Accepted Manuscript

Risk Shaping in Production Planning Problem with Pricing under Random Yield

Saman Eskandarzadeh, Kourosh Eshghi, Mohsen Bahramgiri

PII: S0377-2217(16)30069-8
DOI: 10.1016/j.ejor.2016.02.032
Reference: EOR 13537

To appear in: *European Journal of Operational Research*

Received date: 19 September 2014
Revised date: 30 October 2015
Accepted date: 19 February 2016

Please cite this article as: Saman Eskandarzadeh, Kourosh Eshghi, Mohsen Bahramgiri, Risk Shaping in Production Planning Problem with Pricing under Random Yield, *European Journal of Operational Research* (2016), doi: 10.1016/j.ejor.2016.02.032

Highlights

- Presented a risk shaping model for production planning problem under random yield
- Used Conditional Value-at-Risk as the risk measure
- Developed an efficient solution method for the presented nonconvex model
- The presented model can incorporate risk thresholds of decision makers
- The presented model can be used for obtaining efficient frontiers
- Characterized the set of saddle points by KKT conditions

Risk Shaping in Production Planning Problem with Pricing under Random Yield

Saman Eskandarzadeh[a], Kourosh Eshghi[b], Mohsen Bahramgiri[c]

[a] *IE Department, Sharif University of Technology, Iran,*
eskandar@mehr.sharif.edu
[b]*IE Department, Sharif University of Technology, Iran,*
eshghi@sharif.edu
[c]*Graduate School of Management and Economics, Sharif University of Technology, Iran,*
bahramgiri@sharif.edu

Abstract

In this paper, we address a basic production planning problem with price dependent demand and stochastic yield of production. We use price and target quantity as decision variables to control the risk of low production yield. The value of risk control becomes more important especially for products with short life cycle where high losses are unbearable in the short run. In this cases, optimization of a solely scalar function of profit is not sufficient to control the risk. We apply Conditional Value at Risk (CVaR) measure to model the risk preferences of the producer. The producer is interested in shaping the risk by bounding from below the means of α-tail distributions of profit for different values of α. The resulting model is nonconvex. We propose an efficient solution algorithm and present a sufficient optimality condition.

Keywords: Risk Management, Coherent risk measures, Conditional value at risk, Random yield, Risk Shaping

1. Introduction

In this paper, we consider production planning problem under random yield for a risk averse decision maker. In our setting, a producer sets price before the realization of yield. This is a common practice for products with long production cycle such as vaccines and electronic chips. Both of them also have random yield. The yield of production of a variety of products is random in different industries ranging from agricultural to manufacturing. This could be as a result of such different factors as production process, the quality of raw materials, or natural phenomena. Random yield is also present in procurement from suppliers either due to supply disruptions or their products not meeting quality standards. Random yield imposes high risk on producers. Therefore, controlling the risk in advance is of high value for them.

Nowadays, one important feature of many products is their short life cycle. This could be due to fast technological changes, such as in the case of semiconductors, or perishability and large inventory cost, such as in the case of biological and agricultural products. Therefore, it is unreasonable to use criterion like the expectation of profit to compare different policies. The expectation specifically is equal to the average long term profits and gives decisions which are optimal on average. Whereas products with short life cycle do not have opportunity to be sold in the long-run. Furthermore, it is possible that a few successive run of large losses , for example due to low demand, makes it impossible for a firm to stay in the market. Therefore, considering other measures other than mean becomes necessary.

Risk is an important factor in decision making of many industries. Risk sources are diverse and include any phenomenon that can not be predicted with certainty. Phenomena such as natural disasters, change of regulations, strikes, terrorist attacks and a lot more. Risk can not be defined without associating it with an attribute which is important for the decision maker. For example, for a firm, profit is an important attribute. Moreover, risk associated with profit can not be described just by one number such as variance, mean, quantile and etc. That is the very reason that in financial markets a lot of measures are used to make decision as to whether or not one portfolio is better than another. Perhaps the best way to describe the risk is to use the probability distribution of its associated attribute. We can not claim that probability distribution exactly describes the risk but it is at least our subjective description of the risk. So if we accept the distribution as the description of the risk, the next step is to compare distributions associated with alternatives (actions, policies or strategies) to find the best alternative. For example, consider profit as our desired attribute. Under different policies the distribution function of profit changes accordingly. One important question arises in this situation as to how to compare distributions. It depends on what aspects of the risk is more important to us. We know that the distribution function has all information about the risk associated with profit, but we need to know how to measure one aspect using the distribution function. If we use several measures, it is probable that at the end we can not come up with a single best distribution. Therefore, we usually use a few measures. In the case of one measure, we obtain complete order of distributions. In this paper, we deviate from usual practice in inventory problems and use more than one risk measure to compare distribution functions in a relatively easy but fundamental problem setting. For more study on risk notion, the reader may refer to Arrow [2], Machina [25], Kahneman and Tversky [21].

The expected utility criterion is usually used for modeling the risk aversion of a decision maker but there are serious pitfalls associated with this measure in practical decision making. First, and most importantly, it is hard to elicit utility function of a rational decision maker. Second, the utility values are not sensible in practice. One can compare several alternatives based on expected utility, but can hardly tell exactly how much one is better than another. The utility theory as a general descriptive model of behavior has long shown its weaknesses (e.g. Refer to Kahneman and Tversky [21]). However, it is a sensible model for description of rational behavior or in other words as a normative model.

In this paper, we use a risk measure called Conditional Value at Risk (CVaR). This risk measure has been popularized by Rockafellar and Uryasev (see Rockafellar and Uryasev [30, 31]). Its popularity comes from its convexity property and its relation to Value-at-risk (VaR), a widely used risk measure in finance. Moreover, CVaR is in the class of coherent risk measures (Artzner et al. [3]). CVaR simply computes the conditional mean of the $100(1-\alpha)$ percent of worst-case values of a continuous random variable X. Furthermore it is consistent with first and second order stochastic dominance. Since the introduction of coherent risk measures, a lot of variant classes of risk measures have been introduced for one-period decision problems (see e.g., Föllmer and Schied [16], Rockafellar et al. [32], Ben-Tal and Teboulle [5]).

The remainder of this paper is organized as follows. Section 2 briefly reviews the relevant literature. In section 3 we study production planning problem under yield uncertainty. We use CVaR measure to control the risk of large loss at different risk aversion levels. Since, the constrained problem is nonconvex, we propose an efficient algorithm for finding local maxima based on envelope theorems. Finally, in section 4, we provide a sufficient optimality theorem based on duality theory. It can assure the global optimality of the obtained solution from the proposed algorithm given some conditions are satisfied.

2. Literature Review

Yield uncertainty in inventory problems has been widely studied in the literature. Yano and Lee [44] is an excellent literature review for past studies in this area. Tajbakhsh et al. [34] reviews the literature on supply uncertainty in inventory problems. They classify supply uncertainty into three types: uncertainty in quantity, uncertainty in supply timing, uncertainty in purchase price. They further classify supply uncertainty in quantity into random yield, random supplier availability, and random capacity. Wang et al. [40] introduces similar classification using random disruption term instead of random supplier availability. They also add another source of supply uncertainty namely financial default. Random disruption model can be considered as a special case of random yield model with Bernoulli or binomial distribution. Due to its importance and prevalence, it has received separate consideration in the literature (e.g., Tang et al. [36], Tomlin [37], Dada et al. [12]).

Cho and Tang [9] consider a game setting in which a risk neutral supplier with random yield sells flu vaccines to a risk neutral retailer. They investigate the value of advance selling compared to other selling strategies in different game settings. Advance selling (also referred to as advance booking, pre-booking) means setting price before realization of demand and supply. They also extend their work to multiple retailers. Both for one retailer and multiple retailers they consider linear stochastic demand function. Deo and Corbett [13] investigate influenza vaccine's market in the presence of competition. They employ random yield model and deterministic demand. Tang and Yin [35] employ linear deterministic demand model and discrete random yields. They compare advance selling policy (they call it nonresponsive pricing) to spot selling (they call it responsive pricing ,i.e., selling after realization of uncertainties). Their advance selling model differentiates from our model in two ways. First, they use linear demand function. Second, their model does not consider risk aversion. Li and Zheng [24] study similar model as ours in multi-period setting with stochastic demand function for a risk neutral decision maker. In their model all unmet demand is backlogged allowing them to find a concave representation for profit function. Bakal and Akcali [4] consider a reverse supply chain where End-of-Life products are recovered for their valuable parts and material. The yield of recovery is random. They numerically investigate the effect of random yield on expected profit in different settings.

Risk aversion has been studied extensively in inventory problems with uncertain demand. But comparatively, it has received much less attention in inventory problems with uncertain yield. Modeling risk by coherent risk measures is most recent in this area of research compared to modeling of risk by utility functions. Tomlin and Wang [38] compare different sourcing strategies for a risk averse decision maker in a multiproduct newsvendor problem when resources have random yield. They use CVaR and loss-averse [1] risk measures to quantify risk. They assume that the selling price of products are fixed.

[1]A loss-averse decision maker gives more weight to loss than gain as is introduced by Kahneman and

Their model is a two stage model which in the first stage the sourcing strategy is decided and in the second stage the allocation of the realized resources to each product. Kazaz and Webster [22] is one of the few papers which consider risk aversion in the presence of yield uncertainty. They use a two stage model for obtaining optimal production and pricing decisions for agricultural products with yield dependent cost. The pricing decision is made after realization of yield. In their model demand is deterministic. Their work is different than ours in two ways. First we consider joint optimization of price and quantity before realization of yield and second we use a coherent risk measure to capture risk aversion attitude. Xu and Lu [43] discuss the effect of yield randomness on optimal solution for a price-setting firm with random yield of production using stochastic dominance analysis. Eskandarzadeh et al. [15] investigate price setting and production quantity setting of influenza vaccines in advance for an influenza vaccine producer. Influenza vaccines have random yield. The paper addresses one dimension of risk aversion of a producer by optimization of CVaR of the profit. Our work differs from theirs since we consider multiple dimensions of risk aversion by incorporating risk constraints.

Another stream of research related to our work is the literature on risk-averse newsvendor problems. They model risk preferences of decision makers by VaR or chance constraints (e.g., Gan et al. [18], Özler et al. [28]), mean-CVaR objectives (e.g.,Lau [23], Xu and Li [42]), mean-variance constraints or objectives (e.g.,Gan et al. [18], Martínez-de Albéniz and Simchi-Levi [26], Choi et al. [11]), expected utility objectives (e.g., Lau [23], Eeckhoudt et al. [14], Agrawal and Seshadri [1], Gan et al. [18], Gaur and Seshadri [19],). To control the risk, they use operational instruments individually or in combination such as selling price, production or order quantity, and risk-mitigating contracts (e.g., Gaur and Seshadri [19], Martínez-de Albéniz and Simchi-Levi [26]). Gotoh and Takano [20] investigate newsvendor problem under CVaR and Mean-CVaR criteria. They extend their model to multiproduct case but in the finite scenario space. Chen et al. [8] study the newsvendor problem under CVaR measure. Their paper has similar settings as ours. But, they consider the uncertainty on demand side as opposed to our paper which considers the uncertainty on supply side. They also do not take into account the multifaceted aspect of risk. Moreover, we take a different solution approach than Chen's paper by introducing a change of variable. We take the demand as a factor of the target production quantity. This enables us to greatly simplify the analysis. Xu [41] adopts a similar setting as Chen et al.'s with additionally considering emergency purchase option after demand realization. Choi et al. [10] study multiproduct newsvendor problem. They find optimal orders under correlated and uncorrelated demand distributions. They use coherent risk measures to account for risk aversion. Their paper also has a very good literature review on risk aversion in inventory problems including papers using coherent risk measures to model risk preferences.

Our paper is the first work that uses a coherent risk measure to model multiple risk preferences in production planning problem with general price-dependent demand and yield uncertainty. In our model both selling price and production quantity are set before the realization of yield. The model turns out to be non-convex.

Tversky [21]

3. The Model

In this section, we consider a producer selling a single product. The life cycle of the product is one period and the unsold units of the product have zero salvage value at the end of the selling period. Moreover the yield of production, $\tilde{y}$, is random with Cumulative Distribution Function (CDF) $G(y)$ assuming to be continuous and Probability Distribution Function (PDF) $g(y)$ on support $[0,1]$. We denote the inverse of $G(y)$ with $G^{-1}(x)$. Sign $'$ denotes the first derivative and sign $''$ denotes the second derivative of a one-dimensional function. The decision variables are price, p, and target quantity, q. By target quantity, we mean the maximum quantity of production that will be realized. The demand function, $D(p)$, is a strictly decreasing function of price. There is production cost, c, per each unit of target quantity. We use CVaR as the risk measure. Depending on its parameter value, denoted by α, CVaR for an arbitrary random variable is defined as the mean of $100(1-\alpha)$ percent of worst-case values of that random variable. For instance, for a random variable representing profit, worst case values are smallest values of that random variable. Based on the definition, when $\alpha = 0$, CVaR is equal to the expected value. For our problem, the profit function can be defined as follows

$$f(p,q,\tilde{y}) = p\min\{\tilde{y}q, D(p)\} - cq \tag{1}$$

We denote α-CVaR of random variable $\tilde{x}$ with $\varphi_\alpha(\tilde{x})$. The objective is to control risk of function $f(p,q,\tilde{y})$ by using CVaR measure for different values of α. To simplify the analysis, we define $q := kD(p)$. We call the new variable k the production factor as it shows what multiple of demand will be produced. So, the optimization problem can be written as

$$\begin{aligned} [\mathcal{P}]\ \max & \phi_{\alpha_0}(p,k) \\ & \phi_{\alpha_i}(p,k) \geq v_i, i = 1,\ldots,m \\ & p,k \geq 0 \end{aligned}$$

Where $\phi_\alpha(p,k) := \varphi_\alpha(f(p,kD(p),\tilde{y}))$ and v_i is a desired threshold associated to α_i. Rockafellar and Uryasev [31] were the first to introduce CVaR as a measure to control the risk at different confidence levels. They provide a real example in financial market and show that adding CVaR constraints reduces the underperformance of chosen decision vector in terms of objective function value.

To proceed, we need to compute function $\phi_\alpha(p,k)$. For this purpose, we go over some properties of CVaR which was denoted by $\varphi_\alpha(.)$. It has been proved that $\varphi_\alpha(\tilde{x})$ for random variable $\tilde{x}$ representing cost, is a coherent measure of risk thereby satisfying the following four conditions (Rockafellar and Uryasev [31, Corollary 12]).

1. $\varphi_\alpha(a) = a$ for every scalar $a \in \mathbf{R}$
2. (Convexity)

$$\varphi_\alpha(\lambda\tilde{x}_1 + (1-\lambda)\tilde{x}_2) \leq \lambda\varphi_\alpha(\tilde{x}_1) + (1-\lambda)\varphi_\alpha(\tilde{x}_2)$$

   for arbitrary random variables $\tilde{x}_1$ and $\tilde{x}_2$ and $\lambda \in [0,1]$
3. (Monotonicity) $\varphi_\alpha(\tilde{x}_1) \leq \varphi_\alpha(\tilde{x}_2)$ for $\tilde{x}_1$ and $\tilde{x}_2$ where $\tilde{x}_1 \leq \tilde{x}_2$

4. (Positive homogeneity) $\varphi_\alpha(\lambda\tilde{x}) = \lambda\varphi_\alpha(\tilde{x})$ for $\lambda > 0$

From conditions 1, 2 and 4 the following condition can be deduced

5. $\varphi_\alpha(\tilde{x} + a) = \varphi_\alpha(\tilde{x}) + a$ for all $a \in \mathbf{R}$

Therefore, the α-CVaR of $f(p, q, \tilde{y})$ can be rewritten as

$$\begin{aligned}\phi_\alpha(p,k) =& \varphi_\alpha\left(p\min\{\tilde{y}kD(p), D(p)\} - ckD(p)\right)\\ =& pD(p)\,\varphi_\alpha\left(\min\{\tilde{y}k, 1\}\right) - ckD(p)\end{aligned} \tag{2}$$

The second equality is immediate from conditions 1, 4 and 5. There are different formulation to find α-CVaR of an arbitrary random variable. We use the one based on the following function (e.g. Föllmer and Schied [17])

$$F^I_{\tilde{u}}(z) = \sup\{\zeta\,|F_{\tilde{u}}(\zeta) \le z\}\ \forall z \in [0,1]$$

Where function $F_{\tilde{u}}(.)$ denotes the CDF of random variable $\tilde{u}$. When $\tilde{u}$ is continuous and increasing, function $F_{\tilde{u}}(.)$ is equal to inverse CDF of $\tilde{u}$. But you should note that $\tilde{x}$ may be discontinuous. The α-CVaR of $\tilde{u}$ can be obtained by the following formulation:

$$\varphi_\alpha(\tilde{u}) = \frac{1}{1-\alpha}\int_0^{1-\alpha} F^I_{\tilde{u}}(z)\,dz$$

To find α-CVaR of $\tilde{x} := \min\{\tilde{y}k, 1\}$ which is needed to compute 2, we first find its CDF

$$\begin{aligned}F_{\tilde{x}}(t) &= \text{Prob}\{\tilde{x} \le t\}\\ &= \text{Prob}\{\min\{\tilde{y}k, 1\} \le t\}\\ &= \begin{cases}\text{Prob}\{\tilde{y}k \le t\} & \text{if } t < 1\\ 1 & \text{Otherwise}\end{cases}\\ &= \begin{cases}G\left(\frac{t}{k}\right) & \text{if } t < 1\\ 1 & \text{Otherwise}\end{cases}\end{aligned}$$

The third equality follows from the support of random variable $\tilde{y}$ being interval $[0, 1]$.
Using the definition of $F^I_{\tilde{u}}(z)$ we obtain

$$F^I_{\tilde{x}}(z) = \begin{cases}kG^{-1}(z) & \text{if } z < G\left(\frac{1}{k}\right)\\ 1 & \text{otherwise}\end{cases}$$

We consider two cases to find $\varphi_\alpha(\tilde{x})$:
Case 1: $k < 1$
In this case, $\tilde{x}$ is equal to $\tilde{y}k$ thereby a continuous random variable. Therefore $F^I_{\tilde{x}}(z)$ coincides with the inverse function of $G\left(\frac{t}{k}\right)$ i.e. $kG^{-1}(z)$.

$$\varphi_\alpha(\tilde{x}) = \frac{1}{1-\alpha}\int_0^{1-\alpha} kG^{-1}(z)\,dz \tag{3}$$

By change of variable $t = G^{-1}(z)$,

$$\begin{aligned}\varphi_\alpha(\tilde{x}) &= \frac{1}{1-\alpha}\int_0^{G^{-1}(1-\alpha)} ktg(t)\,dt \\ &= \frac{1}{1-\alpha} kS\left(G^{-1}(1-\alpha)\right)\end{aligned} \tag{4}$$

Where $S(z) := \int_0^z ydG(y)$

Case 2: $k \geq 1$

Here, we consider two subcases. If $1-\alpha < G(1/k)$ then $F^I_{\tilde{x}}(z) = kG^{-1}(z)$ and we have

$$\varphi_\alpha(\tilde{x}) = \frac{1}{1-\alpha} kS\left(G^{-1}(1-\alpha)\right) \tag{5}$$

Otherwise i.e. $1-\alpha \geq G(1/k)$ and we have

$$\begin{aligned}\varphi_\alpha(\tilde{x}) =& \frac{1}{1-\alpha}\int_0^{G\left(\frac{1}{k}\right)} kG^{-1}(z)\,dz + \frac{1}{1-\alpha}\int_{G\left(\frac{1}{k}\right)}^{1-\alpha} dz \\ =& \frac{1}{1-\alpha}\int_0^{\frac{1}{k}} ktg(t)\,dt + \frac{1}{1-\alpha}\left(1-\alpha-G\left(\frac{1}{k}\right)\right) \\ =& \frac{1}{1-\alpha}\left[kS\left(\frac{1}{k}\right)+\left(1-\alpha-G\left(\frac{1}{k}\right)\right)\right]\end{aligned} \tag{6}$$

With replacing for $\varphi(\tilde{x})$ in (2) from equations (4) to (6) we obtain

$$\phi_\alpha(p,k) = \begin{cases} \frac{pD(p)}{1-\alpha} kS\left(G^{-1}(1-\alpha)\right) - ckD(p) \\ \text{if } \frac{1}{G^{-1}(1-\alpha)} > k \\ \frac{pD(p)}{1-\alpha}\left[kS\left(\frac{1}{k}\right)+\left(1-\alpha-G\left(\frac{1}{k}\right)\right)\right] - ckD(p) \\ \text{Otherwise} \end{cases} \tag{7}$$

We define $a_1(k,\alpha)$, $a_2(k,\alpha)$ and $a(k,\alpha)$ as follows.

$$a_1(k,\alpha) := \frac{1}{1-\alpha} kS\left(G^{-1}(1-\alpha)\right) \tag{8}$$

$$a_2(k,\alpha) := \frac{1}{1-\alpha}\left[kS\left(\frac{1}{k}\right)+\left(1-\alpha-G\left(\frac{1}{k}\right)\right)\right] \tag{9}$$

$$a(k,\alpha) := \begin{cases} a_1(k,\alpha) & \text{if } \frac{1}{G^{-1}(1-\alpha)} > k \\ a_2(k,\alpha) & \text{Otherwise} \end{cases} \tag{10}$$

Thereby rewriting 7 in a simpler form as

$$\phi_\alpha(p,k) = a(k,\alpha)pD(p) - ckD(p) \tag{11}$$

We denote the maximum of prices for which there is a positive demand by p_m. Let us define

$$p_l = \frac{(1-\alpha)\, c}{S\left(G^{-1}(1-\alpha)\right)}$$

Assumption 1. $p_m > p_l$

Assumption 1 is necessary for $\phi_\alpha(p,k)$ having positive value for a given α. It simply says that the maximum price should be larger than a threshold otherwise the optimal decision is not to produce. We will show this in the next Lemma.

Lemma 1. *If assumption 1 does not hold, the optimal action for the firm is not to produce anything i.e.* $q^* = k^* D(p^*) = 0$.

Proof. By taking the first derivative of $\phi_\alpha(p,k)$ with respect to k:

$$\begin{aligned}\frac{\partial \phi_\alpha(p,k)}{\partial k} &= \frac{\partial a(k,\alpha)}{\partial k} pD(p) - cD(p) \\ &= \begin{cases} \left(\frac{p}{1-\alpha} S\left(G^{-1}(1-\alpha)\right) - c\right) D(p) \\ \text{if} \quad \frac{1}{G^{-1}(1-\alpha)} > k \\ \left(\frac{p}{1-\alpha} S(1/k) - c\right) D(p) \\ \text{Otherwise} \end{cases}\end{aligned}$$

Since assumption 1 does not hold, feasible prices i.e. $p \in [0, p_m]$ are less than or equal to p_l. Additionally, function $S(1/k)$ is a decreasing function of k. Thus, the derivative is negative everywhere except at $p = p_m$ which is zero. So, the optimal solution should be on boundaries. For all points on lines $p = p_m$ and $k = 0$ the objective value is zero, so the optimal value is zero and optimal solutions are all (k,p) with $k = 0$ or $p = p_m$. □

Assumption 2. *Demand function is strictly decreasing in* $[0, p_m]$ *i.e.* $D'(p) < 0$

Lemma 2. *Let* p_0 *denote the minimum local maximizer of function* $pD(p)$ *, then if assumption 2 holds,* $\phi_\alpha(p,k)$ *is increasing for all* $p < p_0$.

Proof. Function $pD(p)$ is zero at $p = 0$ and positive near $p = 0$. So following the definition of p_0, it is obvious $(pD(p))'$ is positive on $(0, p_0)$. Indeed, we have

$$\begin{aligned}\frac{\partial \phi_\alpha(p,k)}{\partial p} &= a(k,\alpha)\left(D(p) + pD'(p)\right) - cD'(p) \\ &> -cD'(p) > 0\end{aligned}$$

The first inequality follows from positivity of $(pD(p))' = D(p) + pD'(p)$ for $p < p_0$ and the second inequality follows from the demand function being decreasing. □

It immediately follows from Lemma 2 that the optimal price can not be less than p_0. The conclusion is valid for all differentiable and strictly decreasing demand functions.

In model $\mathcal{P}$, without loss of generality, we assume

$$0 \le \alpha_0 < \alpha_1 < \cdots < \alpha_i < \cdots < \alpha_m$$

With this assumption, the v_is can be sorted as follows.

$$v_1 > \cdots > v_i > \cdots > v_m \tag{12}$$

Table 1: Commonly used demand functions

name	function	parameter values
Linear	$a - bp$	$a > 0, b > 0$
Exponential	ae^{-bp}	$a > 0, b > 0$
Constant elasticity	ap^{-b}	$a > 0, b > 0$
Logit	$\frac{e^{-bp}}{1+e^{-bp}}$	$b > 0$

The associated constraints to v_i's violating (12) are redundant and can be removed. To show this, assume for $i, j \in \{1, ..., m\}$ with $i < j$, we have the reverse inequality i.e. $v_i \leq v_j$. Following Lemma Appendix A.1, if for a given point (p, k), constraint $\phi_{\alpha_j}(p, k) \geq v_j$ holds, we can show constraints $\phi_{\alpha_i}(p, k) \geq v_i$ holds too as follows

$$\phi_{\alpha_i}(p, k) \geq \phi_{\alpha_j}(p, k) \geq v_j \geq v_i$$

Therefore, constraint ith is redundant. Problem $\mathcal{P}$ is not convex or quasiconcave [2]. Thus, it is not an easy task to find its global optimum. In the following, we first provide some properties of problem $\mathcal{P}$ and then propose an efficient search algorithm for finding its local maximum solutions. In cases there exist a unique local maximum, our proposed algorithm finds it in polynomial time.

Assumption 3. $2D'(p)^2 - D(p)D''(p) > 0 \; \forall p \in [0, p_m]$

All demand functions in table 1 satisfy assumption 3 for all parameter values except constant elasticity demand function which satisfies the assumption for $b > 1$.

Lemma 3. *Suppose assumptions 2 and 3 hold, then function $\phi_\alpha(k, p)$ is strictly quasiconcave on p*

Proof. There exists different approaches to prove this proposition, here we use lemma Appendix A.2 (Refer to the appendix). First, we take the first derivative of $\phi_\alpha(k, p)$ with respect to p,

$$\frac{\partial \phi_\alpha(p, k)}{\partial p} = a(k, \alpha)(D(p) + pD'(p)) - ckD'(p) = 0 \Longrightarrow p = \frac{ckD'(p) - a(k, \alpha)D(p)}{a(k, \alpha)D'(p)} \quad (13)$$

Then, we take second derivative,

$$\frac{\partial^2 \phi_\alpha(p, k)}{\partial p^2} = a(k, \alpha)(2D'(p) + pD''(p)) - ckD''(p) = 2a(k, \alpha)D'(p) + (a(k, \alpha)p - ck)D''(p) \quad (14)$$

[2] A quasiconcave program is a problem with convex constraint set and quasiconcave objective function

If we substitute p by it equal value 13 in the parenthesis, we obtain

$$
\begin{aligned}
&= a(k,\alpha)\left[2D'(p) + \left(\frac{ckD'(p) - a(k,\alpha)D(p)}{a(k,\alpha)D'(p)}\right) D''(p)\right] - ckD''(p) \\
&= 2a(k,\alpha)D'(p) + \left[a(k,\alpha)\left(\frac{ckD'(p) - a(k,\alpha)D(p)}{a(k,\alpha)D'(p)}\right) - ck\right] D''(p) \qquad (15)\\
&= 2a(k,\alpha)D'(p) + \frac{-a(k,\alpha)D(p)}{D'(p)}D''(p) = \frac{2a(k,\alpha)D'(p)^2 - a(k,\alpha)D(p)D''(p)}{D'(p)} \\
&= \frac{a(k,\alpha)\left(2D'(p)^2 - D(p)D''(p)\right)}{D'(p)} < 0 \qquad (16)
\end{aligned}
$$

The inequality follows from assumptions 2 and 3. □

Let us define

$$k_i := \frac{1}{G^{-1}(1-\alpha_i)} \text{for} i = 0, \ldots, m$$

Proposition 1. *Suppose assumption 1 hold for $\alpha = \alpha_0$ in addition to assumption 3 . If the optimal decision vector of problem $\mathcal{P}$, denoted by (p^*, k^*) is a regular point of the constraint set as well as the feasible set is not empty, then we have*

$$k^* \geq k_0 \textit{ and } p^* > p_0$$

Proof. Refer to Appendix B. □

As a hint in applying Proposition 1, note that regularity is a technical requirement in nonlinear programming and for practical applications we do not need to worry ourselves about its satisfaction. Define sets

$$I\left(k\right) := \left\{i \in \left\{1, \cdots, m\right\} | k < k_i\right\} \text{ and } J\left(k\right) := I(k)^C$$

Where A^C denotes the complement of set A. We can write the extensive form of model $\mathcal{P}$ as

$$
\begin{aligned}
[\mathcal{Q}\left(k\right)] \; & \max a_2\left(k, \alpha_0\right) pD\left(p\right) - ckD\left(p\right) && (17)\\
& \text{s.t.} && (18)\\
& a_1\left(k, \alpha_i\right) pD\left(p\right) - ckD\left(p\right) \geq v_i, i \in I\left(k\right) && (19)\\
& a_2\left(k, \alpha_i\right) pD\left(p\right) - ckD\left(p\right) \geq v_i, i \in J\left(k\right) && (20)\\
& p \geq 0, k \geq 0 && (21)
\end{aligned}
$$

If assumption 3 holds, model $\mathcal{Q}\left(k\right)$ is a quasiconcave program for all k greater than or equal to zero following lemma 3. Therefore it has a unique local maximum (e.g. refer to Simon and Blume [33]). We denote the optimal value of problem $\mathcal{Q}\left(k\right)$ for a given k by function $h\left(k\right)$. Function $h\left(k\right)$ might be indifferentiable at some points but it is continuous everywhere. To be able to present an search algorithm for finding the optimal solution of $h\left(k\right)$, we need to find the set of subgradients of $h\left(k\right)$. Since function $h\left(k\right)$ is one-dimensional, the set of all subgradients at point k is equal to the interval $\left[h'\left(k-\right), h'\left(k+\right)\right]$ where $h'\left(k+\right)$ and $h'\left(k-\right)$ denote the right and left derivative of $h\left(k\right)$ at k respectively.

Theorem 1. *Suppose assumptions 2 and 3 hold and the feasible set of model $\mathcal{Q}(k)$ given k has more than one solution. Then*

1) There is a unique local optimal price at k which we denote it by $p^*(k)$

2) Function $h(k)$ exists and is continuous and bounded and its right and left derivatives are given by

$$h'(k+) = \left(\frac{\partial \phi_{\alpha_0}(p^*(k), k)}{\partial k} - \frac{\frac{\partial \phi_{\alpha_0}(p^*(k),k)}{\partial p}}{\frac{\partial \phi_{\alpha_{i_+}}(p^*(k),k)}{\partial p}} \frac{\partial \phi_{\alpha_{i_+}}(p^*(k), k)}{\partial k} \right)$$

$$h'(k-) = \left(\frac{\partial \phi_{\alpha_0}(p^*(k), k)}{\partial k} - \frac{\frac{\partial \phi_{\alpha_0}(p^*(k),k)}{\partial p}}{\frac{\partial \phi_{\alpha_{i_-}}(p^*(k),k)}{\partial p}} \frac{\partial \phi_{\alpha_{i_-}}(p^*(k), k)}{\partial k} \right)$$

where

$$i_+ = \underset{j}{Argmin} \left\{ \frac{-\frac{\partial \phi_{\alpha_j}(p^*(k),k)}{\partial k}}{\frac{\partial \phi_{\alpha_j}(p^*(k),k)}{\partial p}} \frac{\partial \phi_{\alpha_0}(p^*(k), k)}{\partial p} \middle| \phi_{\alpha_j}(p^*(k), k) - v_j = 0 \right\}$$

$$i_- = \underset{j}{Argmax} \left\{ \frac{-\frac{\partial \phi_{\alpha_j}(p^*(k),k)}{\partial k}}{\frac{\partial \phi_{\alpha_j}(p^*(k),k)}{\partial p}} \frac{\partial \phi_{\alpha_0}(p^*(k), k)}{\partial p} \middle| \phi_{\alpha_j}(p^*(k), k) - v_j = 0 \right\}$$

Proof. proof of part 1) Due to Lemma 3, model $\mathcal{Q}(k)$ is a quasiconcave program and thereby it has a unique local maximizer (e.g. refer to Simon and Blume [33]). So $P^*(k)$ is singleton regarding the constraint set of model $\mathcal{Q}(k)$ is not empty.

proof of part 2) We first write the Lagrange function for model $\mathcal{Q}(k)$ as follows:

$$l(p, k, \lambda) = \phi_{\alpha_0}(p, k) + \sum_i \lambda_i [\phi_{\alpha_i}(p, k) - v_i]$$

We use envelope theorem Appendix A.1 to find the right and left derivatives of $h(k)$. First, we need to show that the set of saddle points is not empty. We use theorem Appendix A.3. Lagrange function $l(p, k, \lambda)$ is quasiconcave on p for all $k, \lambda \geq 0$ thereby having a unique local maximum. The proof of quasiconcavity is similar to that of Lemma 3 and we skip it. The optimal solution of $\mathcal{Q}(k)$ is regular or equivalently speaking it satisfies C.Q. conditions. From regularity, it follows that the optimal solution should satisfies KKT conditions. We need to show at least one type of constraints' qualification conditions hold at global optimum point. If for example the Jacobian matrix of binding constraints of $\mathcal{Q}(k)$ has full row rank at optimum then we are done (e.g. Simon and Blume [33, Theorem 19.2(a)]). Since, model $\mathcal{Q}(k)$ is one dimensional, we can only have at most one binding constraint. Nonnegativity constraint $p \geq 0$ can not be binding in the optimal solution due to proposition 1. So one of constraints $\phi_{\alpha_i}(p, k) \geq v_i$ should be binding. Here Jacobian is equal to the first derivative with respect to p for fixed k and it should be nonzero at optimal point to guaranty it satisfies KKT conditions. Assume the counter case that the binding constraint has zero first derivative. Since $\phi_{\alpha_i}(p, k)$ is a quasiconcave function for each fixed k (Due to Lemma 3) then it has a unique local maximum and the only way that the associated constraint can be binding is that v_i is equal to the maximum. But in this case the constraint $\phi_{\alpha_i}(p,k) \geq v_i$ has only one feasible solution and thereby model $\mathcal{Q}(k)$ has at most one feasible solution which is against the assumption of the theorem. Thus, we proved assumptions of theorem Appendix A.3 hold and more importantly the set of saddle points is equivalent to set of all points satisfying KKT conditions for model $\mathcal{Q}(k)$. So, the set of saddle points is not empty. Rockafellar [29] in Lemma 36.2 finds another characterization for the set of saddle points. Following the Rockafellar's Lemma, the set of saddle points given not empty is equivalent to the set of all points (p,λ) in the product set $P^*(k) \times \Lambda^*(k)$. Where $P^*(k)$ and $\Lambda^*(k)$ are defined by

$$P^*(k) = \underset{p\geq 0}{\mathrm{Argmax}}\inf_{\lambda\geq 0} l(p,k,\lambda)$$
$$\Lambda^*(k) = \underset{p\geq 0}{\mathrm{Argmin}}\sup_{\lambda\geq 0} l(p,k,\lambda)$$

Thus, by using Theorem Appendix A.1 we obtain

$$\begin{aligned} h'(k+) &= \max_{P^*(k)}\min_{\Lambda^*(k)} \frac{\partial l(p,k,\lambda)}{\partial k} \\ &= \max_{P^*(k)}\min_{\Lambda^*(k)} \frac{\partial \phi_{\alpha_0}(p,k)}{\partial k} + \sum_i \lambda_i \frac{\partial \phi_{\alpha_i}(p,k)}{\partial k} \\ &= \min_{\Lambda^*(k)} \frac{\partial \phi_{\alpha_0}(p^*(k),k)}{\partial k} + \sum_{i\in N(k)} \lambda_i \frac{\partial \phi_{\alpha_i}(p^*(k),k)}{\partial k} \end{aligned}$$

Set $\Lambda^*(k)$ can be replaced by KKT conditions. Let $N^*(k)$ be the set of indices which their associated constraints are active at $(p^*(k),k)$. Consider the following problem

$$h'(k+) = \min_{\lambda\geq 0} \frac{\partial \phi_{\alpha_0}(p^*(k),k)}{\partial k} + \sum_{i\in N(k)} \lambda_i \frac{\partial \phi_{\alpha_i}(p^*(k),k)}{\partial k}$$

subject to

$$\frac{\partial \phi_{\alpha_0}(p^*(k),k)}{\partial p} + \sum_{i\in N(k)} \lambda_i \frac{\partial \phi_{\alpha_i}(p^*(k),k)}{\partial p} = 0 \tag{22}$$

The constraint 22 along with nonnegetivity constraint $\lambda \geq 0$ are KKT conditions. Note that the complementary slackness conditions are considered by the definition of set $N^*(k)$. The above program is linear. So, the optimal solution occurs at an extreme point. The above problem is a relaxed knapsack problem and its optimal solution is

$$\lambda_{i_+} = \frac{-\frac{\partial \phi_{\alpha_0}(p^*(k),k)}{\partial p}}{\frac{\partial \phi_{\alpha_i}(p^*(k),k)}{\partial p}} \text{ Where } i_+ = \underset{j\in N(k)}{\mathrm{Argmin}} \frac{-\frac{\partial \phi_{\alpha_j}(p^*(k),k)}{\partial k}}{\frac{\partial \phi_{\alpha_j}(p^*(k),k)}{\partial p}} \frac{\partial \phi_{\alpha_0}(p^*(k),k)}{\partial p}$$

The computation of $h'(k-)$ is much as the same and we skip it.

$$h'(k-) = \min_{P^*(k)}\max_{\Lambda^*(k)} \frac{\partial l(p,k,\lambda)}{\partial k}$$

□

Theorem 1 provides us with a tool to develop a search algorithm which converges to a local maximum given it starts from a feasible solution. Later, we prove a sufficient condition for global optimality of a given local maximum.

We define V to be equal to an upper bound for value of $\max h(k) - h(k)$ in the domain of function $h(k)$ and Δ_0 to be equal to an upper bound for length of domain of function $h(k)$. Furthermore, let k_0 be a defined solution, i.e., function $h(k)$ assigns a value to it. Being defined at a given point for function $h(k)$, is translated to the constraint set of problem $\mathcal{Q}(k_0)$ being nonempty. In the following, we present a bisection algorithm for finding a locally ε-optimal solution[3] of function $h(k)$. The algorithm converges to global optimal solution if $h(k)$ is concave.

Algorithm 1 Bisection Algorithm

Step 0: Set $k = k_0$, $\Delta = 2\Delta_0$ and $N = 0$
Step 1: If $N > \lceil \log_2\left(\frac{V}{\epsilon}\right) \rceil$ stop. local maximum is found with precision ε; otherwise go to step 2;
Step 2: Find $h(k)$

$$
\begin{aligned}
h(k) = \max a_2(k,\alpha_0)\, pD(p) - ckD(p) \\
\text{s.t.} \qquad\qquad \\
a_1(k,\alpha_i)\, pD(p) - ckD(p) \geq v_i, i \in I(k) \\
a_2(k,\alpha_i)\, pD(p) - ckD(p) \geq v_i, i \in J(k) \\
p \geq 0
\end{aligned}
$$

Step 3: If $h(k)$ not defined (i.e. the associated problem is not feasible) go to step 4; otherwise go to step 5
Step 4: $\Delta \leftarrow -\frac{\Delta}{2}$; $k \leftarrow k + \Delta$; $N = N + 1$;go to step 2
Step 5: If $h'(k+) > 0$ then
  $\Delta \leftarrow \left|\frac{\Delta}{2}\right|$;$k \leftarrow k + \Delta$
  $N = N + 1$; go to step 1
 elseif $h'(k-) < 0$
  $\Delta \leftarrow -\left|\frac{\Delta}{2}\right|$;$k \leftarrow k + \Delta$
  $N = N + 1$;go to step 1
 else stop; local maximum is found

Assume the time complexity of computing $h(k)$ is equal to B. The time complexity of Algorithm 1 for reaching a locally ϵ-optimal solution within $\Delta_0/(2^N)$ neighborhood of a local optimal solution, is equal to $O(B \log_2(V/\epsilon))$[4]. If function $h(k)$ is strictly quasiconcave, it converges to a point in $\Delta_0/(2^N)$ neighborhood of optimal point. Finally, if function $h(k)$ is quasiconcave it converges to a point in $\Delta_0/(2^N)$ neighborhood of a locally optimal point (i.e. remember that a locally optimal point is a point, say x, with the condition that $0 \in \partial h(x)$). To evaluate function $h(k)$ at a point, we need to solve

[3] locally ε-optimal means a solution which differs from a local maximum at most ε

[4] Nemirovski, Arkadi, Lecture note on Efficient Methods in Convex Programming, Georgia Institute of Technology, www2.isye.gatech.edu/~nemirovs

a quasiconcave program. In our special case whose constraint set is an interval we can easily develop a specialized algorithm. In general case, to find the optimal solution, a series of convex optimization problems are solved. For more elaboration on quasiconcave optimization refer to Boyd and Vandenberghe [7].

4. Sufficient Optimality Theorem

Algorithm 1 only gives us a locally ϵ-optimal solution. But, we are interested in an ϵ-optimal solution. As we alluded to before, it is not an easy task to find assumptions under which the uniqueness of local optimal solutions is assured. We adopt another approach and find a sufficient optimality condition for checking global optimality of a found local optimal solution from algorithm 1. The path taken toward our goal, is the same path as when we wanted to prove some uniqueness result. But, as you will see, even for this closer goal, the analysis is somewhat tedious. To proceed, we need some preparations. Let us define

$$\begin{aligned} r\left(k\right) :=& \sum_{I(k)} \lambda_i \frac{S\left(G^{-1}\left(1-\alpha_i\right)\right)}{1-\alpha_i} \\ a\left(k\right) :=& \frac{1}{1-\alpha_0} + \sum_{J(k)} \frac{\lambda_i}{1-\alpha_i} \\ b\left(k\right) :=& 1 + \sum_{J(k)} \lambda_i \end{aligned}$$

We denote s_1, the infimum point such that for all $k \geq s_1$ function $g\left(\frac{1}{k}\right) + \frac{1}{k} g'\left(\frac{1}{k}\right)$ is positive. In other words,

$$s_1 = \inf\left\{ r \middle| \, xg\left(x\right) \text{ is increasing, } \forall x \leq 1/r \right\}$$

s_1 can be defined in an another equivalent way as

$$s_1 = \frac{1}{\sup\left\{ r \middle| \, xg(x) \, \text{is increasing, } \forall x \leq r \right\}}$$

Let function $u_0(k)$ be defined as

$$u_0\left(k\right) := 1 - \frac{1}{1-\alpha_0}\left(G\left(\frac{1}{k}\right) + \frac{1}{k} g\left(\frac{1}{k}\right)\right)$$

Define s_3^0 to be equal to the supremum point in such a way that function $u_0\left(k\right)$ is negative for all k less than or equal to s_3^0. Mathematically speaking, it is defined as

$$s_3^0 = \sup\left\{ k \middle| \, u_0\left(r\right) < 0 \, \forall r \leq k \right\}$$

Assumption 4. $s_1 < s_3^0$

The above assumption holds for common distributions.

Lemma 4. *If assumption 4 holds then function $u_0\left(k\right)$ is negative up to s_3^0 and increasing afterwards.*

Proof. Regarding the definition of s_3^0, it follows function $u_0(k)$ is negative up to s_3^0. Differentiating $u_0(k)$, we obtain

$$\frac{du_0(k)}{dk} = \frac{1}{k^2(1-\alpha_0)}\left(2g\left(\frac{1}{k}\right)+\frac{1}{k}g'\left(\frac{1}{k}\right)\right)$$

Due to assumption 4, $g\left(\frac{1}{k}\right)+\frac{1}{k}g'\left(\frac{1}{k}\right)$ is positive for all $k \geq s_3^0$. Therefore, the first derivative of function $u_0(k)$ is positive for all $k \geq s_3^0$. □

We define functions $u_i(k)$ as

$$u_i(k) = \left(\sum_{j=0}^{i}\lambda_j\right) - \left(\sum_{j=0}^{i}\frac{\lambda_j}{1-\alpha_j}\right)\left(G\left(\frac{1}{k}\right)+\frac{1}{k}g\left(\frac{1}{k}\right)\right) \quad \forall i \in \{1,\cdots,m\}$$

Where parameters λ_i are nonnegative real numbers. Moreover, λ_0 is defined to be equal to one. We define s_3^i, in the same way as we defined s_3^0 .

$$s_3^i = \sup\{k|\, u_i(r) < 0\,\forall r \leq k\}$$

Lemma 5. *If assumption 4 holds, function $u_i(k)$ is negative up to point s_3^i and increasing afterwards for all i. Moreover $s_3^0 \leq s_3^1 \leq \cdots \leq s_3^m$*

Proof. For proof we consider functions $w_i(k)$ as defined below instead of functions $u_i(k)$. They have the same properties as functions $u_i(k)$ in the sense of assertions of the theorem.

$$w_i(k) = \frac{u_i(k)}{\left(\sum_{j=0}^{i}\frac{\lambda_j}{1-\alpha_j}\right)} = \frac{\left(\sum_{j=0}^{i}\lambda_j\right)}{\left(\sum_{j=0}^{i}\frac{\lambda_j}{1-\alpha_j}\right)} - \left(G\left(\frac{1}{k}\right)+\frac{1}{k}g\left(\frac{1}{k}\right)\right) \quad \forall i \in \{1,\cdots,m\}$$

Note that functions $w_i(k)$ have the same s_3^i as that of functions $u_i(k)$ while the are easier to work with. In the following we prove that functions $w_i(k)$ are nonincreasing in i for all k. For this purpose, we first show that the constant term of function $w_i(k)$ is non-decreasing. To prove, we rewrite it as follows

$$\frac{1}{\left(\frac{\sum_{j=0}^{i}\frac{\lambda_j}{1-\alpha_j}}{\sum_{j=0}^{i}\lambda_j}\right)} = \frac{1}{\sum_{j=1}^{i}t_j^i\left(\frac{1}{1-\alpha_j}\right)}$$

where $t_j^i = \frac{\lambda_j}{\sum_{j=0}^{i}\lambda_j}$. Terms $a_i := \sum_{j=1}^{i}t_j\left(\frac{1}{1-\alpha_j}\right)$ are convex combination of increasing constants $1/(1-\alpha_j)$. Therefore, a_i's are non-decreasing and their inverse i.e. $(1)(a_i)$ non-increasing. So, $w_i(k)$ are non-increasing in i for all k. We know from definition of s_3^0 that function $u_0(k)$ and thereby $w_0(k)$ are nonpositive up to point s_3^0. Moreover, we proved the graph of function $w_1(k)$ is not above the graph of function $w_0(k)$ or equivalently speaking $w_1(k) \leq w_0(k)$ for all k. Therefore s_3^1 is greater than or equal to s_3^0. Following the same logic we can prove $s_3^1 \leq s_3^2$ and so on and so forth.

If we take derivative of function $u_i(k)$ we obtain

$$\frac{du_i(k)}{dk} = \frac{1}{k^2\left(\sum_{j=0}^{i}\frac{\lambda_j}{1-\alpha_j}\right)}\left(2g\left(\frac{1}{k}\right)+\frac{1}{k}g'\left(\frac{1}{k}\right)\right)$$

Due to $s_3^i \geq s_3^0$ and assumption 4 as well as $\lambda_i s$ being nonnegative, the expression $g\left(\frac{1}{k}\right) + \frac{1}{k} g'\left(\frac{1}{k}\right)$ is positive for all $k \geq s_3^i$. Therefore, the first derivative of function $u_i(k)$ is positive and $u_i(k)$ is increasing for all $k \geq s_3^i$. □

Assumption 5. *The elasticity of demand, $\epsilon(p)$, which is defined as $\epsilon(p) := pD'(p)/D(p)$, is nonincreasing, i.e., $\epsilon'(p) \leq 0$.*

Assumption 5 holds for all demand functions in table 1 with positive parameter values.

We are now in a position to present our sufficient optimality condition based on duality theory. If at one feasible solution of problem $\mathcal{P}$, the sufficient optimality condition holds then the duality gap is zero and that point is the global optimal solution. Since the problem is non-convex it is indeed possible that the duality gap is not zero.

Proposition 2. *(Sufficient optimality condition) Suppose assumptions 4 and 5 hold. Let (p^*, k^*) be a local maximum of problem $\mathcal{P}$ and λ^* be an associated Lagrange multiplier vector. Then (p^*, k^*) is global optimum if*

$$\left(a(k)S\left(\frac{1}{k}\right) + r(k)\right)\left(b(k) - a(k)G\left(\frac{1}{k}\right) - \frac{a(k-)}{k}g(\frac{1}{k})\right) - \frac{1}{k^2}g\left(\frac{1}{k}\right)\left(b(k) - a(k)G\left(\frac{1}{k}\right)\right) < 0 \tag{23}$$

for all $k \in A(\lambda^*)$ where
$A(\lambda^*) = \left\{k_i, i = 1, \cdots, m | \lambda_i^* > 0, \frac{c\left(1+\sum_N \lambda_i^*\right)}{a(k_i-)S\left(\frac{1}{k_i}\right)+r(k_i-)} \leq p_m\right\}$ *, and $f(x-) := \lim_{y \nearrow x} f(x)$*

Proof. The sketch of proof is based on Theorem Appendix A.2. We find a condition which assures the uniqueness of solution for the Lagrange Problem at $\lambda = \lambda^*$. The Dual function at λ^* for problem $\mathcal{P}$ is as follows

$$f(\lambda^*) = \max_{(p,k)\geq 0} L(p, k, \lambda^*)$$

Where

$$L(p, k, \lambda^*) = \left[a_2(k, \alpha_0) + \sum_{I(k)} \lambda_i^* a_1(k, \alpha_i) + \sum_{J(k)} \lambda_i^* a_2(k, \alpha_i)\right] D(p) - ck\left(1 + \sum_N \lambda_i^*\right) D(p)$$

Now, we need to show (p^*, k^*) is the unique solution of the Lagrange Problem for the associated Lagrange vector λ^*. We know that the optimal solution for problem $\max_{(p,k)\geq 0} L(p, k, \lambda^*)$ is interior. Therefore, the first order optimality conditions are necessary and sufficient for being local extrema. Thus, we have

$$\begin{aligned}\partial L(p,k,\lambda^*)/\partial p = &\left[a_2\left(k,\alpha_0\right)+\sum_{I(k)}\lambda_i^* a_1\left(k,\alpha_i\right)+\sum_{J(k)}\lambda_i^* a_2\left(k,\alpha_i\right)\right]\left(D\left(p\right)+pD'\left(p\right)\right)\\ &-ck\left(1+\sum_{N}\lambda_i^*\right)D'\left(p\right)=0 \end{aligned}\tag{24}$$

$$\begin{aligned}\partial L(p,k,\lambda^*)/\partial k = &\left[\frac{\partial a_2\left(k,\alpha_0\right)}{\partial k}+\sum_{I(k)}\lambda_i^*\frac{\partial a_1\left(k,\alpha_i\right)}{\partial k}+\sum_{J(k)}\lambda_i^*\frac{\partial a_2\left(k,\alpha_i\right)}{\partial k}\right]pD\left(p\right)\\ &-c\left(1+\sum_{N}\lambda_i^*\right)D\left(p\right)=0 \end{aligned}\tag{25}$$

Multiplying equation (25) by $-k\frac{D'(p)}{D(p)}$ and adding the outcome to equation (24), we obtain

$$\left[a_2\left(k,\alpha_0\right)+\sum_{I(k)}\lambda_i^* a_1\left(k,\alpha_i\right)+\sum_{J(k)}\lambda_i^* a_2\left(k,\alpha_i\right)\right]D\left(p\right)+\tag{26}$$

$$\left[1-\frac{G\left(\frac{1}{k}\right)}{1-\alpha_0}+\sum_{J(k)}\lambda_i^*\left(1-\frac{G\left(\frac{1}{k}\right)}{1-\alpha_i}\right)\right]pD'\left(p\right)=0\tag{27}$$

$$\left[\frac{\partial a_2\left(k,\alpha_0\right)}{\partial k}+\sum_{I(k)}\lambda_i^*\frac{\partial a_1\left(k,\alpha_i\right)}{\partial k}+\sum_{J(k)}\lambda_i^*\frac{\partial a_2\left(k,\alpha_i\right)}{\partial k}\right]pD\left(p\right)-\tag{28}$$

$$c\left(1+\sum_{N}\lambda_i^*\right)D\left(p\right)=0\tag{29}$$

Rearranging the first and the second equations, we get

$$\frac{a_2\left(k,\alpha_0\right)+\sum_{I(k)}\lambda_i^* a_1\left(k,\alpha_i\right)+\sum_{J(k)}\lambda_i^* a_2\left(k,\alpha_i\right)}{1-\frac{G\left(\frac{1}{k}\right)}{1-\alpha_0}+\sum_{J(k)}\lambda_i^*\left(1-\frac{G\left(\frac{1}{k}\right)}{1-\alpha_i}\right)}+p\frac{D'\left(p\right)}{D\left(p\right)}=0$$

$$p=\frac{c\left(1+\sum_N\lambda_i^*\right)}{\frac{\partial a_2(k,\alpha_0)}{\partial k}+\sum_{I(k)}\lambda_i^*\frac{\partial a_1(k,\alpha_i)}{\partial k}+\sum_{J(k)}\lambda_i^*\frac{\partial a_2(k,\alpha_i)}{\partial k}}$$

Substituting for functions $a_2\left(k,\alpha\right)$, and $a_1\left(k,\alpha\right)$ and simplifying, we obtain

$$1+\frac{k\frac{S\left(\frac{1}{k}\right)}{1-\alpha_0}+\sum_{I(k)}\lambda_i^* k\frac{S\left(G^{-1}(1-\alpha_i)\right)}{1-\alpha_i}+\sum_{J(k)}\lambda_i^* k\frac{S\left(\frac{1}{k}\right)}{1-\alpha_i}}{1-\frac{G\left(\frac{1}{k}\right)}{1-\alpha_0}+\sum\limits_{J(k)}\lambda_i^*\left(1-\frac{G\left(\frac{1}{k}\right)}{1-\alpha_i}\right)}+p\frac{D'(p)}{D(p)}=0$$

$$p=\frac{c\left(1+\sum_N\lambda_i^*\right)}{\frac{S\left(\frac{1}{k}\right)}{1-\alpha_0}+\sum_{I(k)}\lambda_i^*\frac{S(G^{-1}(1-\alpha_i))}{1-\alpha_i}+\sum_{J(k)}\lambda_i^*\frac{S\left(\frac{1}{k}\right)}{1-\alpha_i}}$$

Which can be written as

$$1+\frac{a(k)\,kS\left(\frac{1}{k}\right)+kr(k)}{b(k)-a(k)\,G\left(\frac{1}{k}\right)}+\epsilon(p)=0$$

$$p=\frac{c\left(1+\sum_N\lambda_i^*\right)}{a(k)\,S\left(\frac{1}{k}\right)-r(k)}$$

For simplification, we define

$$\begin{aligned}
f_1(k) &:= 1+\frac{a(k)\,kS\left(\frac{1}{k}\right)+kr(k)}{b(k)-a(k)\,G\left(\frac{1}{k}\right)}\\
f_2(k) &:= \epsilon(p(k))\\
p(k) &:= \frac{c\left(1+\sum_N\lambda_i^*\right)}{a(k)\,S\left(\frac{1}{k}\right)+r(k)}\\
f(k) &:= f_1(k)+f_2(k)
\end{aligned}$$

We need to show function $f(k)$ has a unique root. If we prove function $f(k)$ is strictly decreasing on its domain and it has at least one root then uniqueness follows.For this purpose, We take first derivatives of functions $f_1(k)$ and $f_2(k)$. Note that functions $a(k)$, $b(k)$ and $r(k)$ have zero derivative everywhere except at k_is which their derivatives are undefined. But expressions $a(k)\,S\left(\frac{1}{k}\right)+r(k)$ and $b(k)-a(k)\,G\left(\frac{1}{k}\right)$ are differentiable for all $k>0$. To find first derivative of function $f(k)$, we find first derivatives of functions $f_1(k)$ and $f_2(k)$:

$$\frac{df_1(k)}{dk}=\frac{\left(a(k)S(\frac{1}{k})-\frac{a(k)}{k^2}g(\frac{1}{k})+r(k)\right)\left(b(k)-a(k)G\left(\frac{1}{k}\right)\right)-\frac{a(k)}{k}g\left(\frac{1}{k}\right)\left(a(k)S(\frac{1}{k})+r(k)\right)}{\left(b(k)-a(k)G\left(\frac{1}{k}\right)\right)^2}$$

$$\frac{df_2(k)}{dk}=\epsilon'(p(k))p'(k)=\epsilon'(p(k))\frac{\frac{a(k)}{k^3}g(\frac{1}{k})c[1+\sum_N\lambda_i^*]}{\left(a(k)S\left(\frac{1}{k}\right)+r(k)\right)^2}$$

By assumption 5 $df_2(k)/dk$ is clearly non-positive. If the numerator of $df_1(k)/dk$ is negative everywhere as well, thereby $df(k)/dk$ is negative everywhere and we are done. We denote the numerator of $df_1(k)/dk$ by function $h(k)$. The first derivative of $h(k)$ is

$$
\begin{aligned}
h'(k) = &\left[\frac{a(k)}{k^3} g\left(\frac{1}{k}\right) + \frac{a(k)}{k^4} g'\left(\frac{1}{k}\right)\right]\left(b(k) - a(k)G\left(\frac{1}{k}\right)\right) \\
&+ \frac{a(k)}{k^2} g\left(\frac{1}{k}\right)\left[a(k)S\left(\frac{1}{k}\right) - \frac{a(k)}{k^2} g\left(\frac{1}{k}\right) + r(k)\right] \\
&+ \left[\frac{a(k)}{k^2} g\left(\frac{1}{k}\right) + \frac{a(k)}{k^3} g'\left(\frac{1}{k}\right)\right]\left[a(k)S\left(\frac{1}{k}\right) + r(k)\right] + \frac{a^2(k)}{k^4} g^2\left(\frac{1}{k}\right) \\
= &\frac{a(k)}{k^3}\left[g\left(\frac{1}{k}\right) + \frac{1}{k} g'\left(\frac{1}{k}\right)\right]\left(b(k) - a(k)G\left(\frac{1}{k}\right)\right) + \frac{a(k)}{k^2} g\left(\frac{1}{k}\right)\left[a(k)S\left(\frac{1}{k}\right) + r(k)\right] \\
&+ \frac{a(k)}{k^2}\left[g\left(\frac{1}{k}\right) + \frac{1}{k} g'\left(\frac{1}{k}\right)\right]\left[a(k)S\left(\frac{1}{k}\right) + r(k)\right] \\
= &\frac{a(k)}{k^3}\left[g\left(\frac{1}{k}\right) + \frac{1}{k} g'\left(\frac{1}{k}\right)\right]\left(b(k) - a(k)G\left(\frac{1}{k}\right)\right) + \\
&\frac{a(k)}{k^2}\left[2g\left(\frac{1}{k}\right) + \frac{1}{k} g'\left(\frac{1}{k}\right)\right]\left[a(k)S\left(\frac{1}{k}\right) + r(k)\right]
\end{aligned}
$$

Now, let us get back to function $h(k)$. With rearranging it, we obtain

$$
\begin{aligned}
h(k) = &\left(a(k)S(\frac{1}{k}) + r(k)\right)\left(b(k) - a(k)G\left(\frac{1}{k}\right) - \frac{a(k)}{k} g\left(\frac{1}{k}\right)\right) - \\
&\frac{a(k)}{k^2} g(\frac{1}{k})\left(b(k) - a(k)G\left(\frac{1}{k}\right)\right)
\end{aligned}
$$

The expression $b(k) - a(k)G\left(\frac{1}{k}\right) - \frac{a(k)}{k} g\left(\frac{1}{k}\right)$ is equal to $u_i(k)$ for all $k \in [k_i, k_{i+1})$. We investigate to cases:

Case 1: $s_3^i \geq k_{i+1}$

In this case, based on lemma 5, function $b(k) - a(k)G\left(\frac{1}{k}\right) - \frac{a(k)}{k} g\left(\frac{1}{k}\right)$ is negative in $[k_i, k_{i+1})$. So, $h(k) < 0$ and as the result of that, $df_1(k)/dk$ and $df(k)/dk$ are negative in $[k_i, k_{i+1})$.

Case 2: $s_3^i < k_{i+1}$

We know from lemma 5, the function $u_i(k)$ is negative up to s_3^i. Following lemma 5 and assumption 4, function $h(k)$ is increasing for all $k \geq s_3^i$. Therefore, if $h(k_{i+1}-)$ is negative, it means $h(k)$ is negative for all $k \in [k_i, k_{i+1})$. Condition (23) insures $h(k_i-)$ is negative for all i[5]. Note that we do not need to check the condition for all k_i. The evaluation of the condition only for those k_is which their corresponding λ_i^*s are not zero, is sufficient.

We showed that $f(k)$ is decreasing. So it can at most have one root. At the other hand we know that k^* is a root of $f(k)$.Thus $f(k)$ has a unique root. Using the strong duality theorem Appendix A.2, the given local optimum (i.e., (p^*, k^*)) is indeed global

[5]Since $b(k) - a(k)G\left(\frac{1}{k}\right)$ and $a(k)S(\frac{1}{k}) + r(k)$ are continuous, $h(k-)$ is equal to

$$
h(k-) = \left(a(k)S\left(\frac{1}{k}\right) + r(k)\right)\left(b(k) - a(k)G\left(\frac{1}{k}\right) - \frac{a(k-)}{k} g(\frac{1}{k}) - \frac{1}{k^2} g\left(\frac{1}{k}\right)\left(b(k) - a(k)G\left(\frac{1}{k}\right)\right)\right)
$$

It is also insightful if we investigate $h(k+)$. Following the previous lines and $a(k-)$ being less than or equal to $a(k+)$, it is clear that $h(k+) < h(k-)$

optimum. □

5. An Example

In this section, we present an example to further illustrate the applicability of the model. Assume an influenza vaccine producer wants to determine the optimal price and optimal target production quantity in advance. Assume the variable cost of production is 6$ per dose, demand function is equal to

$$D(p) = 4 \times 10^6 - 10^5 p$$

, The distribution of yield is Beta(4,2) with CDF:

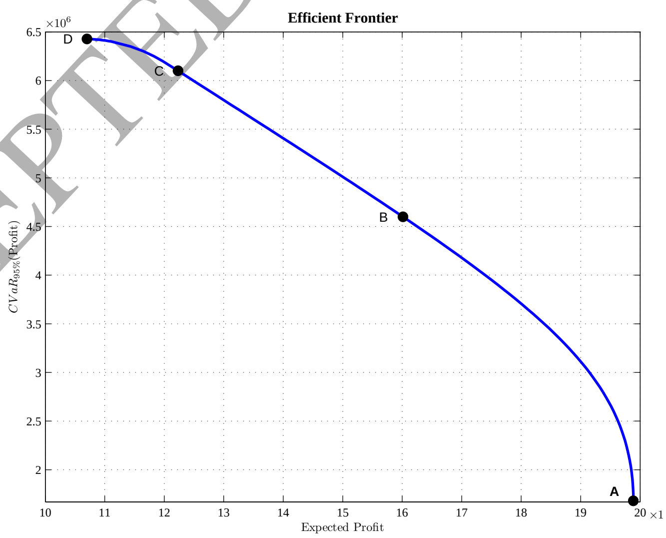
$$G(z) = \int_0^z \frac{\Gamma(6)}{\Gamma(4)\Gamma(2)} z^3(1-z)dz$$

The producer wants to maximize the expected profit while minimizing the risk of low profits. She measures the risk of low profits by conditional expectation of lowest 5 percent of profits i.e., $\alpha = 0.95$. Therefore, the producer's problem is modeled as follows:

$$\begin{aligned} [\mathcal{P}]\, \max & \phi_0\,(p,k) \\ & \phi_{0.95}\,(p,k) \geq v \\ & p,k \geq 0 \end{aligned}$$

We can find the efficient frontier by changing v. The efficient frontier for this producer is shown in Figure 1.

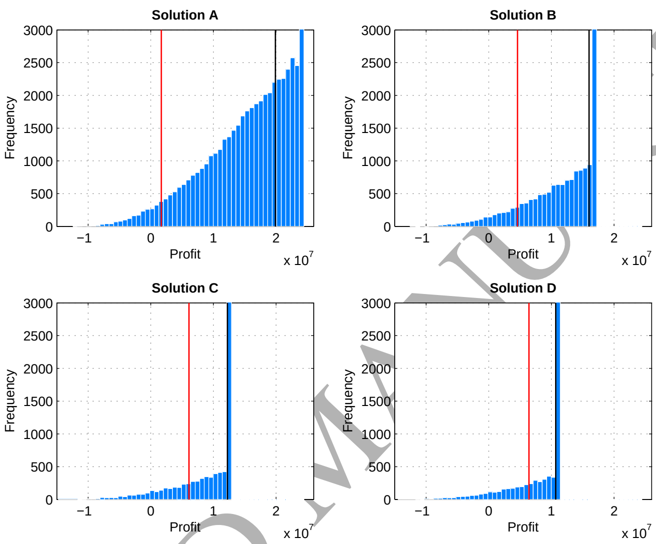
Figure 1: Efficient Frontier



We picked four solutions on the graph to compare. The characteristics of these solutions are presented in Table 2.

Solution A is for a risk neutral person since it is the highest achievable expected profit. Other solutions are for risk averse persons. The more a person is risk averse, the more he tries to avoid high losses. The purpose of the risk constraint is to limit this risk. The higher the risk threshold (i.e., v), the lower the expected value at the optimal solution (i.e., $\phi_0(p^*, k^*)$) as it is apparent in the solutions B,C, and D. Let us compare solutions A and B. A risk averse decision maker prefers Solution B to A since the risk of low profits is much lower. To be exact, $CVaR$ at level 95 percent is about 4.6 millions for solution B compared to about 1.7 million for solution A. It is about 2.7 times that of A. On the other hand, this comes at the price of losing some high profits. The expected profit decreases from about 20 millions to 16 millions (i.e., about 19 percent). Histograms shown in Figure 2 give you a nice perspective on the various aspects of risk implications of solutions A to D. The red and black lines (i.e., left and right vertical lines respectively) denote the expected profit and conditional expected profit at level 95 percent respectively.

Figure 2: Probability Distributions of Profit



As we go from A to D the probability of high losses decreases, at the same time the probability of high profits decreases too.

6. Conclusion

In this paper, we investigated production planning problem with pricing under random yield for a risk averse decision maker. We used CVaR measure to shape the distribution of profit function according to preferences of decision maker. One of the advantages of using CVaR is that it can easily deal with different risk aversion levels in a coherent way.

Table 2: Representative Efficient Solutions

Solution	(2) $E(Profit)$ in Millions	(3) $CVaR_{95\%}(Profit)$ in Millions	(4) Total Cost in Millions	(5) $\frac{(2)}{(3)}$	(6) p^*	(7) k^*	(8) q^*	(9) $D(p^*)$
A	19.88	1.68	13.38	8.4	25	1.49	2,230,638	1,496,430
B	16.01	4.6	14.21	28.7	29.4	2.24	2,367,688	1,055,616
C	12.23	6.1	16.75	49.9	30.3	2.89	2,792,421	966,033
D	10.70	6.43	16.50	60.1	31.2	3.14	2,749,471	875,927

In another word, as we showed in our model, we can incorporate several aspects of risk aversion of a decision maker without considerably increasing computational burden of the problem and without losing the coherency in tackling the risk. For solving our model, we developed an efficient inexact solution algorithm. We also proposed an important saddle point theorem (theorem Appendix A.3) which its usage is far beyond this paper. This theorem enabled us to characterize the set of saddle points by KKT conditions. Having the characterization of the set of saddle points of a Lagrangean function is a requirement in using envelope theorems.

Acknowledgements

This research has benefited greatly from support of Professor Roger Wets (Department of Mathematical Sciences, University of California at Davis) by hosting the first author. We should thank him for all his guidance, and kindness.

References

[1] Vipul Agrawal and Sridhar Seshadri. Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem. *Manufacturing & Service Operations Management*, 2(4):410–423, 2000.

[2] Kenneth J. Arrow. *Essays in the theory of risk-bearing*. Markham Pub. Co., Chicago, 1971.

[3] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

[4] Ismail Serdar Bakal and Elif Akcali. Effects of random yield in remanufacturing with price-sensitive supply and demand. *Production and operations management*, 15(3): 407–420, 2006.

[5] Aharon Ben-Tal and Marc Teboulle. An Old-New Concept Of Convex Risk Measures: The Optimized Certainty Equivalent. *Mathematical Finance*, 17(3):449–476, 2007.

[6] Dimitri P. Bertsekas. *Nonlinear programming*. Athena Scientific, Belmont, Mass., 2nd edition, 1999.

[7] Stephen P. Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, UK ; New York, 2004.

[8] Youhua Chen, Minghui Xu, and Zhe George Zhang. Technical note–a risk-averse newsvendor model under the cvar criterion. *Operations research*, 57(4):1040–1044, 2009.

[9] Soo-Haeng Cho and Christopher S. Tang. Advance Selling in a Supply Chain under Uncertain Supply and Demand. *SSRN eLibrary*, 2011.

[10] Sungyong Choi, Andrzej Ruszczyński, and Yao Zhao. A multiproduct risk-averse newsvendor with law-invariant coherent measures of risk. *Operations Research*, 59 (2):346–364, 2011.

[11] Tsan-Ming Choi, Duan Li, and Houmin Yan. Mean-variance analysis for the newsvendor problem. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 38(5):1169 –1180, 2008.

[12] Maqbool Dada, Nicholas C. Petruzzi, and Leroy B. Schwarz. A newsvendor's procurement problem when suppliers are unreliable. *Manufacturing and Service Operations Management*, 9(1):9–32, 2007.

[13] Sarang Deo and Charles J. Corbett. Cournot competition under yield uncertainty: The case of the u.s. influenza vaccine market. *Manufacturing and Service Operations Management*, 11(4):563–576, 2009.

[14] Louis Eeckhoudt, Christian Gollier, and Harris Schlesinger. The risk-averse (and prudent) newsboy. *Management Science*, 41(5):786–794, 1995.

[15] Saman Eskandarzadeh, Kourosh Eshghi, Mohammad Modarres Yazdi, and Mohsen Bahramgiri. Production planning problem with pricing under random yield: Cvar criterion. *Journal of Systems Science and Systems Engineering*, 23(3):312–328, 2014.

[16] Hans Föllmer and Alexander Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6:429–447, 2002.

[17] Hans Föllmer and Alexander Schied. *Stochastic Finance. An Introduction in Discrete Time*. de Gruyter Studies in Mathematik. Walter de Gruyter, Berlin, New York, 2 edition, 2004.

[18] Xianghua Gan, Suresh P Sethi, and Houmin Yan. Channel coordination with a risk-neutral supplier and a downside-risk-averse retailer. *Production and Operations Management*, 14(1):80–89, 2005.

[19] Vishal Gaur and Sridhar Seshadri. Hedging inventory risk through market instruments. *Manufacturing & Service Operations Management*, 7(2):103–120, 2005.

[20] Jun-ya Gotoh and Yuichi Takano. Newsvendor solutions via conditional value-at-risk minimization. *European Journal of Operational Research*, 179(1):80 – 96, 2007.

[21] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):pp. 263–292, 1979.

[22] Burak Kazaz and Scott Webster. The impact of yield-dependent trading costs on pricing and production planning under supply uncertainty. *Manufacturing and Service Operations Management*, 13(3):404–417, 2011.

[23] Hon-Shiang Lau. The newsboy problem under alternative optimization objectives. *Journal of the Operational Research Society*, pages 525–535, 1980.

[24] Qing Li and Shaohui Zheng. Joint inventory replenishment and pricing control for systems with uncertain yield and demand. *Operations Research*, 54(4):696–705, 2006.

[25] Mark J. Machina. Choice under uncertainty: Problems solved and unsolved. *The Journal of Economic Perspectives*, 1(1):pp. 121–154, 1987.

[26] Victor Martínez-de Albéniz and David Simchi-Levi. Mean-variance trade-offs in supply contracts. *Naval Research Logistics (NRL)*, 53(7):603–616, 2006.

[27] Paul Milgrom and Ilya Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601, 2002.

[28] Aysun Özler, Barış Tan, and Fikri Karaesmen. Multi-product newsvendor problem with value-at-risk considerations. *International Journal of Production Economics*, 117(2):244–255, 2009.

[29] Ralph Tyrrell Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, N.J., 1970.

[30] Ralph Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.

[31] Ralph Tyrrell Rockafellar and Stanislav Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7):1443 – 1471, 2002.

[32] Ralph Tyrrell Rockafellar, Stanislav Uryasev, and Michael Zabarankin. Generalized deviations in risk analysis. *Finance and Stochastics*, 10:51–74, 2006.

[33] Carl P. Simon and Lawrence Blume. *Mathematics for economists.* Norton, New York, 1994.

[34] M Mahdi Tajbakhsh, Saeed Zolfaghari, and Chi-Guhn Lee. Supply uncertainty and diversification: a review. In *Trends in Supply Chain Design and Management*, pages 345–368. Springer, 2007.

[35] Christopher S. Tang and Rui Yin. Responsive pricing under supply uncertainty. *European Journal of Operational Research*, 182(1):239 – 255, 2007.

[36] Ou Tang, S. Nurmaya Musa, and Juan Li. Dynamic pricing in the newsvendor problem with yield risks. *International Journal of Production Economics*, 2011.

[37] Brian Tomlin. On the value of mitigation and contingency strategies for managing supply chain disruption risks. *Management Science*, 52(5):639–657, 2006.

[38] Brian Tomlin and Yimin Wang. On the value of mix flexibility and dual sourcing in unreliable newsvendor networks. *Manufacturing and Service Operations Management*, 7(1):37–57, 2005.

[39] Hirofumi Uzawa. The kuhn-tucker theorem in concave programming. In Kenneth H. Arrow, Leonid Hurwicz, and Hirofumi Uzawa, editors, *Studies in Linear and Nonlinear Programming*, pages 32–37. Stanford University Press, Stanford CA, 1958.

[40] Yimin Wang, Wendell Gilland, and Brian Tomlin. Mitigating supply risk: Dual sourcing or process improvement? *Manufacturing & Service Operations Management*, 12(3):489–510, 2010.

[41] Minghui Xu. A price-setting newsvendor model under cvar decision criterion with emergency procurement. *Journal of Systems Science and Systems Engineering*, 19: 85–104, 2010.

[42] Minghui Xu and Jianbin Li. Optimal decisions when balancing expected profit and conditional value-at-risk in newsvendor models. *Journal of Systems Science and Complexity*, 23(6):1054–1070, 2010.

[43] Minghui Xu and Ye Lu. The effect of supply uncertainty in price-setting newsvendor models. *European Journal of Operational Research*, 2013.

[44] C. A. Yano and H. L. Lee. Lot sizing with random yields: a review. *Operations Research*, 43(2):311–334, 1995.

Appendix A.

Lemma Appendix A.1. *Function $\phi_\alpha(p,k)$ is a decreasing function of α.*

Proof. First, we take first derivative of $a(k,\alpha)$. We have two cases to consider:

Case 1: $k < \frac{1}{G^{-1}(1-\alpha)}$

$$\frac{\partial a(k,\alpha)}{\partial\alpha} = \frac{\partial a_1(k,\alpha)}{\partial\alpha} = \frac{-\frac{1}{g(G^{-1}(1-\alpha))}G^{-1}(1-\alpha)g(G^{-1}(1-\alpha))(1-\alpha)+S(G^{-1}(1-\alpha))}{(1-\alpha)^2}$$
$$= \frac{\frac{S(G^{-1}(1-\alpha))}{1-\alpha} - G^{-1}(1-\alpha)}{(1-\alpha)} < 0$$

The inequality follows from the fact that $\frac{S(G^{-1}(1-\alpha))}{1-\alpha}$ is the conditional expectation of $\tilde{y}$ knowing that $\tilde{y} \le G^{-1}(1-\alpha)$. So it is less than $G^{-1}(1-\alpha)$. In other words, we have:

$$\frac{S(G^{-1}(1-\alpha))}{1-\alpha} = \mathbf{E}\left(\tilde{y}\,\middle|\tilde{y} \le G^{-1}(1-\alpha)\right)$$

Case 2: $k \ge \frac{1}{G^{-1}(1-\alpha)}$

$$\frac{\partial a(k,\alpha)}{\partial\alpha} = \frac{\partial a_2(k,\alpha)}{\partial\alpha} = \frac{kS\left(\frac{1}{k}\right) - G\left(\frac{1}{k}\right)}{(1-\alpha)^2} = kG\left(\frac{1}{k}\right)\frac{\frac{S\left(\frac{1}{k}\right)}{G\left(\frac{1}{k}\right)} - \frac{1}{k}}{(1-\alpha)^2} < 0$$

The inequality follows the same argument as that of case 1. So,

$$\frac{\partial\phi_\alpha(p,k)}{\partial\alpha} = \frac{\partial a(k,\alpha)}{\partial\alpha}pD(p) < 0$$

□

Lemma Appendix A.2. *(e.g. Boyd and Vandenberghe [7]) Function $f(x): R_+ \to R$ is quasiconcave if, $f'(x) = 0 \Rightarrow f''(x) < 0$*

Theorem Appendix A.1 (Milgrom and Segal [27])**.** *Let X and Y be compact spaces and suppose that $f(x,y,t): X \times Y \times [0,1] \to R$ and $\frac{\partial f(x,y,t)}{\partial t}$ are continuous functions. Suppose also that the set of saddle points are not empty for all $t \in [0,1]$. In this case, the set of all saddle points can be represented by set $X^*(t) \times Y^*(t)$ for all $t \in [0,1]$. Then*

$V(t)$ is directionally differentiable, and the directional derivatives are

$$V'(t+) = \lim_{u \downarrow t} V'(u) = \max_{x \in X^*(t)} \min_{y \in Y^*(t)} f_t(x,y,t) = \min_{y \in Y^*(t)} \max_{x \in X^*(t)} f_t(x,y,t) \quad \textit{for all } t < 1$$
$$V'(t-) = \lim_{u \uparrow t} V'(u) = \min_{x \in X^*(t)} \max_{y \in Y^*(t)} f_t(x,y,t) = \max_{y \in Y^*(t)} \min_{x \in X^*(t)} f_t(x,y,t) \quad \textit{for all } t > 0$$

Where

$$\begin{aligned} V(t) &= \sup_{x \in X} \inf_{y \in Y} f(x,y,t), t \in [0,1] \\ X^*(t) &= \underset{x \in X}{Argmax} \min_{y \in Y} f(x,y,t) \\ Y^*(t) &= \underset{y \in Y}{Argmin} \max_{x \in X} f(x,y,t) \end{aligned}$$

The following duality theorem is an important theorem for proving several results in this paper and because of its generality, its application is by far beyond this paper.

Theorem Appendix A.2. *Let functions $f(x)$ and $f_i(x)$, $i = 1 \cdots m$ be $\mathcal{C}^1$ (i.e. have continuous first directional derivatives everywhere) . Then, a regular local maximum point of the problem (we denote it by x^*)*

$$\begin{aligned} &\max f(x) \\ &s.t. \\ &\quad f_i(x) \geq 0, \qquad i = 1, \cdots, m \\ &\quad x \geq 0 \end{aligned} \tag{A.1}$$

is the global maximum point, if for an associated Lagrange multipliers vector $\lambda^* = (\lambda_i^*) \in \mathcal{R}_+^m$ (obtained from KKT conditions), the Lagrange problem

$$\max_{x \geq 0} f(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x)$$

has a unique local maximizer. In that case, the local maximizer of the Lagrange problem will be x^*

Proof. If x^* is a regular local maximum of problem A.1, then there is a Lagrange multiplier vector λ^* that along with x^* satisfies the Karush-Kuhn-Tucker(KKT) conditions (e.g. Bertsekas [6]). In other words,

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) \leq 0 \tag{A.2}$$

$$x_j \left(\nabla_j f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla_j f_i(x^*) \right) = 0, \qquad j = 1, \cdots, n \tag{A.3}$$

$$f_i(x^*) \geq 0, \qquad i = 1, \cdots, m \tag{A.4}$$

$$\lambda_i^* f_i(x^*) = 0, \qquad i = 1, \cdots, m \tag{A.5}$$

$$\lambda^*, x^* \geq 0 \tag{A.6}$$

From weak duality theorem, we know that for every feasible primal-dual pair (x, λ) we have

$$f(x) \leq L(\lambda) = \max_{x \geq 0} f(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$$

So,

$$f(x) \leq L(\lambda^*) = \max_{x \geq 0} f(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x)$$

The KKT necessary optimality conditions for problem $L(\lambda^*)$ are

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x) \leq 0 \tag{A.7}$$

$$x_j \left(\nabla_j f(x) + \sum_{i=1}^{m} \lambda_i^* \nabla_j f_i(x) \right) = 0, \qquad j = 1, \cdots, n \tag{A.8}$$

$$x \geq 0 \tag{A.9}$$

We know from the assumption of theorem there is a unique x satisfies A.7-A.9. Furthermore, we know x^* satisfies them. Therefore,

$$f(x) \leq L(\lambda^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i f_i(x^*) = f(x^*)$$

The equality follows from A.5. $\square$

Theorem Appendix A.3. *Consider the following problem:*

$$\begin{aligned} [\mathcal{H}] \max & f(x) \\ s.t. & \\ & f_i(x) \geq 0, \qquad i = 1, \cdots, m \\ & x \geq 0 \end{aligned}$$

where functions $f(x)$ and $f_i(x)$, $i = 1 \cdots m$ are $\mathcal{C}^1$ i.e. continuously differentiable. We denote the Lagrange function of $\mathcal{H}$ by $l(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$. Assume Lagrange function $l(x, \lambda)$ has a unique local maximum for all $\lambda \geq 0$ and furthermore, one type of constraint qualification conditions (C.Q. conditions) holds at optimal solution. Then

1) if $(\bar{x}, \bar{\lambda})$ is a saddle point of Lagrange function $l(x, \lambda)$ then it satisfies Karush-Kuhn-Tucker conditions (KKT in short)

2) if $(\bar{x}, \bar{\lambda})$ satisfies KKT conditions then it is a saddle point of Lagrange function $L(x, \lambda)$

Proof. Proof of part 1) From Theorem one in Uzawa [39] it immediately follows that if $(\bar{x}, \bar{\lambda})$ is a saddle point of Lagrange function then it is optimal point of $\mathcal{H}$. In the Uzawa's proof, complementary slackness and feasibility conditions of KKT are derived, Therefore, we only need to show that the saddle point $\left(\bar{x}, \bar{\lambda}\right)$ satisfies

$$\nabla f(x) + \sum_i \lambda_i f_i(x) \leq 0 \tag{A.10}$$

$$x^T \left(\nabla f(x) + \sum_i \lambda_i f_i(x)\right) = 0 \tag{A.11}$$

From Lemma 36.2 in Rockafellar [29], we can conclude

$$l\left(\bar{x}, \bar{\lambda}\right) = \max_{x \geq 0} l\left(x, \bar{\lambda}\right)$$

The first order optimality conditions for function $l\left(x, \bar{\lambda}\right)$ are exactly conditions A.10 and A.11. Therefore, along with complementary slackness and feasibility conditions, we have all KKT conditions being satisfied at saddle point $\left(\bar{x}, \bar{\lambda}\right)$

Proof of part 2) From complementary slackness assumption, we deduce

$$l\left(\bar{x}, \bar{\lambda}\right) = f(\bar{x}) + \sum_i \bar{\lambda} f_i(\bar{x}) = f(\bar{x}) \leq f(\bar{x}) + \sum_i \lambda f_i(\bar{x}) = l(\bar{x}, \lambda) \text{ for all } \lambda$$

Furthermore, since $\left(\bar{x}, \bar{\lambda}\right)$ satisfies the first order optimality conditions A.10 and A.11, along with this assumption that the local maximum of Lagrange function $l\left(x, \bar{\lambda}\right)$ is unique, we conclude that

$$l\left(x, \bar{\lambda}\right) \leq l\left(\bar{x}, \bar{\lambda}\right) = \max_{x \geq 0} l\left(x, \bar{\lambda}\right) \text{ for all } x \geq 0$$

To sum up, we proved the validity of the relation

$$l\left(x, \bar{\lambda}\right) \leq l\left(\bar{x}, \bar{\lambda}\right) \leq l(\bar{x}, \lambda) \text{ for all } x, \lambda \geq 0$$

Which is the definition of a saddle point □

Appendix B.

Proof of Proposition 1:

To prove $k^* > k_0$, we argue by contradiction. Suppose the opposite relation is true, i.e. $k^* < k_0$. Due to $G(x)$ being increasing, 1 $G^{-1}(1-\alpha)$ is an increasing function of α. Thus, we can deduce

$$k^* < k_0 < k_1 < \cdots < k_m$$

Considering the above relations, the extensive form of model $\mathcal{P}$ for $k < k^*$ can be expressed as follows:

$$\begin{aligned}
[\mathcal{P}] \max & \left[\frac{S\left(G^{-1}(1-\alpha_0)\right)}{1-\alpha_0} p - c\right] kD(p) \\
& \left[\frac{S\left(G^{-1}(1-\alpha_i)\right)}{1-\alpha_i} p - c\right] kD(p) \geq v_i, i = 1, \ldots, m \\
& p, k \geq 0
\end{aligned}$$

We consider two cases:

Case 1:$k^*D\left(p^*\right) > 0$

We can find a sufficiently small ϵ and δ in such a way that with increasing p^* to $p^* + \epsilon$ and k^* to $k^* + \delta$, the product $kD\left(p\right)$ remains constant and thereby both righthand side of constraints as well as the objective function increase. In other words we find another feasible solution with better objective function which is a contradiction.

Case 2:$k^*D\left(p^*\right) = 0$

In this case, the feasibility of the optimal solution requires that all v_i's are nonpositive. If at least one v_i is positive then this case does not happen. In this case, the objective function and the righthand side expressions are zero. Since assumption 1 holds, there exists point $(p_l + \epsilon, \epsilon)$ for sufficiently small ϵ in such a way that it meets all constraints and has positive objective value. This is against our assumption that (k^*, p^*) is optimal.

To prove $p^* > p$, we only need to partially write down KKT necessary optimality conditions (since we assumed (p^*, k^*) is regular, KKT conditions are necessary optimality conditions) :

$$\frac{\partial \phi_{\alpha_0}\left(p,k\right)}{\partial p} + \sum_i \lambda_i \frac{\partial \phi_{\alpha_i}\left(p,k\right)}{\partial p} \leq 0 \Rightarrow$$

$$\left(a\left(k,\alpha_0\right) + \sum_i \lambda_i a\left(k,\alpha_0\right)\right)\left(pD\left(p\right)\right)' - ck\left(1 + \sum_i \lambda_i\right) D'\left(p\right) \leq 0 \tag{B.1}$$

We know from the definition of p_0 (Refer to Lemma 2) that $\left(pD\left(p\right)\right)'$ is positive on $(0, p_0)$. Due to this fact and the negativity of $D'\left(p\right)$, the left hand side of inequality B.1 will be positive on $(0, p_0)$ which is a contradiction. Thus the optimal price can not be less than p_0 □