem 4 below. Note that in the above arguments, $w(Q)$ must lie in $[f, w_0]$, while Q must lie in $[Q_l, Q_u]$, where w_0, Q_l and Q_u are defined as follows:

Definition 3 Define $w_0 = w(\gamma = 0)$ and $Q_u = Q(f), Q_l = Q(w_0)$.

Putting everything together, we have the following main result:

Theorem 4 The outcome of the Stackelberg game is $Q \in [Q_l, Q_u]^4$ when the retailer chooses

$$\gamma = 1 - \frac{(E(P_s)/2 - f - \alpha(Q))\sigma(D)\beta}{(\beta - \alpha(Q)^2)^{3/2}Q}, \text{ for } Q \in (Q_l, Q_u)$$

where

$$\alpha(Q) = \begin{cases} \sqrt{\beta} / \sqrt{1 + \left(\frac{\sigma(D)}{Q - E(D)}\right)^2}, & \text{if } Q \geq E(D), \\ -\sqrt{\beta} / \sqrt{1 + \left(\frac{\sigma(D)}{Q - E(D)}\right)^2}, & \text{if } Q < E(D). \end{cases}$$

For $Q = Q_l$, then choose $\gamma = 0$ and for $Q = Q_u$, then choose $\gamma = 1$.

This result is interesting as it posits a non-linear connection between $1 - \gamma$, the profit share accrued by the retailer, and the level of capacity investment Q by the retailer and the wholesale price w by the supplier involved in the supply chain.

We can now answer the main question posed in this paper: *How should the retailer determine the profit share parameter γ to maximize the profit the retailer can obtain?* Instead of optimizing over γ , which is difficult, we optimize over α . For a given α , where $\alpha = E(P_s)/2 - w$, the profit that the retailer obtains from the supply chain can be rewritten as a function of α by substituting $\Pi = \Pi_{wst}$:

$$(1 - \gamma)\Pi = \frac{(E(P_s)/2 - f - \alpha)\sigma(D)\beta}{\left(E(D) + \frac{\alpha}{\sqrt{\beta - \alpha^2}}\sigma(D)\right)(\beta - \alpha^2)^{3/2}} \times \left(\alpha E(D) - \sigma(D)\sqrt{\beta - \alpha^2} + E(P_s D)/2\right). \quad (21)$$

By optimizing the above function over $\alpha \in \left[\frac{E(P_s)}{2} - w_u, \frac{E(P_s)}{2} - w_l\right]$ to obtain α^* , we can set the optimal profit share parameter γ^* as

$$\gamma^* = 1 - \frac{(E(P_s)/2 - f - \alpha^*)\sigma(D)\beta}{\left(E(D) + \frac{\alpha^*}{\sqrt{\beta - (\alpha^*)^2}}\sigma(D)\right)(\beta - (\alpha^*)^2)^{3/2}}. \quad (22)$$

The retailer’s profit in Equation (21) is a complicated function in α . The following example shows the general shape of the retailer’s profit function with the change of α and f .

Example 1 Consider the case when demand and price are independent, with $E(D) = 100$, $\sigma(D) = 50$, $E(P_s) = 40$, $\sigma(P_s) = 15$, $\gamma = 0.2$. We can plot the function $(1 - \gamma)\Pi$ in terms of the parameter f and α in Figure 3.

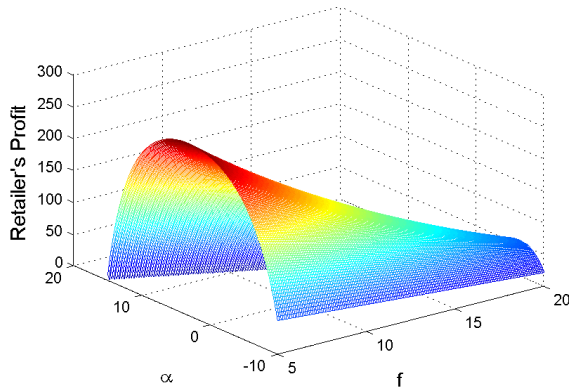


Figure 3: The retailer’s profit as a function of α and f

6 Price Dependent Stochastic Demand Model

In the previous analysis, we focus on the model in which both the demand and selling price are exogenous and random, and can be arbitrarily correlated. The general moments matrix captures the price and demand correlation with the cross moment term $E[P_s D]$. In this section we investigate the case that the selling price is an internal decision of firms as in monopoly markets. In the literature, price dependent stochastic demand models have been widely used to address the joint pricing and inventory decisions (readers may refer to Petruzzi and Dada (1999) for an excellent review of earlier works along this line of research). Raza (2014) investigates the distribution-free approach to the newsvendor problem with pricing decision. With linear price dependent demand model, the author shows that the profit function is quasi-concave in price and demand, and proposes a sequential search process to solve the optimal price and robust inventory decisions. To internalize the pricing decision, we model explicitly the dependency between the

selling price and demand by a linear price dependent stochastic demand function, and show that this problem is a special case of our general model.

To see the connection, consider the case when both demand and selling price are functions of a common factor τ , say $P_s(\tau) = \tau$, and $D(\tau) = a - b\tau + \varepsilon$, where ε is a random component with zero mean and standard deviation σ . The objective is to jointly determine the selling price P_s , i.e. τ , and the order quantity Q to maximize the resulting worst-case expected profit in the distributionally robust setting. This setup implies the following conditions on the first two moments of P_s and D :

$$\begin{aligned} E(P_s) &= \tau \\ E(P_s^2) &= \tau^2 \\ E(D) &= a - b\tau \\ E(D^2) &= (a - b\tau)^2 + \sigma^2 \\ E(P_s D) &= \tau(a - b\tau). \end{aligned}$$

The robust order quantity $Q^*(\tau)$ can be solved by Equation (15) for moments given in the above equations:

$$Q(\tau, w) = a - b\tau + \frac{\tau/2 - w}{\sqrt{\tau^2/4 - (\tau/2 - w)^2}}\sigma,$$

and the worst-case expected profit is given by Equation (16) as follows:

$$\Pi(\tau, w) = (\tau/2 - w)(a - b\tau) - \sigma\sqrt{(\tau^2/4 - (\tau/2 - w)^2)} + \frac{1}{2}\tau(a - b\tau).$$

The optimal price can be solved by the first order condition of function $\Pi(\tau, w)$ for each w . This recovers the solution obtained by Raza (2014) for the pricing problem with limited distribution information, when price and demand are linearly related.

Our approach actually recovers more. By Equations (21) and (22), the retailer can actually jointly optimize the price and the profit sharing formula as follows:

$$\max_{\alpha, \tau} \left\{ \frac{\sigma \frac{\tau^2}{4} \left(\tau/2 - f - \alpha \right)}{\left(a - b\tau + \sigma \frac{\alpha}{\sqrt{\frac{\tau^2}{4} - \alpha^2}} \right) \sqrt{\left(\frac{\tau^2}{4} - \alpha^2 \right)^3}} \times \left((a - b\tau)\alpha - \sigma \sqrt{\frac{\tau^2}{4} - \alpha^2} + \tau(a - b\tau)/2 \right) \right\} \quad (23)$$

to obtain the optimal α^* and τ^* , and the optimal profit sharing formula γ^* is then obtained by

$$\gamma^* = 1 - \frac{\sigma \frac{(\tau^*)^2}{4} \left(\tau^*/2 - f - \alpha^* \right)}{Q(\tau^*, \tau^*/2 - \alpha^*) \sqrt{\left(\frac{(\tau^*)^2}{4} - (\alpha^*)^2 \right)^3}}. \quad (24)$$

7 Numerical Experiments

We use a set of numerical experiments to understand how the proposed technique works. The following common parameters are used in the example: the mean and standard deviation of the demand and price distributions are as follows, $E(D) = 100$, $\sigma(D) = 50$, $E(P_s) = 40$, $\sigma(P_s) = 15$. Since the optimal capacity investment decision is independent of the demand-price correlation, without loss of generality, we set the correlation coefficient $\rho = 0.5$, and the partial demand and price information is given by the following moments matrix:

$$\Sigma = \begin{pmatrix} E(P_s^2) & E(P_s D) & E(P_s) \\ E(P_s D) & E(D^2) & E(D) \\ E(P_s) & E(D) & 1 \end{pmatrix} = \begin{pmatrix} 1825 & 4375 & 40 \\ 4375 & 12500 & 100 \\ 40 & 100 & 1 \end{pmatrix}.$$

Figure 4 depicts the supplier's optimal choice of wholesale price $w^* = w(\gamma)$ as functions of the profit sharing parameter γ . As expected, w^* decreases with γ , reflecting the trade-off between the supplier's two sources of profit. The supplier charges a higher wholesale price w when it takes a smaller share of the profit made by the retailer. In addition, the optimal wholesale price increases with the supplier's internal capacity cost f , for fixed γ . This makes sense, as the supplier needs to charge a higher wholesale price to recover the higher capacity cost.

Figure 5 shows the retailer's optimal capacity investment decision $Q^* = Q(w(\gamma))$ as functions of the profit share parameter γ with the corresponding wholesale price chosen optimally by the supplier. We can observe that Q^* increases with γ , since the wholesale price decreases with γ , and Q^* decreases with the wholesale price. Furthermore, Q^* decreases with the supplier's capacity cost f for fixed γ , due to the higher wholesale price charged by the supplier for higher capacity cost.

Figure 6 and Figure 7 present how the worst-case profits of the supplier and the retailer change with γ . Interestingly, the supplier's worst-case profit increases with γ , so the supplier

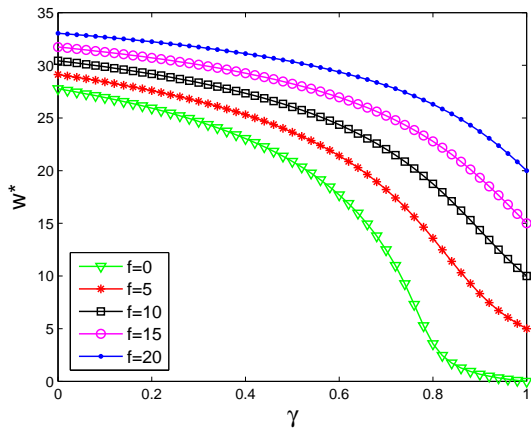


Figure 4: The supplier's optimal wholesale price w^* as functions of γ

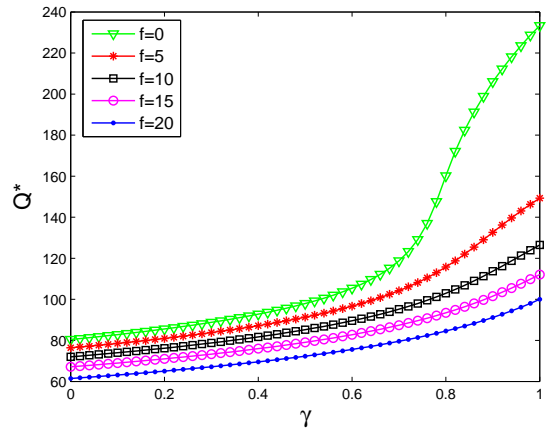


Figure 5: The retailer's optimal capacity investment quantity Q^* as functions of γ

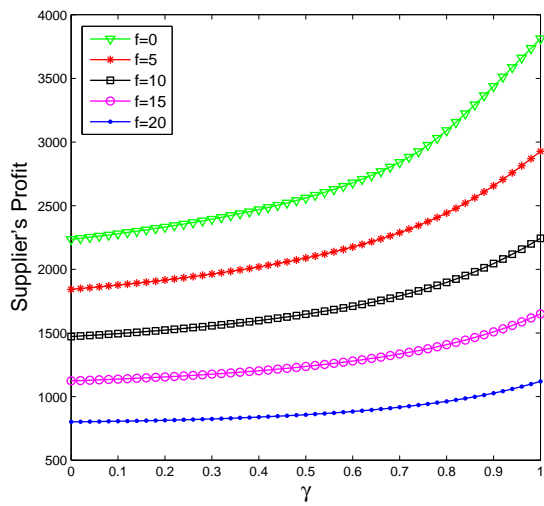


Figure 6: The supplier's worst-case profit as functions of γ

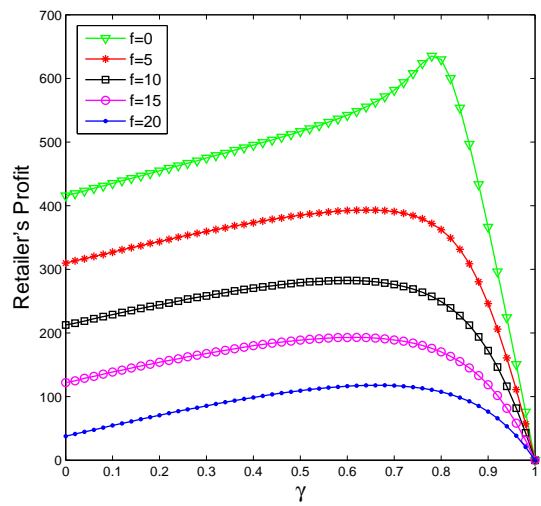


Figure 7: The retailer's worst-case profit as functions of γ

prefers the profit share γ to be as large as possible. However, the retailer's profit is unimodal, attaining the maximum for some γ strictly between 0 and 1. Thus, the retailer would select the profit sharing parameter γ corresponding to this maximum point on its profit performance curve. This point would predict the equilibrium of the three-stage Stackelberg game. It is not surprising that the retailer's profit decreases with the unit cost f .

8 Conclusion

In this paper, we solve the profit sharing agreement problem in an ambiguity averse supply chain with price and demand uncertainty. We formulate the problem as a Stackelberg game, and show that compared with the case without profit sharing ($\gamma = 0$, i.e., the wholesale price case), the Stackelberg equilibrium chooses a profit share parameter γ^* in $(0, 1)$, such that (i) the retailer has a higher worst-case profit, and (ii) the supplier has a higher worst-case profit, and hence the supply chain is more efficient. If the profit share of the supplier is smaller than this threshold γ^* , then both the retailer and the supplier will be worse off in terms of the profit accrued to each of them. On the other hand, a share higher than γ^* accrued to the supplier will only benefit the supplier, but hurt the retailer. Thus a careful calibration of the contractual elements in the profit sharing agreements is crucial for the retailer. Furthermore, the worst-case profit for both the retailer and the supplier dominates the pure wholesale price model (with no profit sharing). These results indicate that in an ambiguity averse supply chain, the profit sharing agreement approach is generally more beneficial to both parties involved, compared with the traditional wholesale price contract.

Our paper also develops new analytical insights on the profit share parameter and the optimal capacity invested in the supply chain. By projecting out the role of wholesale price w , we show that the profit share parameter γ and the corresponding optimal order decision Q should satisfy the relationship

$$1 - \gamma = \frac{\left(E(P_s)/2 - f - \alpha(Q) \right) \sigma_D(E(P_s^2)/4)}{(E(P_s^2)/4 - \alpha(Q)^2)^{3/2} Q},$$

where

$$\alpha(Q) = \begin{cases} \sqrt{E(P_s^2)/4} / \sqrt{1 + \left(\frac{\sigma_D}{Q-E(D)}\right)^2}, & \text{if } Q \geq E(D) \\ -\sqrt{E(P_s^2)/4} / \sqrt{1 + \left(\frac{\sigma_D}{Q-E(D)}\right)^2}, & \text{if } Q < E(D) \end{cases}$$

The results obtained in this paper generalize a long series of work in the area of **distribution-free** newsvendor problem (initiated by Scarf (1958)), including the price optimization variant studied in Raza (2014). These results can also be used to check whether the outcomes (ordering decision, wholesale price) are consistent with a distributionally robust Stackelberg game model, and can be extended to handle price optimization when the mean demand is related to price through a factor model, and to deal with information asymmetry in case the supplier and retailer possess different beliefs on the moment conditions on demand and selling price.

The model proposed in this paper can be extended in different directions. In our approach, we have assumed that the supplier uses a linear wholesale price w to sell to the retailer. In practice, the supplier can also offer quantity discount or other kind of contracts to extract the appropriate profits from the retailer. This can be handled by splitting the order size into distinct segments, with a different wholesale price for each segment. Optimizing the wholesale prices over the distinct segments turns out to be technically challenging, and we are not aware of a tractable solution to this variant of the problem considered in this paper.

Endnotes

1. <http://carnegieendowment.org/2011/01/13/how-to-stop-rise-in-food-price-volatility>
2. The study of completely positive programs has gained considerable interests in recent years. Relevant literature includes Burer (2012) and Dür (2010).
3. In this case, note that the retailer will be indifferent to the choice of w and Q , since it gets no share of the profit. We assume the solution concept adopted is the Strong Stackelberg Equilibrium, i.e., the retailer will pick Q to maximize the payoff of the supplier, if there are multiple optimal solutions available for the retailer.
4. The outcome of the Stackelberg game Q cannot lie outside $[Q_l, Q_u]$ when each party seeks to maximize its worst-case profit.

References

- [1] Afriat, S. N. 1967. The construction of utility functions from expenditure data. *International Economic Review*, **8**(1), 67-77.
- [2] Aghassi, M., D. Bertsimas. 2006. Robust game theory. *Mathematical Programming*. **107**, 231-273.
- [3] Ben-Tal, A., B. Golany, A. Nemirovski, J.-P. Vial. 2005. Retailer-Supplier Flexible Commitments Contracts: A Robust Optimization Approach. *Manufacturing and Service Operations Management*, **7**(3), 248-271.
- [4] Berman, A., N. Shaked-Monderer. 2003. *Completely Positive Matrices*. World Scientific Publishing.
- [5] Boyd, S., L. Vandenberghe. 2004. *Convex Optimization*. Cambridge University Press, Cambridge, UK.
- [6] Brown, D., R. Matzkin. 1996. Testable restrictions on the equilibrium manifold, *Econometrica*, **64**(6), 1249-1262.
- [7] Burer, S. 2012. Copositive programming. In: Anjos, M. F., J. B. Lasserre (eds). *Handbook on Semidefinite, Conic and Polynomial Optimization*. Springer. 201-218.
- [8] Cachon, G. P., M. A. Lariviere. 2005. Supply chain coordination with revenue-sharing contracts: strengths and limitations. *Management Science*, **51**(1),30-44.
- [9] Carroll, G. 2015. Robustness and Linear Contracts. *American Economic Review*. **105**(2), 536-563.
- [10] Carvajal, A., R. Deb, J. Fenske, J. K.-H. Quah. 2013. Revealed preference tests of the Cournot model *Econometrica*. **81**(6), 2351-2379.
- [11] Dúr M. 2010. Copositive programming - A survey. In: Diehl, M., F. Glineur, E. Jarlebring, W. Michiels (eds). *Recent Advances in Optimization and its Applications in Engineering*. Springer Berlin Heidelberg. 3-20.

- [12] Filson, D., D. Switzer, P. Besock. 2005. At the movies: The economics of exhibition contracts. *Economic Inquiry*. **43**(2), 354-369.
- [13] Gallego, G. 1992. A minimax distribution-free for the (Q,R) inventory model. *Operations Research Letters*. **11**, 55-60.
- [14] Gallego, G., I. Moon. 1993. The distribution free newsboy problem: Review and extensions. *Journal of the Operational Research Society*. **44**(8), 825-834.
- [15] Gallego, G., J. K. Ryan, D. Simchi-Levi. 2001. Minimax analysis for finite-horizon inventory models. *IIE Transactions*. **33**, 861-874.
- [16] Han, Q., D. Du, L. F. Zuluaga. 2014. Technical Note - A risk- and ambiguity-averse extension of the max-min newsvendor order formula. *Operations Research*. **62**(3), 535-542.
- [17] Hanasusanto, G. A., D. Kuhn, S.W. Wallace, S. Zymler. 2015. Distributionally robust multi-item newsvendor problems with multimodal demand distributions, *Mathematical Programming*, **152**(1), 1-32.
- [18] Holmström, B., P. Milgrom. 1987. Aggregation and Linearity in the Provision of Intertemporal Incentives. *Econometrica*, **55**(2), 303-328.
- [19] Jiang, H., S. Netessine and S. Savin. 2011 Technical note: robust newsvendor competition under asymmetric information. *Operations Research*. **59**(1), 254-261.
- [20] Lariviere E. L. and M. Porteus. 2001. Selling to the newsvendor: an analysis of price-only contracts, *Manufacturing & Service Operations Management*. **3**, 293-305.
- [21] Mamani, H., S. Nassiri, M. Wagner. 2016. Closed-Form Solutions for Robust Inventory Management, *Management Science*, forthcoming.
- [22] McMillan, M. S., A. Waxman. 2007. Profit Sharing Between Governments and Multinationals in Natural Resource Extraction: Evidence From a Firm-Level Panel. NBER Working Paper No. 13332.
- [23] Moon, I., S. Choi. 1995. The distribution free newsboy problem with balking. *Journal of the Operational Research Society*, **46**(4), 537-542.

- [24] Moon, I, S. Choi. 1997. Distribution free procedures for make-to-order, make-in-advance, and composite policies. *International Journal of Production Economics*, **48**, 21-28.
- [25] Moon, I., G. Gallego. 1994. Distribution free procedures for some inventory models. *Journal of the Operational Research Society*, **45**(6), 651-658.
- [26] Moon, I, E. A. Silver. 2000. The multi-item newsvendor problem with a budget constraint and fixed ordering costs. *Journal of the Operational Research Society*, **51**, 602-608.
- [27] Pasternack, B. A., 2002. Using revenue sharing to achieve channel coordination for a newsboy type inventory model. In: Geunes, J., P. M. Pardalos, H. E. Romeijn (eds.), *Supply Chain Management: Models, Applications, and Research Directions*, Kluwer Academic Publishers, London, 117-136.
- [28] Perakis, G., G. Roels. 2008. Regret in the newsvendor model with partial information. *Operations Research*, **56**(1), 188-203.
- [29] Petruzzi, N. C., M. Dada, (1999) Pricing and the Newsvendor Problem: A Review with Extensions. *Operations Research* **47**(2),183-194.
- [30] Raza, S. A. 2014. A distribution free approach to newsvendor problem with pricing. *4OR - A Quarterly Journal of Operations Research*, **12**(4), 335-358.
- [31] Scarf, H. 1958. A min-max solution to an inventory problem. *Studies in the Mathematical Theory of Inventory and Production*. Stanford University Press, Stanford, CA. 201-209.
- [32] Stiglitz, J.E. 1989. Sharecropping. In: Eatwell, J., M. Milgate, P. Newman (eds). *Economic Development*. The New Pelgrave. Macmillan, London.
- [33] Tang, S. Y., P. Kouvelis. 2014. Pay-Back-Revenue-Sharing Contract in Coordinating Supply Chains with Random Yield. *Production and Operations Management*. **23**(12), 2089-2102.
- [34] Topkis, D. M. 1998. *Supermodularity and Complementarity*. Princeton University Press, Princeton, NJ.
- [35] Wagner, M. 2015. Robust Policies and Information Asymmetry in Supply Chains with a Price-Only Contract, *IIE Transactions*, **47**(8), 819-840.

- [36] Wang, Y., L. Jiang, Z. J. Shen. 2004. Channel Performance Under Consignment Contract with Revenue Sharing. *Management Science*, **50**(1),34-47.
- [37] Wang, Y., Y. Gerchak. 2003. Capacity games in assembly systems with uncertain demand. *Manufacturing and Service Operations Management*, **5**(3), 252-267.
- [38] Weinstein M. 1998. Profit-sharing contracts in Hollywood: Evolution and analysis. *Journal of Legal Studies*, **27**, 67-112.
- [39] Yue, J. F., B. T. Chen, M. C. Wong. 2006. Expected value of distribution information for the newsvendor problem, *Operations Research*, **54**(6), 1128-1136.

Appendix A: Applications

1. Information Asymmetry

Our model can be extended to incorporate information asymmetry in the decentralized supply chain. This is motivated by Wagner (2015), who studied its impact in a decentralized supply chain with price-only contract. In Wagner's model, the supplier and retailer may know either the full distribution or just the first and second moments of the random demand, with the price fixed. Our model with profit sharing is considerably more difficult, because the different information set will lead to different assessment of the newsvendor profit, which will affect how the players share the profits.

More precisely, we consider the case that the supplier and the retailer possess asymmetric information about the market demand and selling price, i.e., the moments matrix Σ . That is, the retailer knows $\Sigma = \Sigma_0$, and is aware that the supplier only knows $\Sigma \in \Theta$ for some set Θ .

Let w_{Σ}^{UB} , $Q_{\Sigma}(w)$ and $\Pi_{\Sigma}(w)$ denote the upper bound for the wholesale price, the capacity decision of Equation (18) and the worst-case profit of Equation (16) as defined in Section 3, for any $\Sigma \in \Theta$. In this case, without knowing the actual value of Σ , the supplier solves for the optimal $w_{\Theta}(\gamma)$ against the worst possible case in the set Θ , for given γ , to maximize the supplier's profit function:

$$w_{\Theta}(\gamma) := \overbrace{\operatorname{argmax}_{w \geq f} \left\{ \min_{\Sigma \in \Theta: w_{\Sigma}^{UB} \geq w} \left[(w - f)Q_{\Sigma}(w) + \gamma\Pi_{\Sigma}(w) \right] \right\}}^{\text{Supplier's Problem}}.$$

In this setting, the retailer has to solve the following problem:

$$(\text{Model } M') : \quad \max_{\gamma \in [0,1]} (1 - \gamma)\Pi_{\Sigma_0}(w_{\Theta}(\gamma)).$$

Let $\gamma_{\Theta}(w) := \gamma$ when $w = w_{\Theta}(\gamma)$. We solve instead the problem on w^5 :

$$\max_{w \in [f, w_{\Theta}(1)]} (1 - \gamma_{\Theta}(w))\Pi_{\Sigma_0}(w).$$

Our solution approach thus depends on the difficulty of the task to infer $\gamma_{\Theta}(w)$. To illustrate the idea, we consider $D = \theta + \varepsilon$, where θ captures the average demand, and ε is a random noise with zero mean and standard deviation σ_D . In the following Example 2, we analyze the case that the retailer and supplier have asymmetric information about the demand size θ . The cases of asymmetric information on other parameters can be analyzed similarly.

Example 2 Asymmetric information about demand size θ

As retailers are closer to market, we assume that the retailer knows the actual value of θ , which is θ_0 when the contract is signed, but the supplier only knows that θ lies in the range of $[\theta_L, \theta_H]$. The information on ε , P_s and ρ are common knowledge. Furthermore, $E(P_s) \geq f$ ensures that the average selling price is greater than the cost of production. Note that for any $\theta \in [\theta_L, \theta_H]$,

$$Q_\theta(w) = \theta + \frac{\alpha(w)}{\sqrt{E(P_s^2)/4 - \alpha(w)^2}} \sigma_D,$$

and

$$\Pi_\theta(w) = \alpha(w)\theta - \sigma_D \sqrt{E(P_s^2)/4 - \alpha(w)^2} + \frac{1}{2} (E(P_s)\theta + \rho\sigma_D\sigma_{P_s}),$$

whenever $w < w_\theta^{UB}$, and $\alpha(w) = E(P_s)/2 - w$. When the market condition θ is unknown, the supplier maximizes the worst-case expected profit as follows:

$$\max_{w \geq f} \min_{\theta \in [\theta_L, \theta_H]: w \leq w_\theta^{UB}} \left\{ [(1 - \gamma)w - f + \gamma E(P_s)]\theta + \frac{(w - f)\alpha(w) - \gamma(E(P_s^2)/4 - \alpha(w)^2)}{\sqrt{E(P_s^2)/4 - \alpha(w)^2}} \sigma_D + \frac{1}{2} \gamma \rho \sigma_D \sigma_{P_s} \right\}.$$

Since $E(P_s)$ and w are both greater than f , the coefficient of θ is positive and therefore the above problem reduces to

$$\max_{w \geq f} \left\{ [(1 - \gamma)w - f + \gamma E(P_s)]\theta_L + \frac{(w - f)\alpha(w) - \gamma(E(P_s^2)/4 - \alpha(w)^2)}{\sqrt{E(P_s^2)/4 - \alpha(w)^2}} \sigma_D + \frac{1}{2} \gamma \rho \sigma_D \sigma_{P_s} \right\}.$$

The retailer can determine the optimal profit sharing formula in this case by solving

$$\max_{w \geq f} (1 - \gamma_{\theta_L}(w)) \Pi_{\theta_0}(w),$$

where

$$\gamma_{\theta_L}(w) = 1 - \frac{\sigma(D) \frac{E(P_s^2)}{4} (w - f)}{Q_{\theta_L}(w) \sqrt{(\frac{E(P_s^2)}{4} - (E(P_s)/2 - w)^2)^3}}.$$

If θ_0 is indeed in $[\theta_L, \theta_U]$, then $Q_{\theta_L}(w) \leq Q_{\theta_0}(w)$, and thus $\gamma_{\theta_L}(w) \leq \gamma_{\theta_0}(w)$. We have

$$\max_{w \geq f} (1 - \gamma_{\theta_L}(w)) \Pi_{\theta_0}(w) \geq \max_{w \geq f} (1 - \gamma_{\theta_0}(w)) \Pi_{\theta_0}(w).$$

Hence the retailer can always exploit the information asymmetry to generate more profits in the supply chain, compared to the symmetric setting. On the other hand, if the supplier's information set is not accurate, for example $\theta_0 < \theta_L$, then it is better off for the retailer to share her information on θ_0 in order to generate more profits in the supply chain.

2. Consistency Test

Afriat (1967) initiates the research on the following question: “what restrictions on the data set are necessary and sufficient for it to be consistent with observations drawn from a utility-maximizing consumer?” A large body of literature, both theoretical and empirical, has been built around this theme. Brown and Matzkin (1996) extend this research to derive testable restrictions on outcomes that are consistent with a Walrasian equilibrium in an exchange economy. Carvajal et al. (2013) develop tests for outcomes (prices and quantities) in a class of Cournot game to be consistent with convex ordering costs, by solving an LP.

We can use the results obtained in this paper to derive testable conditions for the distributionally robust Stackelberg game model. For instance, given a set of T observations $\{(P_s[t], Q[t], w[t], \gamma[t])\}_{t=1}^T$ on prices, order quantities, wholesale prices and profit sharing terms in a decentralized supply chain, can we test whether the data are consistent with the following model in a distributionally robust setting:

production cost of the supplier is a linear function, and the average demand of the product is a convex decreasing function of price P_s ?

Using Equations (18) and (22), and assuming P_s is deterministic, the model output must satisfy the following relationship:

$$\sigma(D_t) = \frac{(1 - \gamma[t])Q[t]\sqrt{\left(\frac{P_s^2[t]}{4} - \left(\frac{P_s[t]}{2} - w[t]\right)^2\right)^3}}{\frac{P_s^2[t]}{4}\left(w[t] - f\right)}, \quad (25)$$

so

$$E(D_t) = Q[t] - \left\{ \frac{\frac{P_s[t]}{2} - w[t]}{\sqrt{\frac{P_s^2[t]}{4} - \left(\frac{P_s[t]}{2} - w[t]\right)^2}} \sigma(D_t) \right\} \quad (26)$$

must be a convex decreasing function in $P_s[t]$, for some production cost parameter f . If we assume further that $P_s[1] \leq P_s[2] \dots \leq P_s[T]$, then the above reduces the consistency test to checking whether

$$\frac{E(D_{t+1}) - E(D_t)}{P_s[t+1] - P_s[t]}$$

is non-decreasing in t .

3. Estimating Mean Demand

We can also use our results to perform approximate inference of the (implicit) demand model in the Stackelberg game, with a few simple observations of the output. As an illustration, suppose we observe the following characteristics and outputs in 10 different profit sharing supply chains, operating under a Stackelberg model as described above. The pricing information (P_s, f, w^*) , ordering decision Q^* etc. are publicly available. Suppose the profit sharing formula is also known and as shown in Table 1. Can we find the corresponding mean and variance of the demand in the supply chains?

Case	$E(D)$	$E(P_s)$	$\sigma(D)$	$\sigma(P_s)$	ρ	f	γ	w^*	Q^*
1		120		30		5	0.80	45.77	221.18
2		120		30		5	0.60	74.93	190.48
3		120		30		5	0.40	88.22	175.03
4		120		30		5	0.20	95.54	165.13
5		120		30		15	0.40	91.55	170.71
6		120		30		25	0.40	94.73	166.31
7		120		30		40	0.40	99.25	159.45
8		120		30		55	0.60	97.00	162.99
9		120		30		55	0.60	96.58	157.04
10		120		30		55	0.40	96.05	151.92

Table 1: What is the mean and standard deviation of demand?

In fact, the values obtained above are generated in the case when the ordering and wholesale price decisions are constructed by solving a Stackelberg game model, in which price and demand are bivariate normal, and the demand has a mean of 200 and standard deviation of 50, for $\rho = 0.5$ in all instances except cases 9 and 10, where $\rho = 0$ and -0.5 respectively.

Formula (25) and (26), extended to the case when P_s is random, provide a quick way to estimate the mean and standard deviation of the unknown demand based on a distributionally robust model. Interestingly, even without knowing ρ , our formula derived for the distributionally robust model returns the following estimates for the mean and standard deviation of the demand in the 10 instances:

The estimate for mean demand using our approach performs extremely well, compared to the naive heuristic of using order quantity to approximate the mean demand.

Case	1	2	3	4	5	6	7	8	9	10
$E(D)$	206.56	205.79	203.23	199.87	201.93	200.33	197.31	199.87	192.97	187.16
$\sigma(D)$	61.85	61.59	54.99	49.46	52.65	50.14	46.09	49.39	48.99	49.11

Table 2: Estimates for the mean and standard deviation

Appendix B: Proofs

Proof of Proposition 1: Given $(P_s, D) \in \Phi(\Sigma_0)$, using conditional expectation, we see from (8) that the objective function of the original inner problem (denoted by LHS hereafter) is equal to the objective function of the CPP problem (denoted by RHS hereafter). With $\mathbf{M}_i, i = 1, 2$, defined using variables (7), it is easy to see by their definitions that $\mathbf{M}_1 + \mathbf{M}_2 = \Sigma_0$. Also, $\mathbf{M}_i \geq 0, i = 1, 2$, since P_s and D are nonnegative random variables. Furthermore, since

$$\mathbf{M}_i = E_{P_s, D}((P_s \ D \ 1)^T (P_s \ D \ 1) | I(i)) Prob(I(i)) \text{ for } i = 1, 2,$$

they are positive semi-definite. Hence, $\mathbf{M}_i, i = 1, 2$, defined this way are feasible to the RHS. It is then clear that LHS \geq RHS.

We next show that LHS \leq RHS, by constructing from the optimal solution of the RHS a distribution in $\Phi(\Sigma_0)$ attaining at least the optimal value of the LHS.

Let \mathbf{M}_i^* denote the optimal solution obtained by solving the RHS, where

$$\mathbf{M}_i^* = \begin{pmatrix} PP^*(i) & PD^*(i) & P^*(i) \\ PD^*(i) & DD^*(i) & D^*(i) \\ P^*(i) & D^*(i) & y^*(i) \end{pmatrix}, \quad i = 1, 2.$$

Since \mathbf{M}_i^* is positive semidefinite with nonnegative entries, it is a doubly nonnegative matrix of order 3. By Theorem 2.4 of Berman and Shaked-Monderer (2003), \mathbf{M}_i^* is completely positive. Therefore, \mathbf{M}_i^* can be represented as a sum of rank 1 completely positive matrices (Proposition 2.1, Berman and Shaked-Monderer (2003)),

$$\mathbf{M}_i^* = \sum_{j=1}^{k_i} \begin{pmatrix} b_j^*(i)_1 \\ b_j^*(i)_2 \\ b_j^*(i)_3 \end{pmatrix} \begin{pmatrix} b_j^*(i)_1 \\ b_j^*(i)_2 \\ b_j^*(i)_3 \end{pmatrix}^T,$$

where $b_j^*(i)_1, b_j^*(i)_2, b_j^*(i)_3 \geq 0, j = 1, \dots, k_i, i = 1, 2$.

We use the above decomposition of \mathbf{M}_i^* , $i = 1, 2$, to construct a discrete distribution, which is feasible to the LHS, as follows:

$$P \left((P_s, D) = \left(\frac{b_j^*(i)_1}{b_j^*(i)_3}, \frac{b_j^*(i)_2}{b_j^*(i)_3} \right) \right) = b_j^*(i)_3^2, \quad j = 1, \dots, k_i, i = 1, 2,$$

with probability of (P_s, D) taking other values equal to zero. If $b_j^*(i)_3 = 0$ for some $j = 1, \dots, k_i$, $i = 1, 2$, we let the probability above equal to zero. With this distribution, (P_s, D) is in $\Phi(\Sigma_0)$. We also have

$$\begin{aligned} & E_{P_s, D} \left(\min [P_s D, P_s Q] \right) \\ &= \sum_{i=1}^2 \sum_{j=1}^{k_i} \min \left[\frac{b_j^*(i)_1 b_j^*(i)_2}{b_j^*(i)_3^2}, \frac{b_j^*(i)_1}{b_j^*(i)_3} Q \right] b_j^*(i)_3^2 \\ &\leq PD^*(1) + QP^*(2). \end{aligned}$$

Hence, LHS \leq RHS. Since we also have LHS \geq RHS, therefore, LHS = RHS.

Note that in above, we have also obtained a distribution in $\Phi(\Sigma_0)$ that solves the LHS. \square

Proof of Theorem 2: Note that solving Problem (13) is the same as solving the following minimization problem

$$\begin{aligned} & \min_{Q \geq 0, \mathbf{Y}} \quad wQ + \text{Tr}(\Sigma_0 \mathbf{Y}) - E(P_s D) & (27) \\ & \text{s.t.} \\ & \mathbf{Y} \succeq \frac{1}{2} \begin{pmatrix} 0 & 1 & -Q \\ 1 & 0 & 0 \\ -Q & 0 & 0 \end{pmatrix}, \quad \mathbf{Y} \succeq 0, \end{aligned}$$

We will work on the above minimization problem. By transformation of variables, $\mathbf{S}_1 = \Sigma_0^{1/2} \mathbf{Y} \Sigma_0^{1/2}$, solving Problem (27) is equivalent to solving the following problem (here, we ignore the constant

term $-E(P_s D)$ temporarily):

$$\begin{aligned} \min_{Q \geq 0, \mathbf{S}_1} \quad & wQ + \text{Tr}(\mathbf{S}_1) \\ \text{s.t.} \quad & \\ & \mathbf{S}_1 \succeq \frac{1}{2} \Sigma_0^{1/2} \begin{pmatrix} 0 & 1 & -Q \\ 1 & 0 & 0 \\ -Q & 0 & 0 \end{pmatrix} \Sigma_0^{1/2}, \\ & \mathbf{S}_1 \succeq 0 \end{aligned} \tag{28}$$

There exist orthogonal matrix $\mathbf{U}(Q)$ and diagonal matrix $\mathbf{D}(Q) = \begin{pmatrix} d_1(Q) & 0 & 0 \\ 0 & d_2(Q) & 0 \\ 0 & 0 & d_3(Q) \end{pmatrix}$

such that

$$\frac{1}{2} \Sigma_0^{1/2} \begin{pmatrix} 0 & 1 & -Q \\ 1 & 0 & 0 \\ -Q & 0 & 0 \end{pmatrix} \Sigma_0^{1/2} = \mathbf{U}(Q) \mathbf{D}(Q) \mathbf{U}(Q)^T.$$

$d_i(Q), i = 1, 2, 3$, are the eigenvalues of the left hand side matrix.

Let $\mathbf{S} = \mathbf{U}(Q)^T \mathbf{S}_1 \mathbf{U}(Q)$, we can further transform Problem (28) as

$$\begin{aligned} \min_{Q \geq 0, \mathbf{S}} \quad & wQ + \text{Tr}(\mathbf{S}) \\ \text{s.t.} \quad & \\ & \mathbf{S} \succeq \mathbf{D}(Q), \\ & \mathbf{S} \succeq 0. \end{aligned} \tag{29}$$

For any feasible solution (Q, \mathbf{S}) to Problem (29), suppose s_i is the main diagonal entry of \mathbf{S} for $i = 1, 2, 3$. Then we must have $s_i \geq \max\{d_i(Q), 0\}, 1 \leq i \leq 3$. From this observation, an optimal solution to Problem (29), (Q_r^*, \mathbf{S}_r^*) , has

$$\mathbf{S}_r^* = \begin{pmatrix} \max\{d_1(Q_r^*), 0\} & 0 & 0 \\ 0 & \max\{d_2(Q_r^*), 0\} & 0 \\ 0 & 0 & \max\{d_3(Q_r^*), 0\} \end{pmatrix}$$

with Q_r^* an optimal solution to the following minimization problem in one dimension:

$$\min_{Q \geq 0} wQ + \sum_{i=1}^3 \max\{d_i(Q), 0\}, \tag{30}$$

Observe that $d_i(Q), i = 1, 2, 3$, are also eigenvalues of the following matrix:

$$\begin{aligned}\Sigma_1 &= \frac{1}{2}\Sigma_0 \begin{pmatrix} 0 & 1 & -Q \\ 1 & 0 & 0 \\ -Q & 0 & 0 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} E(P_s D) - QE(P_s) & E(P_s^2) & -QE(P_s^2) \\ E(D^2) - QE(D) & E(P_s D) & -QE(P_s D) \\ E(D) - Q & E(P_s) & -QE(P_s) \end{pmatrix}.\end{aligned}$$

Computing the eigenvalues of Σ_1 by solving the equation $\det(\lambda I - \Sigma_1) = 0$ for λ , we have:

$$\begin{aligned}d_1(Q) &= \frac{1}{2} \left[E(P_s D) - QE(P_s) + \sqrt{E(P_s^2)(Q^2 - 2E(D)Q + E(D^2))} \right] \\ d_2(Q) &= \frac{1}{2} \left[E(P_s D) - QE(P_s) - \sqrt{E(P_s^2)(Q^2 - 2E(D)Q + E(D^2))} \right] \\ d_3(Q) &= 0.\end{aligned}$$

Since Σ_0 is positive semi-definite, it implies that

$$\det(\Sigma_0) = E(P_s^2)E(D^2) + 2E(P_s)E(D)E(P_s D) - E(P_s)^2E(D^2) - E(P_s^2)E(D)^2 - E(P_s D)^2 \geq 0, \quad (31)$$

from which we have

$$(E(P_s D) - QE(P_s))^2 \leq E(P_s^2)(Q^2 - 2E(D)Q + E(D^2)),$$

for all Q . The latter inequality holds since the discriminant of $E(P_s^2)(Q^2 - 2E(D)Q + E(D^2)) - (E(P_s D) - QE(P_s))^2$ in Q is nonpositive by (31), with its coefficient of Q^2 positive. It then follows that $d_1(Q) \geq 0$ and $d_2(Q) \leq 0$ for all Q .

Let $g(Q)$ denote the objective function of Problem (30). We have

$$\begin{aligned}g(Q) &= wQ + d_1(Q) \\ &= wQ + \frac{1}{2} \left[E(P_s D) - QE(P_s) + \sqrt{E(P_s^2)(Q^2 - 2E(D)Q + E(D^2))} \right].\end{aligned}$$

Taking differentiation, we have

$$g'(Q) = w - \frac{1}{2}E(P_s) + \frac{1}{2}\sqrt{E(P_s^2)}(Q^2 - 2E(D)Q + E(D^2))^{-1/2}(Q - E(D)), \quad (32)$$

and

$$g''(Q) = \frac{1}{2} \sqrt{E(P_s^2)} (Q^2 - 2E(D)Q + E(D^2))^{-3/2} (E(D^2) - E(D)^2) > 0.$$

Hence, $g(Q)$ is a strictly convex function of Q , and since $g(Q) \rightarrow \infty$ as $Q \rightarrow \infty$, there exists a unique optimal solution $Q_r^* \geq 0$ to Problem (30). Observe that $Q_r^* > 0$ if and only if $g'(0) < 0$, in which case, Q_r^* satisfies $g'(Q_r^*) = 0$. We have $g'(0) < 0$ if and only if

$$w < \frac{1}{2} \left[\sqrt{\frac{E(P_s^2)}{E(D^2)}} E(D) + E(P_s) \right].$$

Let α and β be as given in Definition 1. Setting $g'(Q_r^*) = 0$. From (32), we have

$$\left[w - \frac{1}{2} E(P_s) \right] ((Q_r^*)^2 - 2E(D)Q_r^* + E(D^2))^{1/2} = - \left[\frac{1}{2} \sqrt{E(P_s^2)} (Q_r^* - E(D)) \right], \quad (33)$$

from which we obtain

$$\alpha^2 ((Q_r^*)^2 - 2E(D)Q_r^* + E(D^2)) = \beta ((Q_r^*)^2 - 2E(D)Q_r^* + E(D^2)).$$

Solving for Q_r^* in the above equation, we have

$$\begin{aligned} Q_r^* &= E(D) \pm \sqrt{\frac{\alpha^2 (E(D^2) - E(D)^2)}{\beta - \alpha^2}} \\ &= E(D) \pm \frac{|\alpha|}{\sqrt{\beta - \alpha^2}} \sigma_D \end{aligned} \quad (34)$$

where $\sigma(D)$ is the standard deviation of demand D . Note that (34) implies that we must have $\beta > \alpha^2$. Substituting the above expression (34) for Q_r^* into the right-hand side of (33) and comparing with its left-hand side, we obtain

$$Q_r^* = E(D) + \frac{\alpha}{\sqrt{\beta - \alpha^2}} \sigma(D).$$

We next obtain the worst-case profit, Π_{wst} , which equals to the negative of the optimal objective value of Problem (28) or equivalently, Problem (30), $g(Q_r^*)$, plus the constant term $E(P_s D)$.

$$\Pi_{wst} = -wQ_r^* - \frac{1}{2} \left[E(P_s D) - Q_r^* E(P_s) + \sqrt{E(P_s^2) (Q_r^{*2} - 2E(D)Q_r^* + E(D^2))} \right] + E(P_s D)$$

Since

$$\begin{aligned}
(Q_r^*)^2 &= E^2(D) + \frac{\alpha^2}{\beta - \alpha^2} \text{Var}(D) + 2E(D) \frac{\alpha}{\sqrt{\beta - \alpha^2}} \sigma(D) \\
2E(D)Q_r^* &= 2E^2(D) + 2E(D) \frac{\alpha}{\sqrt{\beta - \alpha^2}} \sigma(D) \\
(Q_r^*)^2 - 2E(D)Q_r^* + E(D)^2 &= E^2(D) + \frac{\alpha^2}{\beta - \alpha^2} \text{Var}(D) - 2E^2(D) + E(D)^2 \\
&= \frac{\alpha^2}{\beta - \alpha^2} \text{Var}(D) + \text{Var}(D) \\
&= \frac{\beta}{\beta - \alpha^2} \text{Var}(D)
\end{aligned}$$

We have

$$\begin{aligned}
\Pi_{wst} &= \left(\frac{1}{2} E(P_s) - w \right) Q_r^* - \frac{1}{2} \sqrt{E(P_s^2) \left((Q_r^*)^2 - 2E(D)Q_r^* + E(D)^2 \right)} + \frac{1}{2} E(P_s D) \\
&= \alpha Q_r^* - \sqrt{\beta \cdot \frac{\beta}{\beta - \alpha^2}} \sigma(D) + \frac{1}{2} E(P_s D) \\
&= \alpha \left(E(D) + \frac{\alpha}{\sqrt{\beta - \alpha^2}} \sigma(D) \right) - \frac{\beta}{\sqrt{\beta - \alpha^2}} \sigma(D) + \frac{1}{2} E(P_s D) \\
&= \alpha E(D) - \left(-\frac{\alpha^2}{\sqrt{\beta - \alpha^2}} \sigma(D) + \frac{\beta}{\sqrt{\beta - \alpha^2}} \sigma(D) \right) + \frac{1}{2} E(P_s D) \\
&= \alpha E(D) - \sqrt{(\beta - \alpha^2)} \sigma(D) + \frac{1}{2} E(P_s D)
\end{aligned}$$

□

Before we present the proof of Theorem 3, we proceed with a few results that are needed in the proof of the theorem.

Remark 3 *From the proof of Theorem 2, we can also obtain \mathbf{Y}_r^* :*

$$\mathbf{Y}_r^* = \boldsymbol{\Sigma}_0^{-1/2} \mathbf{U}(Q_r^*) \begin{pmatrix} d_1(Q_r^*) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{U}(Q_r^*)^T \boldsymbol{\Sigma}_0^{-1/2},$$

where

$$\mathbf{U}(Q_r^*) \begin{pmatrix} d_1(Q_r^*) & 0 & 0 \\ 0 & d_2(Q_r^*) & 0 \\ 0 & 0 & d_3(Q_r^*) \end{pmatrix} \mathbf{U}(Q_r^*)^T = \frac{1}{2} \boldsymbol{\Sigma}_0^{1/2} \begin{pmatrix} 0 & 1 & -Q_r^* \\ 1 & 0 & 0 \\ -Q_r^* & 0 & 0 \end{pmatrix} \boldsymbol{\Sigma}_0^{1/2},$$

$$\begin{aligned}
d_1(Q_r^*) &= \frac{1}{2} \left[E(P_s D) - Q_r^* E(P_s) + \sqrt{E(P_s^2)((Q_r^*)^2 - 2E(D)Q_r^* + E(D^2))} \right] \\
d_2(Q_r^*) &= \frac{1}{2} \left[E(P_s D) - Q_r^* E(P_s) - \sqrt{E(P_s^2)((Q_r^*)^2 - 2E(D)Q_r^* + E(D^2))} \right] \\
d_3(Q_r^*) &= 0.
\end{aligned}$$

Here, $\mathbf{U}(Q_r^*) = (u_1, u_2, u_3)$ is an orthogonal matrix. Hence, $u_i, i = 1, 2, 3$, are orthogonal eigenvectors of $\frac{1}{2}\Sigma_0^{1/2} \begin{pmatrix} 0 & 1 & -Q_r^* \\ 1 & 0 & 0 \\ -Q_r^* & 0 & 0 \end{pmatrix} \Sigma_0^{1/2}$, with $d_i(Q_r^*)$, $i = 1, 2, 3$, being their corresponding eigenvalues.

Note that the dual problem to Problem (13) is given by:

$$\begin{aligned}
\min_{\mathbf{X}_1, \mathbf{X}_2} \quad & -(\mathbf{X}_1)_{12} + E(P_s D) \\
\text{s.t.} \quad & \\
& \mathbf{X}_1 + \mathbf{X}_2 = \Sigma_0 \\
& (\mathbf{X}_1)_{13} \leq w \\
& \mathbf{X}_1, \mathbf{X}_2 \succeq 0.
\end{aligned} \tag{35}$$

We have the following complementary slackness conditions on Problems (13) and (35):

$$\left(\frac{1}{2} \begin{pmatrix} 0 & 1 & -Q \\ 1 & 0 & 0 \\ -Q & 0 & 0 \end{pmatrix} - \mathbf{Y} \right) \mathbf{X}_1 = 0, \tag{36}$$

$$\mathbf{Y} \mathbf{X}_2 = 0, \tag{37}$$

$$Q(w - (\mathbf{X}_1)_{13}) = 0, \tag{38}$$

Note that primal, dual feasibility conditions and the above complementary slackness conditions are sufficient for optimality to Problem (13) and its dual, Problem (35), with zero duality gap.

Let $(\mathbf{X}_1^*, \mathbf{X}_2^*)$ be an optimal solution to Problem (35). We present an explicit expression for $(\mathbf{X}_1^*, \mathbf{X}_2^*)$ in the following proposition by showing that primal, dual feasibility conditions and the above complementary slackness conditions are satisfied by this expression for $(\mathbf{X}_1^*, \mathbf{X}_2^*)$.

Proposition 5 $(\mathbf{X}_1^*, \mathbf{X}_2^*)$ is given by

$$\begin{aligned}\mathbf{X}_1^* &= (\boldsymbol{\Sigma}_0^{1/2} u_1)(\boldsymbol{\Sigma}_0^{1/2} u_1)^T, \\ \mathbf{X}_2^* &= (\boldsymbol{\Sigma}_0^{1/2} u_2)(\boldsymbol{\Sigma}_0^{1/2} u_2)^T + (\boldsymbol{\Sigma}_0^{1/2} u_3)(\boldsymbol{\Sigma}_0^{1/2} u_3)^T,\end{aligned}$$

where if $Q_r^* > 0$,

$$\boldsymbol{\Sigma}_0^{1/2} u_1 = a_1 \begin{pmatrix} 1 \\ \frac{\Delta}{E(P_s^2)} \left[1 + \frac{wQ_r^*}{d_1(Q_r^*)} \right] \\ \frac{w\Delta}{E(P_s^2)d_1(Q_r^*)} \end{pmatrix}, \quad (39)$$

$$\boldsymbol{\Sigma}_0^{1/2} u_2 = a_2 \begin{pmatrix} 1 \\ \frac{\Delta}{E(P_s^2)} \left[-1 + \frac{Q_r^*}{d_2(Q_r^*)} (w - E(P_s)) \right] \\ \frac{\Delta}{E(P_s^2)d_2(Q_r^*)} (w - E(P_s)), \end{pmatrix}, \quad (40)$$

$$\boldsymbol{\Sigma}_0^{1/2} u_3 = a_3 \begin{pmatrix} 0 \\ 1 \\ \frac{1}{Q_r^*} \end{pmatrix}, \quad (41)$$

with $a_i > 0, i = 1, 2, 3$, given by

$$\begin{aligned}a_1 &= \sqrt{\frac{E(P_s^2)d_1(Q_r^*)}{\Delta}}, \\ a_2 &= \sqrt{\frac{-E(P_s^2)d_2(Q_r^*)}{\Delta}}, \\ a_3 &= Q_r^* \sqrt{1 - \frac{\Delta}{E(P_s^2)} \left[\frac{w^2}{d_1(Q_r^*)} - \frac{(w - E(P_s))^2}{d_2(Q_r^*)} \right]}.\end{aligned}$$

Here, $\Delta = \sqrt{E(P_s^2)((Q_r^*)^2 - 2E(D)Q_r^* + E(D^2))} > 0$.

Proof: With $\mathbf{Y} = \mathbf{Y}_r^*$, whose explicit expression is given in Remark 3, we can derive $(\mathbf{X}_1, \mathbf{X}_2)$ using (36), (37) and $\mathbf{X}_1 + \mathbf{X}_2 = \boldsymbol{\Sigma}_0$, as shown below

$$\begin{aligned}\mathbf{X}_1 &= (\boldsymbol{\Sigma}_0^{1/2} u_1)(\boldsymbol{\Sigma}_0^{1/2} u_1)^T, \\ \mathbf{X}_2 &= (\boldsymbol{\Sigma}_0^{1/2} u_2)(\boldsymbol{\Sigma}_0^{1/2} u_2)^T + (\boldsymbol{\Sigma}_0^{1/2} u_3)(\boldsymbol{\Sigma}_0^{1/2} u_3)^T.\end{aligned}$$

We use these as potential candidates for $\mathbf{X}_1^*, \mathbf{X}_2^*$. Except condition (38), we can check easily that $Q = Q_r^*, \mathbf{Y} = \mathbf{Y}_r^*$ and $\mathbf{X}_1, \mathbf{X}_2$ given above satisfy the complementary slackness conditions, (36),

(37), as well as the primal and dual feasibility conditions, except $w \geq (\mathbf{X}_1)_{13}$. To verify condition (38) and that $w \geq (\mathbf{X}_1)_{13}$, we need to derive a closed form expression for $(\mathbf{X}_1)_{13}$.

Suppose $Q_r^* > 0$. Let us calculate $\Sigma_0^{1/2} u_i$, $i = 1, 2, 3$. We know by Remark 3 that

$$\frac{1}{2} \Sigma_0^{1/2} \begin{pmatrix} 0 & 1 & -Q_r^* \\ 1 & 0 & 0 \\ -Q_r^* & 0 & 0 \end{pmatrix} \Sigma_0^{1/2} u_i = d_i(Q_r^*) u_i, \quad i = 1, 2, 3.$$

Then,

$$\frac{1}{2} \Sigma_0 \begin{pmatrix} 0 & 1 & -Q_r^* \\ 1 & 0 & 0 \\ -Q_r^* & 0 & 0 \end{pmatrix} \Sigma_0^{1/2} u_i = d_i(Q_r^*) \Sigma_0^{1/2} u_i, \quad i = 1, 2, 3.$$

Hence, $\Sigma_0^{1/2} u_i$ is an eigenvector of $\frac{1}{2} \Sigma_0 \begin{pmatrix} 0 & 1 & -Q_r^* \\ 1 & 0 & 0 \\ -Q_r^* & 0 & 0 \end{pmatrix}$, with eigenvalue $d_i(Q_r^*)$, $i = 1, 2, 3$.

Therefore, $\Sigma_0^{1/2} u_i$, $i = 1, 2, 3$, can be obtained by solving

$$\left(\frac{1}{2} \Sigma_0 \begin{pmatrix} 0 & 1 & -Q_r^* \\ 1 & 0 & 0 \\ -Q_r^* & 0 & 0 \end{pmatrix} - d_i(Q_r^*) I \right) x = 0, \quad i = 1, 2, 3,$$

for x . Upon solving the above and after algebraic manipulations, using $Q_r^* = E(D) + \frac{\alpha}{\sqrt{\beta - \alpha^2}} \sigma(D)$ and the identity $\Delta = \frac{1}{2} \frac{E(P_s^2) \sigma(D)}{\sqrt{\beta - \alpha^2}} = \frac{1}{2} E(P_s^2) \frac{Q_r^* - E(D)}{\alpha}$, we obtain expressions (39)-(41) for $\Sigma_0^{1/2} u_i$, $i = 1, 2, 3$.

Expressions for a_i , $i = 1, 2$, given in the proposition are obtained by substituting (39) and (40) respectively into the following equation:

$$\frac{1}{2} (\Sigma_0^{1/2} u_i)^T \begin{pmatrix} 0 & 1 & -Q_r^* \\ 1 & 0 & 0 \\ -Q_r^* & 0 & 0 \end{pmatrix} \Sigma_0^{1/2} u_i = d_i(Q_r^*),$$

where we have used the fact that $\|u_i\| = 1$ to obtain the equation, $i = 1, 2$.

Finally, the expression for a_3 in the proposition is obtained by observing that the (3, 3) entries of \mathbf{X}_1 , \mathbf{X}_2 sum up to 1 and then substituting what is known of $\Sigma_0^{1/2} u_i$, $i = 1, 2, 3$, ((39)-(41)) and a_i , $i = 1, 2$ into these (3, 3) entries.

Using the above derivations, we have:

$$(\mathbf{X}_1)_{13} = a_1^2 \frac{w\Delta}{E(P_s^2)d_1(Q_r^*)} = \frac{E(P_s^2)d_1(Q_r^*)}{\Delta} \frac{w\Delta}{E(P_s^2)d_1(Q_r^*)} = w.$$

Thus condition (38) and $w \geq (\mathbf{X}_1)_{13}$ are satisfied.

Since $Q = Q_r^*$, $\mathbf{Y} = \mathbf{Y}_r^*$ and $\mathbf{X}_1, \mathbf{X}_2$ satisfy the complementary slackness conditions as well as primal and dual feasibility conditions, these imply that $\mathbf{X}_1^* = \mathbf{X}_1 = (\boldsymbol{\Sigma}_0^{1/2}u_1)(\boldsymbol{\Sigma}_0^{1/2}u_1)^T$ and $\mathbf{X}_2^* = \mathbf{X}_2 = (\boldsymbol{\Sigma}_0^{1/2}u_2)(\boldsymbol{\Sigma}_0^{1/2}u_2)^T + (\boldsymbol{\Sigma}_0^{1/2}u_3)(\boldsymbol{\Sigma}_0^{1/2}u_3)^T$. \square

Observe that the original problem (PR), which we are interested in solving, is equivalent to the dual problem to Problem (12), i.e., Problem (35) with the additional constraints that $\mathbf{X}_1 \geq 0, \mathbf{X}_2 \geq 0$. Hence, if $\boldsymbol{\Sigma}_0^{1/2}u_i, i = 1, 2, 3$, given in the statement of Proposition 5 are nonnegative, then $(\mathbf{X}_1^*, \mathbf{X}_2^*)$ is also an optimal solution to the dual of Problem (12), with $Q^* = Q_r^*, \mathbf{Y}_1^* = 0, \mathbf{Y}_2^* = 0, \mathbf{Y}^* = \mathbf{Y}_r^*$ solving Problem (12).

With the above, we are now ready to prove Theorem 3.

Proof of Theorem 3: It is clear that $\boldsymbol{\Sigma}_0^{1/2}u_1 \geq 0$ in (39) and $\boldsymbol{\Sigma}_0^{1/2}u_3 \geq 0$ in (41).

For $\boldsymbol{\Sigma}_0^{1/2}u_2$ in (40), it is easy to see that if the second entry of the right-hand side vector is nonnegative, then the last entry of the right-hand side vector is also nonnegative, and thus $\boldsymbol{\Sigma}_0^{1/2}u_2 \geq 0$. We next derive the condition under which the former is true.

It can be shown upon algebraic manipulations that

$$-1 + \frac{Q_r^*}{d_2(Q_r^*)}(w - E(P_s)) \geq 0 \Leftrightarrow wQ_r^* + d_1(Q_r^*) \leq E(P_sD).$$

We would now like to know under what condition on w we have

$$wQ_r^* + d_1(Q_r^*) \leq E(P_sD), \tag{42}$$

that is,

$$wQ_r^* + \frac{1}{2}[E(P_sD) - Q_r^*E(P_s) + \sqrt{E(P_s^2)((Q_r^*)^2 - 2E(D)Q_r^* + E(D^2))}] \leq E(P_sD),$$

where Q_r^* is given in Theorem 2. Substituting the expression of Q_r^* into the above inequality, we have

$$\sqrt{(\beta - \alpha^2)\sigma(D)} \leq \frac{1}{2}E(P_sD) + \alpha E(D). \tag{43}$$

By taking the square on both sides of (43), using the definitions of α and β in Definition 1, and upon algebraic manipulations, we get

$$w \leq \frac{1}{2} \left[E(P_s) + \frac{E(P_s D)E(D) - \sigma(D)\sqrt{E(P_s^2)E(D^2) - E(P_s D)^2}}{E(D^2)} \right], \quad (44)$$

or

$$w \geq \frac{1}{2} \left[E(P_s) + \frac{E(P_s D)E(D) + \sigma(D)\sqrt{E(P_s^2)E(D^2) - E(P_s D)^2}}{E(D^2)} \right]. \quad (45)$$

It can be shown by substitution into (43) that only (44) satisfies (43), and hence (42). Equality holds for (44) if and only if equality holds for (42).

Note that if condition (44) is satisfied, then it can be shown that w satisfies (14). Hence, the above arguments, which use the explicit expression of Q_r^* given in Theorem 2, are valid, and we have the needed $\Sigma_0^{1/2}u_2 \geq 0$. Therefore, we can conclude that $(\mathbf{X}_1^*, \mathbf{X}_2^*)$ solves the dual problem to Problem (12), with $Q^* = Q_r^*$, $\mathbf{Y}_1^* = 0$, $\mathbf{Y}_2^* = 0$, $\mathbf{Y}^* = \mathbf{Y}_r^*$ solving Problem (12), and Problem (12) having the optimal value as Problem (13). Hence, $Q^* = E(D) + \frac{\alpha}{\sqrt{\beta - \alpha^2}}\sigma(D)$ solves the robust problem (PR), when w satisfies (44).

When

$$w = \frac{1}{2} \left[E(P_s) + \frac{E(P_s D)E(D) - \sigma(D)\sqrt{E(P_s^2)E(D^2) - E(P_s D)^2}}{E(D^2)} \right], \quad (46)$$

the following equality holds,

$$-wQ_r^* - d_1(Q_r^*) + E(P_s D) = 0. \quad (47)$$

In this case, we know that the left-hand side of (47) is the optimal value to Problem (13), which is also the optimal value to Problem (12). When w satisfies (46), the optimal value to Problem (12) is therefore zero. Hence, for

$$w \geq \frac{1}{2} \left[E(P_s) + \frac{E(P_s D)E(D) - \sigma(D)\sqrt{E(P_s^2)E(D^2) - E(P_s D)^2}}{E(D^2)} \right],$$

the optimal value to Problem (12) must be equal to 0 due to its decreasing nature as w increases, by Proposition 2, and also because it is always greater than or equal to zero for all $w \geq 0$, since it is the optimal worst-case profit for the retailer. Now, 0 is attained by the objective function

in Problem (12) when $Q = 0$, $\mathbf{Y}_1 = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $\mathbf{Y}_2 = 0$ and $\mathbf{Y} = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, which are feasible to Problem (12). We conclude then that when

$$w > \frac{1}{2} \left[E(P_s) + \frac{E(P_s D)E(D) - \sigma(D)\sqrt{E(P_s^2)E(D^2) - E(P_s D)^2}}{E(D^2)} \right],$$

$Q^* = 0$. □

Proof of Proposition 3: It can be calculated that for $f \leq w \leq w^{UB}$,

$$\frac{dQ}{dw}(w) = -\frac{\beta\sigma(D)}{(\beta - \alpha^2)^{3/2}} < 0. \quad (48)$$

□

Proof of Proposition 4: Note that

$$\frac{d\Pi}{dw}(w) = -\left(E(D) + \sigma(D)\frac{\alpha}{\sqrt{\beta - \alpha^2}} \right) = -Q(w), \quad (49)$$

$$\frac{d\Pi_1}{dw}(w) = Q(w) + (w - f)\frac{dQ}{dw}(w). \quad (50)$$

The first order derivative of Π_S with respect to w is then given by

$$\frac{\partial \Pi_S}{\partial w}(w, \gamma) = (1 - \gamma)Q(w) + (w - f)\frac{dQ}{dw}(w). \quad (51)$$

Taking the cross partial derivative, we have

$$\frac{\partial^2 \Pi_S}{\partial w \partial \gamma}(w, \gamma) = -Q(w) < 0.$$

The supplier's worst-case profit function is therefore submodular in (w, γ) . It follows from Topkis's theorem (Theorem 2.8.2, Topkis, 1998) that $w(\gamma)$ is decreasing in $\gamma \in [0, 1)$.

Since $f < w^{UB}$, we either have $w(\gamma) = w^{UB}$, in which case, $f < w(\gamma)$, or the optimal $w(\gamma)$ satisfies the first order condition $\frac{\partial \Pi_S}{\partial w}(w, \gamma) = 0$. In the latter case, the first term on the right-hand side of Equation (51) is positive, and from Proposition 3, we have $\frac{dQ(w)}{dw} < 0$, so we also have $w(\gamma) > f$. □

Proof of Theorem 4: This expression for $\gamma(Q)$ can be found by substituting Equations (48) to (50) into Equation (20), and note that $w(Q) = E(P_s)/2 - \alpha(Q)$:

$$\gamma(Q) = \frac{Q(w) - (w - f) \frac{\beta\sigma(D)}{(\beta - \alpha^2)^{3/2}}}{Q(w)} = 1 - \frac{(E(P_s)/2 - \alpha(Q) - f)\beta\sigma(D)}{(\beta - \alpha(Q)^2)^{3/2}Q}.$$

□

Proposition 6 *If $0 < w \leq w^{UB}$, in which case $Q^* > 0$, the worst-case distribution $(P_s^*, D^*) \in \Phi(\Sigma_0)$ for the robust problem (PR) is given by*

$$\begin{aligned} P\left((P_s^*, D^*) = \left(\frac{E(P_s^2)d_1(Q^*)}{w\Delta}, \frac{d_1(Q^*) + wQ^*}{w}\right)\right) &= \left(\frac{a_1w\Delta}{E(P_s^2)d_1(Q^*)}\right)^2, \\ P\left((P_s^*, D^*) = \left(\frac{E(P_s^2)d_2(Q^*)}{(w - E(P_s))\Delta}, \frac{-d_2(Q^*) + Q^*(w - E(P_s))}{w - E(P_s)}\right)\right) \\ &= \left(\frac{a_2\Delta}{E(P_s^2)d_2(Q^*)}(w - E(P_s))\right)^2, \\ P\left((P_s^*, D^*) = (0, Q^*)\right) &= \left(\frac{a_3}{Q^*}\right)^2, \end{aligned}$$

and probability of (P_s^*, D^*) taking other values is zero.

Here, $a_i, i = 1, 2, 3$, have explicit expressions given in Proposition 5 by replacing Q_r^* with Q^* and $\Delta = \sqrt{E(P_s^2)((Q^*)^2 - 2E(D)Q^* + E(D^2))} > 0$. Expressions for $d_i(Q^*), i = 1, 2$, are given in Remark 3, where Q_r^* is replaced by Q^* .

Proof: By the proof of Theorem 3, we see that if

$$w \leq w^{UB} = \frac{1}{2} \left[E(P_s) + \frac{E(P_s D)E(D) - \sigma(D)\sqrt{E(P_s^2)E(D^2) - E(P_s D)^2}}{E(D^2)} \right],$$

$(\mathbf{X}_1^*, \mathbf{X}_2^*)$ given in Proposition 5 solves the dual problem to Problem (12). $(\mathbf{X}_1^*, \mathbf{X}_2^*), \mathbf{X}_1^*, \mathbf{X}_2^* \geq 0$, together with $Q^*, \mathbf{Y}^*, \mathbf{Y}_1^*$ and \mathbf{Y}_2^* satisfy the complementary slackness conditions (36)-(38) and primal, dual feasibility conditions for Problem (12) and its dual. Note that $\mathbf{X}_1^*, \mathbf{X}_2^*, \mathbf{Y}^*, \mathbf{Y}_1^*$ and \mathbf{Y}_2^* then satisfy primal, dual, complementary slackness conditions for Problem (9) and its dual with $Q = Q^*$. Hence, $(\mathbf{X}_1^*, \mathbf{X}_2^*)$ solves Problem (9) with $Q = Q^*$, so we have $(\mathbf{M}_1^*, \mathbf{M}_2^*) = (\mathbf{X}_1^*, \mathbf{X}_2^*)$, where $(\mathbf{M}_1^*, \mathbf{M}_2^*)$ denote an optimal solution to Problem (9) in the proof of Proposition 1. From the proof of Proposition 1, we know how to obtain the worst-case distribution from the rank

1 completely positive decomposition of $\mathbf{M}_1^*, \mathbf{M}_2^*$. Since $\mathbf{M}_1^* = \mathbf{X}_1^*, \mathbf{M}_2^* = \mathbf{X}_2^*$, these rank 1 completely positive decomposition are given in Proposition 5.

The condition on w ensures that the last entry of $\Sigma_0^{1/2} u_i, i = 1, 2, 3$, in Proposition 5 are positive. The latter is to ensure that the worst-case distribution given in the statement of this proposition is well-defined. □