A Data Driven Functionally Robust Approach for Coordinating Pricing and Order Quantity Decisions with Unknown Demand Function

Jian Hu, Junxuan Li

Department of Industrial and Manufacturing Systems Engineering

University of Michigan - Dearborn

jianhu@umich.edu, junxuanl@umich.edu

Sanjay Mehrotra

Department of Industrial Engineering and Management Sciences

Northwestern University

mehrotra@northwestern.edu

Abstract

We consider a retailer's problem of optimal pricing and inventory stocking decisions for a product. We assume that the price-demand curve is unknown, but data is available that loosely specifies the price-demand relationship. We propose a conceptually new framework that simultaneously considers pricing and inventory decisions without a priori fitting a function to the price-demand data. The framework introduces a novel concept of functional robustness. We consider two situations: (i) where the price-demand function is decreasing convex, and (ii) where the price-demand function is decreasing concave. Under these assumptions the decision problem is to simultaneously specify the demand function, the optimal selling price and the order quantity. The solution of the proposed maxmin optimization model returns the demand function and the desired pricing and order quantity decisions. A cutting surface algorithm is developed for the convex case; and a monolithic reformulation is given for the concave case. The features of the model and the algorithm are illustrated using public data on porterhouse beef price and demand.

Key Words: Functionally robust optimization, Newsvendor problem, Coordinating pricing and inventory decision

1 Introduction

The profit of a retailer, who has purchased inventory in amount x at price p, sells the product at price s, and assumes that the sales price-demand relationship is given by a function $d(s)$, is given by

$$s \min\{x, d(s)\} - px.$$

Through out the paper we will assume that the purchase price p is set by the wholesaler. If we assume that $d(s)$ is known, the problem of determining the optimal selling price and the order quantity is straight forward. Now assume, however, that we have data points $(t_i, d(t_i)), i = 1, \ldots, n$, that provide selling price-demand relationship. An approach would be to assume the form (linear, quadratic, etc.) of the price-demand function, and find parameters that specify this function by a best fit model. Now consider a situation where we do not know the form of the price-demand function. In this paper we propose and study a conceptually novel framework, called functionally robust optimization, and use this framework to simultaneously decide the selling price s and the order quantity x. A major advantage of the proposed approach is that it allows us to consider both data-rich and data-poor situations. The later is often the case when price-demand data is obtained from expensive market studies.

The model studied in this paper is different from the classical newsvendor modeling approach, which assumes that the demand is stochastic with a known or partially known (e.g., using moments) probability distribution. The literature on the order quantity problem taking the stochastic approach (e.g., see Whitin (1955), Tayur et al. (1999), Yano and Gilbert (2004), Chan et al. (2004), and therein), that takes the newsvendor approach (e.g., see Petruzzi and Dada (1999), Kocabiyikoğlu and Popescu (2011) and therein) is rich. In the known approaches, the demand may be characterized in the additive, multiplicative, or mixed structured function form with random parameters capturing market relationship (Mills (1959), Karlin and Carr (1962), Young (1978), Chen and Simchi-Levi (2004)). A distributionally-robust model has been proposed to handle situations where the demand probability distribution is only partially specified (Scarf (1958), Gallego and Moon (1993)), and models have also been studied to minimize the maximum regret from ordering a suboptimal quantity (see, e.g., Perakis and Roels (2008), and Zhu et al. (2013). The distributionally robust model of Scarf (1958) has recently (e.g., see Pflug and Wozabal (2007), Bertsimas et al. (2010), Delage and Ye (2010), Mehrotra and Zhang (2014), Mehortra and Papp (2014)) provided a key motivation for studying problems outside the inventory domain, and in more general multivariate setting. In this sense, the functionally robust framework of this paper may have other extensions and applications.

Since we only assume that the price-demand data is known; and the demand function is unknown – we make the model well posed by specifying an uncertainty set that preserves the shape of the price-demand function, and a bound on the fitting error in the function obtained by the optimization model. The two cases we study in this paper assume that the shape of price-demand curve is either decreasing-convex, or decreasing-concave (e.g. Mankiw (2008), Malueg (1994)). Convex price-demand functions reflect the common market behavior where most customers will buy a product at low price, while a few will continue to purchase at higher prices (e.g.Allen (2008), Mas-Colell et al. (1995)). In comparison, concave price-demand curves arise in some imperfect competition situations where a large number of consumers will continue to hold on to the purchase with increasing price (Greenhut et al. (1987),Malueg (1994),Klemperer and Meyer (1986)), and then the demand drops suddenly.

1.1 Model Formulation

We now present our problem and the proposed decision framework more formally. Let us assume that $\mathfrak{S} := [\underline{s}, \overline{s}] \subseteq \mathbb{R}_+$ is the decision region of sales price, and $\mathfrak{D}$ describes a set of demand functions. There is no restriction on the order quantity. The functionally robust profit (FRP) model is as follows:

$$\max_{(s,x)\in\mathfrak{S}\times\mathbb{R}_+} \min_{d\in\mathfrak{D}} \left\{ s\min\{x, d(s)\} - px \right\}. \tag{FRP}$$

We now discuss a specification of the set $\mathfrak{D}$. Assume that t_i, for $i = 1, \ldots, n$, are ordered so that $t_1 < t_2 < \cdots < t_n$. The demand observed at price point t_i is represented by y_{ij} for $j = m_1, \ldots, m_n$, where m_i represents the number of different demand observations at t_i. Let the error in a function

$d(s)$ be given by

$$\mathrm{Err}(d) := \left(\frac{1}{\sum_{i=1}^{n} m_i} \sum_{i=1}^{n} \sum_{j=1}^{m_i} (y_{ij} - d(t_i))^2 \right)^{1/2}. \tag{1}$$

It is possible to use other metrics (e.g. ℓ_1-norm, ℓ_∞-norm) of fitting error, but in this paper we restrict to the case of the most commonly used least-squares metric (1). Let V denote the set of all positive, continuous, decreasing, and convex functions over $\mathbb{R}_+$, and Λ denote the set of all positive, continuous, decreasing, and concave functions over $\mathbb{R}_+$. In the classical approach the model form (e.g., linear, quadratic, etc.) of $d(\cdot)$ is specified, and the parameters of this form are determined by minimizing $\mathrm{Err}(d)$ as follows:

$$\min_{d \in \mathfrak{F}} \mathrm{Err}(d), \tag{2}$$

where $\mathfrak{F} \subseteq V$(resp., Λ). Now (FRP) is specified by considering the set of convex and concave demand functions:

$$\mathfrak{D}_V := \{d \in V : \mathrm{Err}(d) \le \epsilon\}, \tag{3}$$

$$\mathfrak{D}_\Lambda := \{d \in \Lambda : \mathrm{Err}(d) \le \epsilon\}. \tag{4}$$

In our context, we make a few additional assumptions: $p < \underline{s}$, i.e., the optimal selling price is greater than the purchase price; and to properly model the boundary behavior of the unknown demand function that

$$\underline{s} \ge t_2, \quad \overline{s} \le t_{n-1}, \tag{5}$$

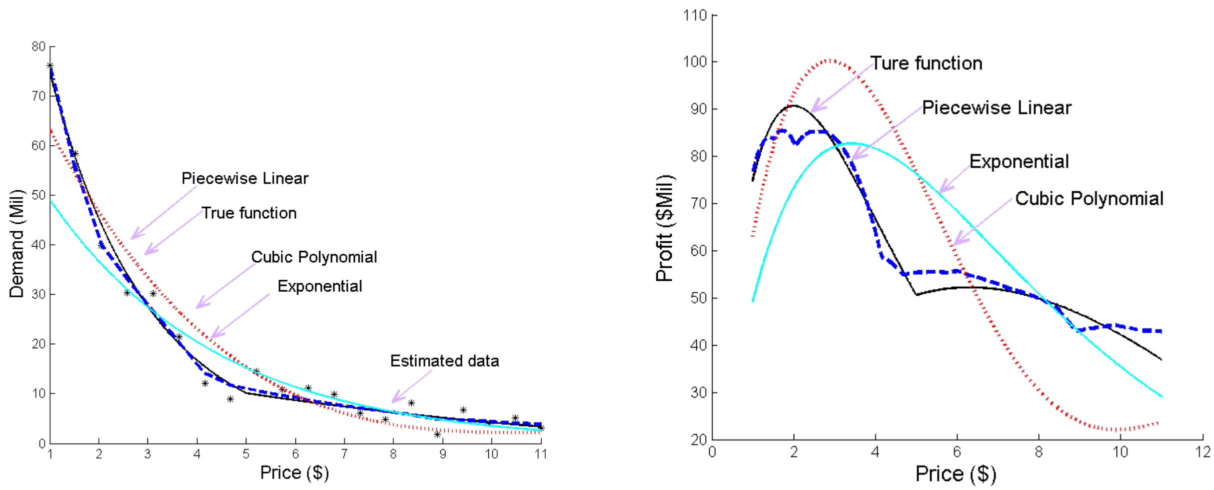
i.e., price-demand data is collected at points t_1 and t_n, which are outside the range in which we would like to set the selling price.

1.2 A Motivating Example

We now illustrate the potential value in taking the non-parametric 'function-free' approach in the functionally robust framework with the help of a synthetic example. For simplicity assume that the purchase price $p = 0$, and that the true demand function is given by $h(s)$ as follows:

$$h(s) := \begin{cases} \frac{55}{2} e^{-\frac{s-3}{2}} & 1 \le s \le 5, \\ \frac{55}{2e} - 15\left(1 - e^{-\frac{s-5}{10}}\right) & 5 < s \le 11, \end{cases} \qquad g(s) := \sqrt{150 - 10s}.$$

The optimal price in this case is \$2, and the maximum profit is \$90.68 million. Figure 1a gives



(a) (b)

Figure 1: Data and Performance of Different Model Fitting Techniques

possible demand responses at twenty different price points ($n = 20$) between \$ 1 and \$ 11 from a hypothetical marketing study. The demand data shown in Figure 1a is a sample generated from the function $d(s) := h(s) + \xi g(s)$, where ξ is a random variable following a shifted Beta distribution given by Beta$(2, 2) - 1/2$. Figure 1a gives the corresponding profit values with a best fitted price-demand function to the data. Figure 1a also illustrates the differences in the profit function and the optimal price that are obtained after fitting a parametric model using linear, piecewise-linear, cubic polynomial, and exponential function form to the price-demand data.

We generated 100 demand data sets using ξ, each of which representing a possible collection of data. In practice only one such set is available. The mean relative fitting error (calcuated as the average over 100 scenarios of the ratio of the sum of absolute difference between the fitted function and true demand function divided by the value of the true demand function at each price point) for the quadratic, cubic, exponential and piecewise linear fit are 48.9%, 20.1%, 33.7%, and 8.7%, respectively. The profit evaluated using $h(\cdot)$ at the price recommended by optimizing the quadratic, cubic, exponential, and piecewise linear fitted functions (compared to the best possible profit obtained from the knowledge of exact demand function) has a mean relative error of 20.3%, 16.9%, 14.2%, and 2.4%, respectively. We now study the implications of model (FRP) to this decision problem. Assuming that the decision region of the sales price is in the interval $[1.5, 10.5]$, the (FRP) model is given as:

$$\max_{s\in[1.5,10.5]} \min_{d\in\mathfrak{D}_V} sd(s), \tag{6}$$

where the set $\mathfrak{D}_V$ is given by (3). A solution method to optimally solve (FRP) is discussed in Section 3.1. We now compare the quality of solutions produced by (6) with those obtained by the data fitting approach. Recall that the use of piecewise linear price-demand function has the smallest mean relative error (2.4%) in the above example. However, over the 100 scenarios the maximum relative error for this function is 20.6%, and the standard deviation in the error is 3.4%. Let us denote by $d^*(\cdot)$ the best fitted piecewise linear function in (2). We set $\epsilon = \kappa\text{Err}(d^*)$, and considered model (6) with $\kappa = 1.01$, 1.1, and 1.2, respectively. The corresponding mean relative errors in the profit over the 100 scenarios are 2.4%, 2.0%, and 2.0%. A small value of $\kappa = 1.01$ gives tight feasible region for the possible choices of a demand function, and as expected, shows a small difference. The maximum and the standard deviation of the relative error in this case are 20.5% and 3.1%. The value of $\kappa = 1.1$ moderately relaxes the feasible region, and shows a moderate reduction in the profit error variation. The maximum error and the standard deviation of the relative error in this case are 16.9% and 2.5%. Finally, for $\kappa = 1.2$, the maximum and the standard deviation of the relative error are 8.7% and 1.6%, showing a significant improvement in the worst case and the standard deviation of optimal pricing decision error. These results illustrate the importance of a function robust model in its ability to protect against large errors in decision making.

1.3 Literature Review

We now put our proposed approach in the context of exiting research on pricing and inventory decisions, and that in the area of robust optimization. The coordination of pricing and inventory decisions has been studied extensively. The classical literature on these problems is reviewed by Tayur et al. (1999), Petruzzi and Dada (1999), Yano and Gilbert (2004), and Chan et al. (2004). Past studies focus on stochastic modeling approaches using specially structured demand functions of the sales price and random market impact factors. The additive form for the demand function was first proposed by Mills (1959), the multiplicative form was proposed by Karlin and Carr (1962), and the additive-multiplicative form is used in Young (1978) and Chen and Simchi-Levi (2004). Raz and Porteus (2006) approximate a random demand function as a finite number of fractiles, each of which is assumed to be a piecewise linear function of the selling price. Stochastic optimization based approaches ignore the difficulty in exactly estimating the probability distribution of random market impact factors. A robust characterization was given by Scarf (1958) to first represent all valid distributions as an uncertainty set specified by the first and second moments, and subsequently considering the newsvendor problem as a maxmin problem over the set of distributions. This use of maximin approach in the literature is for the non-coordinating pricing problems, where the demand is assumed to be a random variable unrelated to the sales price. A detailed discussion on this modeling framework is given in Gallego and Moon (1993). Recently, Perakis and Roels (2008) and Zhu et al. (2013) discussed different characterizations of the newsvendor problem with regret objective using boundary and moment conditions. Chehrazi and Weber (2010) consider the robust pricing problem while specifying a shape-preserving set of demand functions using B-spline approximation. Kocabiyikoğlu and Popescu (2011) consider a generalized demand function relaxing assumptions on the functional form, however, their work is limited to the discussion of structural properties of a retailer's profit.

Over the years, parameter-based robust optimization has been studied and used in a wide variety of disciplines, such as control, energy, environment, finance, logistics, and statistics (Beyer and Sendhoff (2007), Bertsimas et al. (2011), Gabrel et al. (2013) and references therein). A major focus has been on the construction of the uncertainty set, reformulations, and computational complexity of robust optimization models (Bertsimas et al. (2011)). For robust linear optimization and least-squares, Ben-Tal and Nemirovski (1999) and El Ghaoui and Lebret (1997) consider uncertainty sets formulated by an ellipsoid. The corresponding problems are reformulated as linear, second-order cone, and semidefinite programming problems, respectively. Bertsimas and Sim (2004) and Bertsimas et al. (2004) define cardinality constrained polyhedral uncertainty set, and a norm-based uncertainty set. Ben-Tal, Nemirovski, and El Ghaoui, among others, extend studies to the robust counterparts of quadratic and semidefinite optimization (Ben-Tal and Nemirovski (1998); Ben-Tal et al. (2000); Ben-Tal and Nemirovski (2002); Bertsimas and Sim (2006)).

A currently active research area in robust optimization uses a set of distributions of random parameters. The idea of distributionally robust optimization originates in Scarf (1958) in the context of the newsvendor model. Dupačová (1987), Prékopa (1995), Bertsimas and Popescu (2005), Bertsimas et al. (2010), and Delage and Ye (2010) use linear or conic constraints to describe the set of distributions with moments. Shapiro and Ahmed (2004) define a probability ambiguity set with measure bounds and general moment constraints. Pflug and Wozabal (2007) consider the probability ambiguity set defined by Kantorovich distance from a reference probability measure. Bertsimas et al. (2010) use a piecewise linear utility with first and second moment equality constraints and showed that the corresponding problem has semidefinite programming reformulations. Delage and Ye (2010) give general conditions for polynomial time solvability of a distributionally robust model with constraints on first and second derivatives. Shapiro and Ahmed (2004) give stochastic programming reformulations of their model. Mehrotra and Zhang (2014) give reformulation of different distributionally robust models for the linear least-squares problem as semi-definite programs and second order cone programs.

To the best of our knowledge, the only works that study a functionally robust optimization problem where the uncertainty set is a function set are the research papers by Hu and Mehrotra (2015, 2012). The research in these papers is motivated from the fact that it is difficult to specify a decision maker's utility in a economic decision making setting. Hu and Mehrotra (2015, 2012) propose the use of a set of utilities to model this utility function ambiguity.

1.4 Contributions and Organization of this Paper

This paper makes the following contributions. First, it introduces a novel function-free decision approach to the problem of optimal product pricing and demand estimation. Next, in Section 2 it studies the properties of model (FRP), and subsequently discusses the relationship between the optimal pricing, order quantity, and profit under the proposed maximin model. Specifically, it shows how the model (FRP) exhibits the product pricing and demand principles for a conservative retailer. Section 3 develops solution methods for model (FRP). In the case where that demand functions are convex, it presents a novel cutting surface algorithm in Section 3.1 that generates functional cuts. We also show in this section that our algorithm terminates after a finite number of iterations ensuring desired accuracy. The case where demand functions are concave is studied in Section 3.2. In this case we reformulate (FRP) as a second-order cone program. Section 4 provides an illustration of the proposed model and algorithm using real-world data. Specifically, it considers a grocery store's pricing and order quantity decision using porterhouse beef price and demand data. Concluding remarks are given in Section 5.

2 Model and Solution Properties

We present properties of model (FRP), and its solution in this section. We first show that the sets $\mathfrak{D}_V$ and $\mathfrak{D}_\Lambda$ defined in (3) and (4), respectively, are convex. We then show that the demand functions in these sets are uniformly bounded and Lipschitz continuous. We next study the relationship between the optimal sales price and the inventory order quantity given by (FRP). The relationship suggests that the solution of (FRP) can be decomposed into two steps: in the first step the retailer decides the optimal sales price in relationship to the purchase price using the market price-demand data; and in the second step the optimal order quantity is obtained. Finally, we study the impact of the purchase price on the retailer's profit. We also show that in the (FRP) modeling framework an increase in the purchase price results in a reduction in the retailer's profit. Finally, we show that the marginal impact on retailer's profit decreases with an increase in the purchase price.

Let $\widehat{V}$(resp., $\widehat{\Lambda}$) denote the set of all piecewise linear positive, decreasing, convex (resp., concave) functions $d : \mathbb{R}_+ \to \mathbb{R}_+$, with break points t_i for $i = 1, \dots, n$. Note that $\widehat{V}$(resp., $\widehat{\Lambda}$) is a subset of V(resp., Λ). Let ϵ_{min} be the minimum value attainable in (2) over the set of piecewise linear functions $\widehat{V}$(resp., $\widehat{\Lambda}$), i.e.,

$$\epsilon_{min} := \min_{d \in \widehat{V}(\text{resp., } \widehat{\Lambda})} \text{Err}(d). \tag{7}$$

Proposition 2.1 *Let* $\mathfrak{D}_V$ *(*resp., $\mathfrak{D}_\Lambda$*) be defined in* (3) *(*resp., (4)*),* $\epsilon \geq \epsilon_{min}$*. Then, the set* $\mathfrak{D}_V$ *(*resp., $\mathfrak{D}_\Lambda$*) is nonempty and convex. Furthermore, if* $\epsilon < \epsilon_{min}$*, then the set* $\mathfrak{D}_V$ *(*resp., $\mathfrak{D}_\Lambda$*) is empty.*

Proof: By the definition of ϵ_{min}, the set $\mathfrak{D}_V$ is nonempty when $\epsilon \geq \epsilon_{min}$. Now assume that $d_1(\cdot), d_2(\cdot) \in \mathfrak{D}_V$. For any $\lambda \in (0,1)$, let $d_\lambda(\cdot) = \lambda d_1(\cdot) + (1-\lambda) d_2(\cdot)$. Then $d_\lambda(\cdot) \in V$(resp., Λ). By the Minkowski inequality, we have that

$$\mathrm{Err}(d_\lambda) \leq \lambda \mathrm{Err}(d_1) + (1-\lambda)\mathrm{Err}(d_2) \leq \epsilon.$$

It follows that $d_\lambda(\cdot) \in \mathfrak{D}_V$. Now assume that $\epsilon < \epsilon_{min}$, and the set $\mathfrak{D}$ is nonempty. Then, we have $d(\cdot) \in \mathfrak{D}$ such that $d(\cdot) \in V$. Let us consider the piecewise linear function

$$\tilde{d}(s) := \begin{cases} \frac{d(t_2)-d(t_1)}{t_2-t_1} s + \frac{t_2 d(t_1) - t_1 d(t_2)}{t_2 - t_1} & 0 \leq s < t_2, \\ \frac{d(t_{i+1})-d(t_i)}{t_{i+1}-t_i} s + \frac{t_{i+1} d(t_i) - t_i d(t_{i+1})}{t_{i+1} - t_i} & t_i \leq s < t_{i+1},\ i = 2, \ldots, n-1, \\ \frac{d(t_n)-d(t_{n-1})}{t_n-t_{n-1}} s + \frac{t_n d(t_{n-1}) - t_{n-1} d(t_n)}{t_n - t_{n-1}} & s \geq t_{n-1}. \end{cases}$$

Clearly, $\tilde{d}(\cdot) \in \widehat{V}$. Now note that the definition $\mathrm{Err}(\cdot)$ only uses the evaluation of a function at points $t_1, \ldots, t_n$. Since by construction $\tilde{d}(s) = d(s)$ at $t_1, \ldots, t_n$, $\mathrm{Err}(\tilde{d}) \leq \epsilon$. This contradicts the definition of ϵ_{min}. The stated properties of the set $\mathfrak{D}_\Lambda$ can be proved similarly. □

The following proposition shows uniform boundedness and Lipschitz continuity of the demand functions in the sets $\mathfrak{D}_V$ and $\mathfrak{D}_\Lambda$.

Proposition 2.2 *Any function* $d(s) \in \mathfrak{D}_V$ *(resp.,* $\mathfrak{D}_\Lambda$*) is uniformly bounded as follows:*

$$d(s) \leq \frac{1}{m_1} \sum_{j=1}^{m_1} y_{1j} + \epsilon \left(\frac{\sum_{i=1}^{n} m_i}{m_1} \right)^{1/2} =: B, \quad \textit{for all } s \in [t_1, t_n],\ d \in \mathfrak{D}_V \ (\text{resp. } \mathfrak{D}_\Lambda).$$

Furthermore, any function $d(\cdot) \in \mathfrak{D}_V$ *is Lipschitz continuous in* $\mathfrak{S}$ *with the Lipschitz constant*

$$L := \frac{B}{t_2 - t_1}, \tag{8}$$

and any $d(s) \in \mathfrak{D}_\Lambda$ is Lipschitz continuous in $\mathfrak{S}$ with the Lipschitz constant

$$L := \frac{B}{t_n - t_{n-1}}. \tag{9}$$

Proof: Since $d(\cdot)$ is decreasing, showing $d(t_1) \leq B$ implies that $d(s) \leq B$, for any $s \in [t_1, t_n]$. Let $d(\cdot) \in \mathfrak{D}_V$, and note that $d(t_j) \geq 0$. Now, using Holder's inequality, and (3) we have

$$\left(\sum_{j=1}^{m_1}(y_{1j} - d(t_1))\right)^2 \leq m_1 \sum_{j=1}^{m_1}(y_{1j} - d(t_1))^2 \leq m_1 \sum_{i=1}^{n}\sum_{j=1}^{m_1}(y_{ij} - d(t_i))^2 \leq m_1\epsilon^2 \sum_{i=1}^{n} m_i,$$

which by taking the square root of both sides implies $d(t_1) \leq B$. It also follows by decreasingness of $d(\cdot)$ and its convexity that, for any $t', t'' \in \mathfrak{S}$ (note that $\mathfrak{S} \subseteq [t_2, t_{n-1}]$), $t'' > t'$,

$$d(t') \leq \frac{t'' - t'}{t'' - t_2} d(t_2) + \frac{t' - t_2}{t'' - t_2} d(t''),$$

and

$$d(t_2) \leq \frac{t' - t_2}{t' - t_1} d(t_1) + \frac{t_2 - t_1}{t' - t_1} d(t').$$

The above two inequalities lead to

$$\frac{d(t') - d(t'')}{t'' - t'} \leq \frac{d(t_1) - d(t_2)}{t_2 - t_1} = L,$$

which ensures Lipschitz continuity of $d(s)$. The proof for $d(s) \in \mathfrak{D}_\Lambda$ is similar to that for the $\mathfrak{D}_V$ case. □

We now discuss the relationship between the optimal selling price and order quantity determined by (FRP). By this relation, we solve model (FRP) in two steps to decide optimal selling price and order quantity separately.

Theorem 2.3

(1). Let us consider (FRP)*, with $\mathfrak{D} := \mathfrak{D}_V$ (resp., $\mathfrak{D} := \mathfrak{D}_\Lambda$), and (s^*, x^*) be its optimal solution. Then,*

$$x^* = \min_{d \in \mathfrak{D}_V} d(s^*) \quad (\textit{resp.},\ x^* = \min_{d \in \mathfrak{D}_\Lambda} d(s^*)). \tag{10}$$

(2). The optimal value of model (FRP) *equals to the optimal value of the problem*

$$\max_{s\in\mathfrak{S}} \min_{d\in\mathfrak{D}_V} (s-p)d(s) \quad \left(resp.,\ \max_{s\in\mathfrak{S}} \min_{d\in\mathfrak{D}_\Lambda} (s-p)d(s)\right). \tag{FRP-s}$$

Let s^* *be an optimal solution of* (FRP-s)*, and let*

$$x^* = \min_{d\in\mathfrak{D}_V} d(s^*) \quad \left(resp.,\ x^* = \min_{d\in\mathfrak{D}_\Lambda} d(s^*)\right). \tag{FRP-x}$$

Then, (s^*, x^*) *is an optimal solution of* (FRP)*.*

Proof: (1). Suppose that $x^* \neq \min_{d\in\mathfrak{D}_V} d(s^*)$, and $x^* > \min_{d\in\mathfrak{D}_V} d(s^*)$. The optimal value of model (FRP) is given by $s^* \min_{d\in\mathfrak{D}_V} d(s^*) - px^* < s^* \min_{d\in\mathfrak{D}_V} d(s^*) - p \min_{d\in\mathfrak{D}_V} d(s^*)$, a contradiction. On the other hand, if $x^* < \min_{d\in\mathfrak{D}_V} d(s^*)$, the optimal value of model (FRP) equals $(s^* - p)x^*$. The assumption that $p < \underline{s}$, $d(s) > 0$, and $x \geq 0$ implies that optimal profit is positive, and hence, $(s^* - p) > 0$. Now, $(s^* - p)x^* < (s^* - p) \min_{d\in\mathfrak{D}_V} d(s^*)$.

(2). Let s^* be the optimal solution of (FRP-s), and x^* be given by (FRP-x). Since (s^*, x^*) gives a feasible solution of model (FRP), the objective value of (FRP) at (s^*, x^*), which is the optimal value of (FRP-s) (i.e., $\min_{d\in\mathfrak{D}_V}(s^* - p)d(s^*)$) is below the optimal value of model (FRP). On the other hand, let $(\hat{s}, \hat{x})$ be an optimal solution of (FRP). Then, $\hat{x} = \min_{d\in\mathfrak{D}} d(\hat{s})$. The optimal value of (FRP) is thus given by $\min_{d\in\mathfrak{D}}(\hat{s} - p)d(\hat{s})$. Since $\hat{s}$ is a feasible solution of (FRP-s), the optimal value of (FRP) is bounded above by the optimal value of (FRP-s). Hence, (s^*, x^*) is an optimal solution of model (FRP), and models (FRP) and (FRP-s) are equivalent. The proof for the case where $\mathfrak{D} := \mathfrak{D}_\Lambda$ is similar. □

We now study the influence of purchase price p on retailer's profit. We show that the retailer profit is a decreasing convex function of the purchase price. Hence, the marginal impact on profit reduces with increase in the purchase price.

Theorem 2.4 *Let* $p \in [0, \overline{s})$*,* $\epsilon \geq \epsilon_{min}$*, and* $r^*(p, \epsilon)$ *be the optimal value of model* (FRP)*, with* $\mathfrak{D} := \mathfrak{D}_V$ *(resp.,* $\mathfrak{D} := \mathfrak{D}_\Lambda$*). Let* $X^*(p, \epsilon)$ *be the set of optimal order quantities given by model* (FRP)*. Then,*

(1). $r^*(p, \cdot)$ *is decreasing and convex in* $[\epsilon_{min}, \infty)$*.*

(2). $r^*(\cdot,\epsilon)$ *is decreasing and convex in* $[0,\overline{s})$.

(3). $r^*(\cdot,\epsilon)$ *is locally Lipschitz continuous, directionally differentiable in* $[0,\overline{s})$, *and its right and left derivatives are given by*

$$r^{*\prime}_{p+}(p,\epsilon) := \lim_{\Delta p\to 0}\frac{r^*(p+\Delta p,\epsilon)-r^*(p,\epsilon)}{\Delta p} = -\min\ \{x : x\in X^*(p,\epsilon)\},$$
$$r^{*\prime}_{p-}(p,\epsilon)(p,\epsilon) ::= \lim_{\Delta p\to 0}\frac{r^*(p-\Delta p,\epsilon)-r^*(p,\epsilon)}{\Delta p} = \max\ \{x : x\in X^*(p,\epsilon)\},$$

i.e., the set of subgradients of $r^*(p,\epsilon)$ *with respect to* p *is given as follows:*

$$\partial r^*(p,\epsilon) = [-\min\ \{x : x\in X^*(p,\epsilon)\}, \max\ \{x : x\in X^*(p,\epsilon)\}].$$

In particular, if the set $X^*(p,\epsilon)=\{x^*(p,\epsilon)\}$ *is a singleton, then* $r^*(\cdot,\epsilon)$ *is differentiable at* p, *and its derivative is given by*

$$r^{*\prime}_{p}(p,\epsilon) = -x^*(p,\epsilon).$$

Proof: (1). In the proof we denote by $\mathfrak{D}_V(\epsilon)$ the set of demand functions defined by (3) for a given ϵ. It follows by Theorem 2.3 that model (FRP) is equivalent to model (FRP-s), which can be further written as

$$r^*(p,\epsilon) = \max_{s\in\mathfrak{S}}\left[R(s,p,\epsilon) := (s-p)\min_{d\in\mathfrak{D}_V(\epsilon)} d(s)\right].$$

Since $\mathfrak{D}_V(\epsilon_1)\subseteq\mathfrak{D}_V(\epsilon_2)$ for $\epsilon_{min}\le\epsilon_1<\epsilon_2$, $r^*(p,\cdot)$ is decreasing in $[\epsilon_{min},\infty)$. Let $\lambda\in[0,1]$. For sets $\mathfrak{A}$ and $\mathfrak{B}$, let $\lambda\mathfrak{A} := \{\lambda a : a\in\mathfrak{A}\}$ and $\mathfrak{A}\oplus\mathfrak{B} := \{a+b : a\in\mathfrak{A}, b\in\mathfrak{B}\}$. For any $d_1\in\mathfrak{D}_V(\epsilon_1)$ and $d_2\in\mathfrak{D}_V(\epsilon_2)$, we have $\lambda d_1+(1-\lambda)d_2\in\mathfrak{D}_V(\lambda\epsilon_1+(1-\lambda)\epsilon_2)$. It implies that

$$\lambda\mathfrak{D}_V(\epsilon_1)\oplus(1-\lambda)\mathfrak{D}_V(\epsilon_2)\subseteq\mathfrak{D}_V(\lambda\epsilon_1+(1-\lambda)\epsilon_2).$$

Since

$$\begin{aligned} R(s,p,\lambda\epsilon_1+(1-\lambda)\epsilon_2) &= (s-p)\min_{d\in\mathfrak{D}_V(\lambda\epsilon_1+(1-\lambda)\epsilon_2)} d(s) \\ &\le (s-p)\min_{d\in\lambda\mathfrak{D}_V(\epsilon_1)\oplus(1-\lambda)\mathfrak{D}_V(\epsilon_2)} d(s) \\ &= (s-p)\left(\min_{d_1\in\mathfrak{D}_V(\epsilon_1)} \lambda d_1(s) + \min_{d_2\in\mathfrak{D}_V(\epsilon_2)} (1-\lambda)d_2(s)\right) \\ &= \lambda R(s,p,\epsilon_1) + (1-\lambda)R(s,p,\epsilon_2), \end{aligned}$$

it follows that $r^*(p,\cdot)$ is convex in $[\epsilon_{min},\infty)$. Note that in the above the second equality results from the independence of the sets $\mathfrak{D}_V(\epsilon_1)$ and $\mathfrak{D}_V(\epsilon_2)$.

(2) By the fact that $R(s,p,\epsilon)$ is a decreasing linear function of p, we have that $r^*(\cdot,\epsilon)$ is decreasing and convex in $[0,\overline{s})$.

(3). We now show the directional differentiability. Note that the purchase price $p \in [0,\overline{s}]$, which is a compact region. $R(s,\cdot,\epsilon)$ is linear for every $s \in \mathfrak{S}$, and its partial derivate for the purchase price p is

$$R'_p(s,p,\epsilon) = -\min_{d\in\mathfrak{D}_V(\epsilon)} d(s).$$

For any given $\tau > 0$, let $\delta = \tau/L$ where L is given by (8). For any $(s_1,p_1),(s_2,p_2) \in \mathfrak{S}\times[0,\overline{s}]$ satisfying $\|(s_1,p_1)-(s_2,p_2)\|_2 \le \delta$, we have that

$$|R'_p(s_1,p_1,\epsilon) - R'_p(s_2,p_2,\epsilon)| = \left|\min_{d\in\mathfrak{D}_V(\epsilon)} d(s_1) - \min_{d\in\mathfrak{D}_V(\epsilon)} d(s_2)\right|.$$

Let $d_1^*(\cdot)$ and $d_2^*(\cdot)$ be the optimal solutions of $\min_{d\in\mathfrak{D}_V(\epsilon)} d(s_1)$ and $\min_{d\in\mathfrak{D}_V(\epsilon)} d(s_2)$, respectively. Without loss of generality, we assume $d_1^*(s_1) \ge d_2^*(s_2)$. It follows that

$$|R'_p(s_1,p_1,\epsilon) - R'_p(s_2,p_2,\epsilon)| = d_1^*(s_1) - d_2^*(s_2) \le d_2^*(s_1) - d_2^*(s_2) \le \max_{d\in\mathfrak{D}_V(\epsilon)} |d(s_1)-d(s_2)|,$$

which further implies that, by Proposition 2.2,

$$|R'_p(s_1,p_1,\epsilon) - R'_p(s_2,p_2,\epsilon)| \le \max_{d\in\mathfrak{D}_V(\epsilon)} |d(s_1)-d(s_2)| \le L|s_1-s_2| \le L\delta = \tau.$$

It shows that $R'_p(\cdot,\cdot,\epsilon)$ is continuous in $\mathfrak{S}\times[0,\overline{s}]$. Let $S^*(p,\epsilon)$ be the set of the optimal solutions of model (FRP-s), and then by Theorem 2.3 we have

$$X^*(p,\epsilon) = \left\{ x \in \mathbb{R}_+ : x = \min_{d \in \mathfrak{D}(\epsilon)} d(s^*), \text{ for all } s^* \in S^*(p,\epsilon) \right\}.$$

It follows from Danskin's theorem (stated in Theorem 7.21 of Shapiro et al. (2009)) that

$$r^{*\prime}_{p+}(p,\epsilon) = \max_{s \in S^*(p,\epsilon)} R'_p(s,p,\epsilon) = -\min_{s \in S^*(p,\epsilon)} \min_{d \in \mathfrak{D}_V(\epsilon)} d(s) = -\min_{x \in X^*(p,\epsilon)} x,$$
$$r^{*\prime}_{p-}(p,\epsilon) = \max_{s \in S^*(p,\epsilon)} -R'_p(s,p,\epsilon) = \max_{s \in S^*(p,\epsilon)} \min_{d \in \mathfrak{D}_V(\epsilon)} d(s) = \max_{x \in X^*(p,\epsilon)} x.$$

If $X^*(p,\epsilon) = \{x^*(p,\epsilon)\}$ is a singleton, then $r^{*\prime}_{p+}(p,\epsilon) = -x(p,\epsilon)*$, and $r^{*\prime}_{p-}(p,\epsilon) = x^*(p,\epsilon)$. Hence, we have $r^{*\prime}_{p}(p,\epsilon) = -x^*(p,\epsilon)$.

The proof for the case where $\mathfrak{D} := \mathfrak{D}_\Lambda$ is similar. □

Theorem 2.4 indicates a conservative retailer's behavior. It indicates that the maximum profit from the model is a decreasing convex function of the demand function modeling error ϵ. In addition, with ϵ, the optimal order quantity decreases by Theorem 2.3. This reflects the fact that a conservative retailer will reduce the order quantities to avoid losses resulting from an overestimation of demand. On the other hand, an increase in the purchase price incurs a reduction in the maximum profit, and the optimal order quantity. This is also consistent with the expected behavior of a conservative retailer: a high purchase price makes the retailer reduce the order quantity.

3 Solution Methods

Model (FRP) can be solved in two steps to decide optimal selling price and order quantity separately. Model (FRP-s)in Theorem 2.3 specifies the retailer's optimal selling price and maximum profit. In this section we develop solution methods for model (FRP-s). The two subsections consider the cases when the demand functions are convex and concave, respectively.

3.1 The Model with Convex Demand Functions

We now develop an algorithm for model (FRP-s) when demand functions are required to be convex. First consider the inner minimization problem of model (FRP-s) to seek the worst-case demand function at a given selling price. The following theorem reformulates the inner minimization problem of model (FRP-s) as a quadratically constrained linear program.

Theorem 3.1 *For a given $s \in \mathfrak{S}$, suppose that $t_k \leq s < t_{k+1}$ for some $k \in 2, \ldots, n-2$. Then the inner minimization problem of model* (FRP-s),

$$\min_{d \in \mathfrak{D}_V} d(s), \tag{11}$$

is equivalent to

$$\begin{aligned}
&\min_{u,z} \; z && \text{(12a)}\\
&s.t. \; \frac{1}{\sum_{i=1}^{n} m_i} \sum_{i=1}^{n} \sum_{j=1}^{m_i} (y_{ij} - u_i)^2 \leq \epsilon^2 && \text{(12b)}\\
&\quad \frac{u_{i+2} - u_{i+1}}{t_{i+2} - t_{i+1}} \geq \frac{u_{i+1} - u_i}{t_{i+1} - t_i}, \quad i = 1, \ldots, n-2 && \text{(12c)}\\
&\quad u_{n-1} \geq u_n && \text{(12d)}\\
&\quad z \geq \frac{u_k - u_{k-1}}{t_k - t_{k-1}}(s - t_{k-1}) + u_{k-1} && \text{(12e)}\\
&\quad z \geq \frac{u_{k+2} - u_{k+1}}{t_{k+2} - t_{k+1}}(s - t_{k+1}) + u_{k+1} && \text{(12f)}\\
&\quad u_i \geq 0, \quad i = 1, \ldots, n. && \text{(12g)}
\end{aligned}$$

Denote by $(u_1^*, \ldots, u_n^*, \widehat{z}^*)$ an optimal solution of problem (12)*. Then the piecewise linear function*

$$d_s^*(t) := \begin{cases}
\frac{u_2^* - u_1^*}{t_2 - t_1}(t - t_1) + u_1^* & 0 \leq t < t_1,\\
\frac{u_{i+1}^* - u_i^*}{t_{i+1} - t_i}(t - t_i) + u_i^* & t_i \leq t < t_{i+1}, \; i = 1, \ldots, k,\\
\frac{\widehat{z}^* - u_k^*}{s - t_k}(t - t_k) + u_k^* & t_k \leq t < s,\\
\frac{u_{k+1}^* - \widehat{z}^*}{t_{k+1} - s}(t - s) + \widehat{z}^* & s \leq t \leq t_{k+1},\\
\frac{u_{i+1}^* - u_i^*}{t_{i+1} - t_i}(t - t_i) + u_i^* & t_i \leq t < t_{i+1}, \; i = k+1, \ldots, n-1,\\
\max\{0, \; \frac{u_n^* - u_{n-1}^*}{t_n - t_{n-1}}(t - t_{n-1}) + u_{n-1}^*\} & t \geq t_n,
\end{cases} \tag{13}$$

is an optimal solution of problem (11)*.*

Proof: Note that the feasible region $\mathfrak{S}$ is nonempty. Let $\tilde{d}(\cdot)$ be an optimal solution of (11). Let $\tilde{u}_i := \tilde{d}(t_i)$, and $\tilde{z} := \tilde{d}(s)$. Since $\tilde{d}(\cdot)$ is a positive, decreasing, and convex function, $(\tilde{u}_1, \ldots, \tilde{u}_n, \tilde{z})$

satisfies conditions (12c) - (12g). $(\tilde{u}_1, \ldots, \tilde{u}_n, \tilde{z})$ also satisfies condition (12b), because $Err(\tilde{d}) \leq \epsilon$. It thus follows that $(\tilde{u}_1, \ldots, \tilde{u}_n, \tilde{z})$ is feasible to (12). Hence, the optimal value $\tilde{z}$ of (11) is no greater than the optimal value of (12).

On the other hand, the feasibility of (12) shown above implies the existence of $(u_1^*, \ldots, u_n^*, \widehat{z}^*)$. It is easy to see that the function $d_s^*(\cdot)$ is positive, continuous, and convex in $\mathbb{R}_+$, and also satisfies $Err(d_s^*) \leq \epsilon$. It follows that $d_s^*(\cdot) \in \mathfrak{D}_V$. The optimal value $\widehat{z}^*$ of (12) is below the optimal value $\tilde{z}$ of (11). Hence, (11) and (12) have the same optimal value, and $d_s^*(\cdot)$ is a worst-case demand function in (11) which reaches the lowest demand at the given price s. □

We next give an algorithm to solve model (FRP-s), which is motived from the cutting surface algorithm of Mehortra and Papp (2014). This algorithm uses a sequence of functional cuts, where each cut is given by a function in $\mathfrak{D}_V$. We solve a sequence of relaxations of model (FRP-s) over a subset of $\mathfrak{D}_V$ as

$$\begin{aligned} \max_{r,s}\ & r \\ \text{s.t. } & r \leq (s-p)d(s), \quad d(\cdot) \in \{d_1(\cdot), \cdots, d_\ell(\cdot)\} \subset \mathfrak{D}_V, \\ & s \in \mathfrak{S}. \end{aligned} \tag{14}$$

The feasible set in (14) is not convex. In order to identify a new cut, we construct a lower envelope function of $d_1(\cdot), \cdots, d_\ell(\cdot)$ on $\mathfrak{S}$ as

$$\widehat{d}(s) := \min\{d_1(s), \cdots, d_\ell(s)\}, \quad \text{for all } s \in \mathfrak{S}.$$

Using $\widehat{d}(\cdot)$, problem (14) is now reformulated as

$$\widehat{r}^* := \max_{s \in \mathfrak{S}} (s-p)\widehat{d}(s). \tag{15}$$

At an optimal solution $\widehat{s}^*$ of (15), we consider (11). Let $\widehat{z}^*$ be the optimal value of (11), and $d_{\widehat{s}^*}^*(\cdot)$ be an optimal solution given by (13) by solving the equivalent problem (12). If $\widehat{r}^* \geq (\widehat{s}^* - p)\widehat{z}^*$, we have a solution of (FRP-s); otherwise, corresponding to the solution $d_{\widehat{s}^*}^*(\cdot)$ of (11), the constraint $r \leq (s-p)d_{\widehat{s}^*}^*(s)$ is added to the master problem (14) as a valid functional cut. The addition of this cut updates the lower envelop function by

$$\widehat{d}(s) = \min\{\widehat{d}(s), d^*_{\widehat{s}^*}(s)\}, \quad \text{for all } s \in \mathfrak{S},$$

and (15) is solved again. Algorithm 1 formally describes this cutting surface algorithm. We will show in Theorem 3.2 that, in a finite number of iterations, Algorithm 1 correctly generate a δ-optimal solution. A δ-optimal solution of an optimization problem is a feasible solution at which the difference between the corresponding objective value and true optimal value is less than δ.

Algorithm 1 A Cutting Surface Algorithm for (FRP)

---

0. Choose $\delta > 0$. Let $\widehat{r}^* = \infty$, $\widehat{s}^* = (\max\{p, \underline{s}\} + \overline{s})/2$, and the lower envelope function $\widehat{d}(s) = \infty$ for all $s \in \mathfrak{S}$.

1. Solve (12) for the given price $\widehat{s}^*$. Let $\widehat{z}^*$ be the optimal value and $d^*_{\widehat{s}^*}(\cdot)$ be the worst-case demand function given by (13).

2. If $\widehat{r}^* \geq (\widehat{s}^* - p)\widehat{z}^* + \delta$, go to step 3; otherwise, exit.

3. Update the lower envelope function $\widehat{d}(s) = \min\{\widehat{d}(s), d^*_{\widehat{s}^*}(s)\}$ for all $s \in \mathfrak{S}$, and solve (15). Set $\widehat{r}^*$ to be the optimal value of (15) and $\widehat{s}^*$ to be an optimal solution. Then go to step 1.

At step 3 in Algorithm 1, (15) is solved for function $\widehat{d}(\cdot)$. Problem (15) is not a convex program when the objective profit function $(s - p)\widehat{d}(s)$ is not concave in $s \in \mathfrak{S}$. A method to find the optimal value and solution of (15) will be given in Theorem 3.5 after we discuss the termination of Algorithm 1 in the following theorem.

Theorem 3.2 *Algorithm 1 terminates with a δ-optimal solution in finitely many iterations.*

Proof: Let $d^*_j(\cdot)$ $(j = 1, \ldots, \ell)$ be the worst demand functions generated by Algorithm 1 at the first $\ell(< \infty)$ iterations. The current envelop function is

$$\widehat{d}(s) = \min\{d^*_1(s), \ldots, d^*_\ell(s)\} \quad \text{for all } s \in \mathfrak{S}.$$

Let

$$\hat{S} := \int_{\mathfrak{S}} \widehat{d}(s) ds$$

be the area generated by $\widehat{d}(\cdot)$ over $\mathfrak{S}$. It can be seen that $\hat{S}$ is finite since Proposition 2.2 guarantees

$$\hat{S} \leq B(\overline{s} - \underline{s}) < \infty.$$

We first claim the Lipschitz continuity of $\widehat{d}(\cdot)$ in the following lemma.

Lemma 3.3 *The lower envelop function $\widehat{d}(\cdot)$ generated at the ℓth iteration of the algorithm is Lipschitz continuous in $\mathfrak{S}$ with Lipschitz constant L given by* (8)*.*

Proof: From Theorem 3.1, each $d_j^*(\cdot)$, for $j = 1, \ldots, \ell$, is a decreasing piecewise linear function with at most $(n-1)$ break points. It implies that $\widehat{d}(\cdot)$ is a decreasing and piecewise linear function with at most $(n-1)\ell$ break points. We arbitrarily choose $s_1, s_2 \in \mathfrak{S}$ with $s_1 < s_2$. Suppose that b_i for $i = 1, \ldots, k$ are all the break points of $\widehat{d}(\cdot)$ in (s_1, s_2) and $s_1 < b_1 < \cdots < b_k < s_2$. Let $b_0 := s_1$ and $b_{k+1} := s_2$. There is thus some $j \in \{1, \ldots, \ell\}$ such that $\widehat{d}(s) = d_j^*(s)$ for all $s \in [b_i, b_{i+1}]$ and $i = 0, \ldots, k$ such that, by Proposition 2.2,

$$\widehat{d}(b_i) - \widehat{d}(b_{i+1}) = d_j^*(b_i) - d_j^*(b_{i+1}) \leq L(b_{i+1} - b_i).$$

It follows that

$$\widehat{d}(s_1) - \widehat{d}(s_2) = \sum_{i=0}^{k} \widehat{d}(b_i) - \widehat{d}(b_{i+1}) \leq \sum_{i=0}^{k} L(b_{i+1} - b_i) = L \sum_{i=0}^{k} b_{i+1} - b_i = L(s_2 - s_1).$$

□

We next show that Algorithm 1 terminates in finitely many iterations. Let $d_{\ell+1}^*(\cdot)$ be the worst demand function for current $\widehat{s}^*$ generated at step 2 of the $(\ell+1)$th iteration, and suppose that Algorithm 1 does not exit at step 3. Since $\widehat{r}^* = (\widehat{s}^* - p)\widehat{d}(\widehat{s}^*)$, and $\widehat{z}^* = d_{l+1}^*(\widehat{s}^*)$, it implies that

$$(\widehat{s}^* - p)\widehat{d}(\widehat{s}^*) \geq (\widehat{s}^* - p)d_{l+1}^*(\widehat{s}^*) + \delta.$$

The assumption that $p < \overline{s}$ implies that $\widehat{s}^* > p$. Hence, we have

$$\widehat{d}(\widehat{s}^*) - d_{l+1}^*(\widehat{s}^*) \geq \frac{\delta}{\widehat{s}^* - p} \geq \frac{\delta}{\overline{s} - p} =: \gamma > 0.$$

Let

$$s' := \max\left\{\underline{s},\ \widehat{s}^* - \frac{\gamma}{4L}\right\}, \quad s'' := \min\left\{\overline{s},\ \widehat{s}^* + \frac{\gamma}{4L}\right\}.$$

Because of the Lipschitz continuity of $d^*_{\ell+1}(\cdot)$ and $\widehat{d}(\cdot)$ shown by Proposition 2.2 and Lemma 3.3, we have that, for any $s \in [s', s'']$,

$$\begin{aligned}
|\widehat{d}(\widehat{s}^*) - d^*_{l+1}(\widehat{s}^*)| \le & |\widehat{d}(s) - \widehat{d}(\widehat{s}^*)| + |\widehat{d}(s) - d^*_{l+1}(s)| + |d^*_{l+1}(\widehat{s}^*) - d^*_{l+1}(s)| \\
\le & 2L|s - \widehat{s}^*| + |\widehat{d}(s) - d^*_{l+1}(s)| \\
\le & \frac{\gamma}{2} + |\widehat{d}(s) - d^*_{l+1}(s)|,
\end{aligned}$$

and hence,

$$|\widehat{d}(s) - d^*_{l+1}(s)| \ge \frac{\gamma}{2}.$$

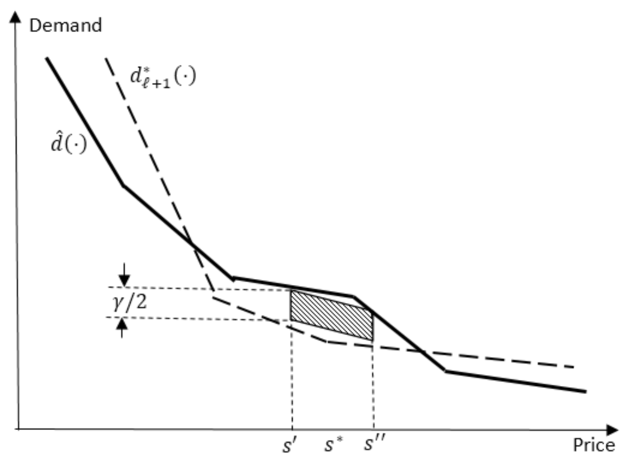


Figure 2: Generation of the lower envelop function

At step 4 we update the lower envelop function given as

$$\hat{\hat{d}}(s) = \min\{\widehat{d}(s), d^*_{\ell+1}(s)\}, \quad \text{for all } s \in \mathfrak{S}.$$

Let

$$\hat{\hat{S}} := \int_{\mathfrak{S}} \hat{\hat{d}}(s) ds,$$

be the area generated by $\hat{\hat{d}}(\cdot)$ over $\mathfrak{S}$. Between $\widehat{d}(\cdot)$ and $\hat{\hat{d}}(\cdot)$ there must be a parallelogram, shown in Figure 2 as the shaded region, of which the area is at least $\frac{\gamma}{2} \times \frac{\gamma}{4L} = \frac{\gamma^2}{8L}$. It follows that

$$\hat{S} - \hat{\hat{S}} \geq \frac{\gamma^2}{8L}.$$

This means that the update of the lower envelop function at each iteration results in the reduction of the generated area $\hat{S}$ by at least $\frac{\gamma^2}{8L}$. Therefore, the algorithm must stop in at most $\frac{8LB(\bar{s}-\underline{s})}{\gamma^2}$ iterations.

Finally, we show that the solution generated when Algorithm 1 stops is a δ-optimal solution of (FRP). Let r^* be the optimal value of (FRP), and thus it is also the optimal value of (FRP-s) as indicated by Theorem (2.3). Let $\widehat{r}^*_{last}$, $\widehat{s}^*_{last}$, and $\widehat{z}^*_{last}$ to be the last $\widehat{r}^*$, $\widehat{s}^*$, and $\widehat{z}^*$, respectively, generated when Algorithm 1 stops at the ℓ_{last}-th iteration. Also let $\widehat{\mathfrak{D}}_V := \{d^*_1(\cdot), \ldots, d^*_{\ell_{last}}(\cdot)\}$ be the set of all the generated cuts in the algorithm. The stopping criterion at step 2 ensures that

$$\widehat{r}^*_{last} < (\widehat{s}^*_{last} - p)d(\widehat{z}^*_{last}) + \delta = \min_{d \in \mathfrak{D}_V} (\widehat{s}^*_{last} - p)d(\widehat{s}^*_{last}) + \delta \leq r^* + \delta.$$

On the other hand, since $\widehat{\mathfrak{D}}_V \subseteq \mathfrak{D}_V$, we have

$$\widehat{r}^*_{last} = \max_{s \in \mathfrak{S}} \min_{d \in \widehat{\mathfrak{D}}_V} (s - p)d(s) \geq r^*.$$

□

Algorithm 1 stops in finitely many iterations. This means that the total number of cuts generated by Algorithm 1 is finite. From Theorem 3.1, these cuts are decreasing piecewise linear functions with at most $n-1$ break points given by $\widehat{s}^*$ and $t_2, \ldots, t_{n-1}$. Therefore, as stated in the following corollary, the lower envelop functions $\widehat{d}(\cdot)$ generated at each iteration are decreasing piecewise linear functions with finitely many break points.

Corollary 3.4 *The lower envelop function $\widehat{d}(\cdot)$ at each iteration in Algorithm 1 is decreasing and piecewise linear. The number of the break points of $\widehat{d}(\cdot)$ is finite at each iteration till the termination of Algorithm 1.*

Step 3 in Algorithm 1 needs to solve a non-convex program where the objective profit function is not concave. Corollary 3.4 guarantees that, when running Algorithm 1, the generated lower envelop function $\widehat{d}(\cdot)$ is piecewise linear, and has finitely many break points. Based on the piecewise linearity of $\widehat{d}(\cdot)$, the following theorem gives an approach to solve (15).

Theorem 3.5 *Let $\underline{s} = b_1 < \cdots < b_m = \overline{s}$ be the break points of the piecewise linear lower envelop function $\widehat{d}(\cdot)$ in $\mathfrak{S}$. For $i = 1, \ldots, m-1$, let*

$$\begin{aligned} h_i &:= \frac{\widehat{d}(b_i) - \widehat{d}(b_{i+1})}{b_{i+1} - b_i}, \\ q_i &:= \frac{b_{i+1}\widehat{d}(b_i) - b_i\widehat{d}(b_{i+1})}{b_{i+1} - b_i}, \\ s_i &:= \begin{cases} b_i, & \text{if } \frac{ph_i + q_i}{2h_i} \le b_i, \\ \frac{ph_i + q_i}{2h_i}, & \text{if } b_i < \frac{ph_i + q_i}{2h_i} \le b_{i+1}, \\ b_{i+1}, & \text{otherwise}, \end{cases} \\ r_i &:= \begin{cases} -h_i b_i^2 + (ph_i + q_i)b_i - pq_i, & \text{if } \frac{ph_i + q_i}{2h_i} \le b_i, \\ \frac{(ph_i - q_i)^2}{4h_i}, & \text{if } b_i < \frac{ph_i + q_i}{2h_i} \le b_{i+1}, \\ -h_i b_{i+1}^2 + (ph_i + q_i)b_{i+1} - pq_i, & \text{otherwise.} \end{cases} \end{aligned}$$

The optimal value of (15) *is*

$$\widehat{r}^* = \max_{i=1,\ldots,m-1} r_i,$$

of which a maximizer is denoted by i^*. An optimal solution of (15) *is*

$$\widehat{s}^* = s_{i^*}.$$

Proof: We partition the decision region $\mathfrak{S}$ into the subintervals $[b_i, b_{i+1}]$, for $i = 1, \ldots, m$, on each of which $\widehat{d}(\cdot)$ is a linear function written as

$$\widehat{d}(s) = -h_i s + q_i.$$

The maximization of the profit function over each subinterval is thus an one-dimensional quadratic programming problem given by

$$\max_{s\in[b_i,b_{i+1}]} -h_i s^2 + (ph_i + q_i)s - pq_i \;=\; \max_{s\in[b_i,b_{i+1}]} -h_i \left(s - \frac{ph_i + q_i}{2h_i}\right)^2 + \frac{(ph_i - q_i)^2}{4h_i}.$$

Hence, the optimal solution and value of each subproblem are s_i and r_i. The optimal solution s_i is a local optimum of (15). It follows by Corollary 3.4 that $m < \infty$ at any iteration of Algorithm 1. Hence, the number of the local optimums is finite. We search the maximum of the all local optimums to get the global optimal value and solution of (15) as $\widehat{r}^*$ and $\widehat{s}^*$. $\square$

3.2 The Model with Concave Demand Functions

We now consider the case that demand functions are decreasing and concave in selling price given by $\mathfrak{D}_\Lambda$. The following theorem provides a second-order cone programming formulation of model (FRP-s) in this case.

Theorem 3.6 *Consider model* (FRP) *with* $\mathfrak{D} := \mathfrak{D}_\Lambda$. *Let*

$$Q := \begin{bmatrix} \sqrt{m_1} & & \\ & \ddots & \\ & & \sqrt{m_n} \end{bmatrix}, \quad q := \begin{bmatrix} \frac{1}{m_1}\sum_{j=1}^{m_1} y_{1j} & \cdots & \frac{1}{m_n}\sum_{j=1}^{m_n} y_{nj} \end{bmatrix},$$

$$c := \sqrt{\sum_{i=1}^{n}\left[\epsilon^2 m_i + \frac{1}{m_i}\left(\sum_{j=1}^{m_i} y_{ij}\right)^2 - \sum_{j=1}^{m_i} y_{ij}^2\right]},$$

$$A := \begin{bmatrix} t_2 - t_1 & t_3 + t_1 - 2t_2 & t_3 - t_2 & & & \\ & t_3 - t_2 & t_4 + t_2 - 2t_3 & t_4 - t_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & t_{n-1} - t_{n-2} & t_n + t_{n-2} - 2t_{n-1} & t_n - t_{n-1} \\ & & & & 1 & -1 \end{bmatrix},$$

and

$$b := [0 \;\, t_2 - t_1 \;\, \cdots \;\, t_n - t_1]^T.$$

Also denote by e *the vector of all ones. For* $\epsilon > \epsilon_{min}$, *model* (FRP) *can be reformulated as a* *second-order cone program:*

$$\max_{s,\phi,\psi,\theta,\pi,\eta} \; q^T A^T \phi + q^T \psi - q^T \theta - \eta \tag{16a}$$

$$s.t. \; \|\pi\|_2 \leq \eta \tag{16b}$$

$$cQ^{-1}A^T\phi + cQ^{-1}\psi - cQ^{-1}\theta - \pi = 0 \tag{16c}$$

$$b^T\psi \geq s^2 - (p + t_1)s + pt_1 \tag{16d}$$

$$e^T\psi \leq s - p \tag{16e}$$

$$\phi \geq 0, \;\; \psi \geq 0, \;\; \theta \geq 0. \tag{16f}$$

Proof: We denote by $\widetilde{\mathfrak{D}}$ the subset of $\mathfrak{D}_\Lambda$ consisting of all piecewise linear decreasing concave functions with break points t_i for $i = 1, \ldots, n$. For a given $s \in \mathfrak{S}$, let $\tilde{d}(\cdot)$ be an optimal solution of the inner minimization problem of model (FRP),

$$\min_{d \in \mathfrak{D}_\Lambda} (s - p)d(s). \tag{17}$$

We construct a piecewise linear decreasing concave function as

$$\tilde{d}_n(s) := \sum_{i=1}^{n-1} \left(\frac{\tilde{d}(t_{i+1}) - \tilde{d}(t_i)}{t_{i+1} - t_i} s + \frac{t_{i+1}\tilde{d}(t_i) - t_i\tilde{d}(t_{i+1})}{t_{i+1} - t_i} \right) \mathbf{1}\left\{t_i \leq s \leq t_{i+1}\right\},$$

where $\mathbf{1}\{\cdot\}$ is an indicator function. Functions $\tilde{d}(\cdot)$ and $\tilde{d}_n(\cdot)$ have the same values at all common t_i, and at non-common points, the values of $\tilde{d}_n(\cdot)$ is not greater than $\tilde{d}(\cdot)$. It is straightforward to see that $\tilde{d}_n(\cdot) \in \mathfrak{D}_\Lambda$, and $\tilde{d}_n(s) \leq \tilde{d}(s)$ for all $s \in \mathfrak{S}$. Therefore, $\tilde{d}_n(\cdot)$ is also a minimizer of (17). It implies that (17) is equivalent to

$$\min_{d \in \widetilde{\mathfrak{D}}} (s - p)d(s),$$

which we claim can be formulated by the quadratic constraint program as

$$\min_{z,u} \; (-s^2 + (p+t_1)s - pt_1)z_1 + (s-p)z_2 \tag{18a}$$

$$\text{s.t. } \frac{1}{\sum_{i=1}^n m_i} \sum_{i=1}^n \sum_{j=1}^{m_i} (y_{ij} - u_i)^2 \le \epsilon^2 \tag{18b}$$

$$\frac{u_{i+2} - u_{i+1}}{t_{i+2} - t_{i+1}} \le \frac{u_{i+1} - u_i}{t_{i+1} - t_i}, \quad i = 1, \ldots, n-2 \tag{18c}$$

$$u_{n-1} \le u_n \tag{18d}$$

$$u_i + (t_i - t_1)z_1 - z_2 \le 0, \quad i = 1, \ldots, n \tag{18e}$$

$$z_1 \ge 0, \; z_2 \ge 0, \; u_i \ge 0, \quad i = 1, \ldots, n. \tag{18f}$$

Conditions (18c) and (18d) ensure that feasible piecewise linear demand functions with break points t_i, and values u_i at the break points

$$d(s) := \sum_{i=1}^{n-1} \left(\frac{u_{i+1} - u_i}{t_{i+1} - t_i} s + \frac{t_{i+1}u_i - t_i u_{i+1}}{t_{i+1} - t_i} \right) \mathbf{1}\left\{ t_i \le s \le t_{i+1} \right\}.$$

are decreasing and concave. Under conditions (18c) and (18d), the function $d(\cdot)$ is written as the minimization of all subgradients at points (t_i, u_i)

$$\begin{aligned} d(s) = \min_{z} \; & -z_1(s - t_1) + z_2 \\ \text{s.t. } & u_i + (t_i - t_1)z_1 - z_2 \le 0, \quad i = 1, \ldots, n, \\ & z \ge 0. \end{aligned}$$

This gives the objective function (18a) and condition (18e). Let $u := (u_1, \ldots, u_n)^T$ and $v := c^{-1}Q(u - q)$. To simplify the mathematical representation, we write problem (18) in the matrix form as

$$
\begin{aligned}
\min_{z,v}\ & (-s^2 + (p+t_1)s - pt_1)z_1 + (s-p)z_2 \\
\text{s.t.}\ & \|v\|_2 \le 1 \\
& -cAQ^{-1}v \ge Aq \\
& -cQ^{-1}v - bz_1 + ez_2 \ge q \\
& cQ^{-1}v \ge -q \\
& z_1 \ge 0,\ z_2 \ge 0.
\end{aligned}
$$

The corresponding dual problem is given by

$$
\begin{aligned}
\max_{\phi,\psi,\theta,\pi,\eta}\ & q^T A^T \phi + q^T \psi - q^T \theta - \eta \\
\text{s.t.}\ & \|\pi\|_2 \le \eta \\
& cQ^{-1}A^T\phi + cQ^{-1}\psi - cQ^{-1}\theta - \pi = 0 \\
& b^T\psi \ge s^2 - (p+t_1)s + pt_1 \\
& e^T\psi \le s - p \\
& \phi \ge 0,\ \psi \ge 0,\ \theta \ge 0.
\end{aligned}
\tag{19}
$$

By the assumption that $\epsilon > \epsilon_{min}$, we now prove that the Slater condition holds for the primal problem (18) such that the primal problem (18) and the dual problem (19) satisfy strong duality. The decision variables z_1 and z_2 can be chosen to make condition (18e) hold strictly. To show the Slater condition, we only need to find a demand function $d(\cdot) \in \widetilde{\mathfrak{D}}_\Lambda$ such that conditions (18b) - (18d) strictly hold for all $u_i = d(t_i)$. We now construct a piecewise linear $\hat{d}(\cdot)$ in $[t_1, t_n]$ with break points t_i for $i = 1, \ldots, n$, which satisfies

$$
\begin{aligned}
& \frac{\hat{d}(t_{i+2}) - \hat{d}(t_{i+1})}{t_{i+2} - t_{i+1}} < \frac{\hat{d}(t_{i+1}) - \hat{d}(t_i)}{t_{i+1} - t_i}, \quad i = 1, \ldots, n-2, \\
& \hat{d}(t_{n-1}) < \hat{d}(t_n), \\
& 0 \le \hat{d}(t_i) < \infty, \quad i = 1, \ldots, n.
\end{aligned}
$$

Let $\hat{\epsilon} := \text{Err}(\hat{d}) < \infty$. If $\hat{\epsilon} < \epsilon$, we have $\hat{d}(\cdot) \in \widetilde{\mathfrak{D}}$, and conditions (18b) - (18d) strictly hold for all $u_i = \hat{d}(t_i)$. Hence, the Slater condition holds for problem (18). Suppose $\hat{\epsilon} \geq \epsilon$. Let $\tilde{d}(\cdot)$ be an optimal solution of the data fitting problem (7). Thus $\tilde{d}(\cdot) \in \widetilde{\mathfrak{D}}$. Choose $\lambda \in \left(0, \frac{\hat{\epsilon}-\epsilon}{\hat{\epsilon}-\epsilon_{min}}\right)$, and construct a function $d_\lambda(\cdot) := \lambda\tilde{d}(\cdot) + (1-\lambda)\hat{d}(\cdot)$. Using the Minkowski inequality, it follows that

$$\text{Err}(d_\lambda) \leq \lambda\text{Err}(\tilde{d}) + (1-\lambda)\text{Err}(\hat{d}) = \lambda\epsilon_{min} + (1-\lambda)\hat{\epsilon} < \epsilon.$$

It is also straightforward to see that

$$\begin{aligned}
&\frac{d_\lambda(t_{i+2}) - d_\lambda(t_{i+1})}{t_{i+2} - t_{i+1}} < \frac{d_\lambda(t_{i+1}) - d_\lambda(t_i)}{t_{i+1} - t_i}, \quad i = 1, \dots, n-2, \\
&d_\lambda(t_{n-1}) < d_\lambda(t_n), \\
&d_\lambda(t_i) \geq 0, \quad i = 1, \dots, n.
\end{aligned}$$

We have that $d_\lambda(\cdot) \in \widetilde{\mathfrak{D}}$ and that conditions (18b) - (18d) strictly hold for all $u_i = d_\lambda(t_i)$. Hence, (18) satisfies the Slater condition. Therefore, model (FRP) is equivalent to (16). $\square$

4 An Example Using Porterhouse Beef Data

In this section we present an application of model (FRP) to a grocery store's decision on selling pricing and order quantity of porterhouse beef. Figure 3 plots the monthly selling price and market demand of porterhouse beef from January 2001 to July 2005 published by the Economic Research Service of the US Department of Agriculture in cooperation with the Livestock Marketing Information Center. Both the selling price and the market demand have a large variation. The selling price per pound varies from \$4.81/lb to \$8.60/lb, and the demand varies between 25 Mil lbs and 157 Mil lbs. We model this problem using (FRP), where the uncertainty set of demand functions is specified as in (3). In what follows, the model is tested by changing the error tolerance and the purchase price. Insights and analyses of computational results are now discussed to demonstrate the effectiveness of (FRP).

4.1 Computational Performance

We coded Algorithm 1 using MATLAB and Gurobi 6.0.0, and tests were performed on a laptop with Inter(R) Core (TM) i7-3610QM CPU (2.30GHz). We set $\delta = 10^{-5}$ in the algorithm, and use $\kappa := \epsilon/\epsilon_{min}$ as the relative error tolerance level. For our dataset, $\epsilon_{min} = 18.04$. Table 1 reports the number of iterations and the running time of Algorithm 1 for $\kappa = 1, 1.04, \ldots, 1.24$, but the purchase price is fixed at \$6. The increase of κ expands the uncertainty set of demand functions. The number of iterations in the algorithm increases from 2 to 202, but the running time increases from 2.023 seconds to 9.621 seconds. The increase in the number of iterations appears linear in the value of κ. Table 2 records the computational results when the purchase price p is changed from \$4.5/lb to \$7.5/lb but κ is fixed at 1.08. The number of iterations and the running time are around 48 and 4.69 seconds, and appear independent of p.

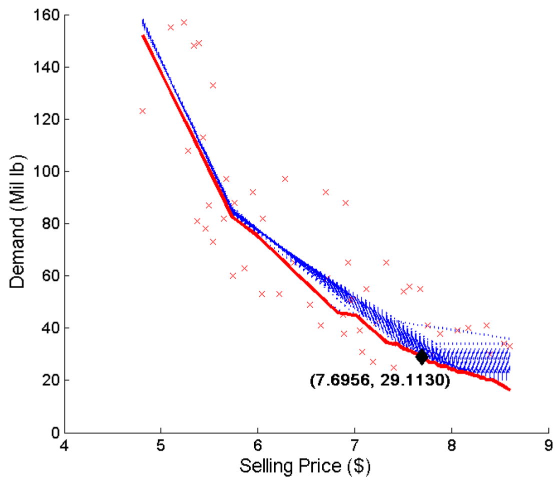
In Figure 3 the dotted curves show the generated cuts, and the solid curves represent the lower envelop functions produced by Algorithm 1. The optimal selling prices and order quantities given by Algorithm 1 are located on the lower envelop functions. These cuts are guaranteed to be decreasing piecewise linear convex functions, while the lower envelop functions are decreasing piecewise linear but non-convex. In addition, these cuts differ in the region where the selling price is above the given purchase price $p = \$6.0/lb$. As κ increases, the generated cuts are more spread out, and the lower envelop functions drop significantly for a larger selling price s.

Purchase price (\$/lb)	Relative error tolerance level κ	Number of functional cuts	Running time (s)	Optimal selling price (\$/lb)	Optimal order (Mil lb)	Optimal Value (Mil \$)
6.0	1.00	2	2.023	7.69	36.08	61.04
	1.04	21	2.518	7.70	29.11	49.36
	1.08	49	4.122	7.76	23.75	41.91
	1.12	83	6.033	7.58	23.55	37.14
	1.16	130	7.489	7.45	22.82	33.09
	1.20	168	8.640	7.45	20.77	30.12
	1.24	202	9.621	7.45	18.88	27.38

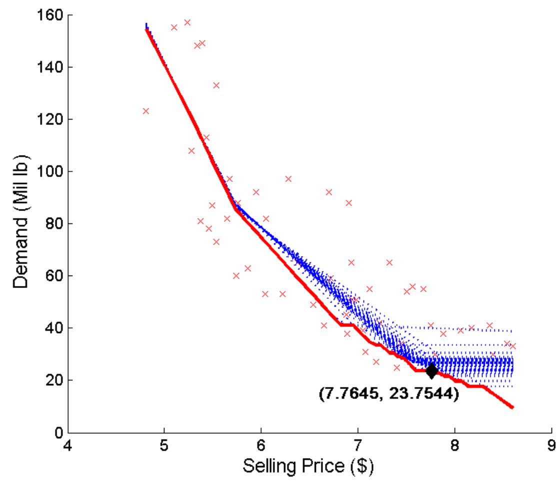
Table 1: Computational results for the relative error tolerance level

Relative error tolerance level κ	Purchase price (\$/lb)	Number of functional cuts	Running time (s)	Optimal selling price (\$/lb)	Optimal order (Mil lb)	Optimal value (Mil \$)
	4.5	41	4.492	6.72	48.87	108.46
	5.0	52	5.581	6.99	41.68	83.10
	5.5	45	4.016	7.20	35.67	60.73
1.08	6.0	49	4.122	7.76	23.75	41.91
	6.5	50	4.893	7.87	22.34	30.62
	7.0	51	5.078	8.20	20.48	24.51
	7.5	49	4.651	8.45	19.64	18.68

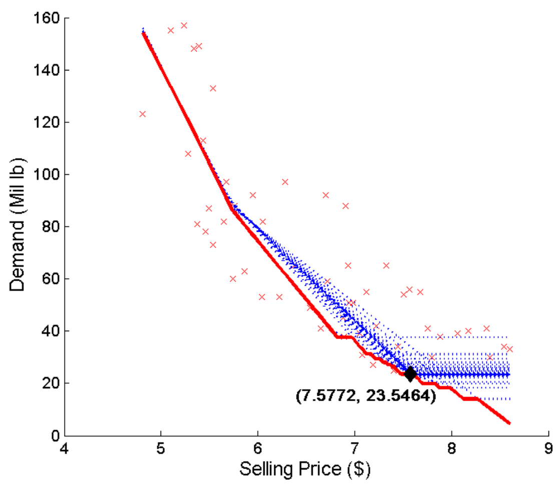
Table 2: Computational results for the purchase price



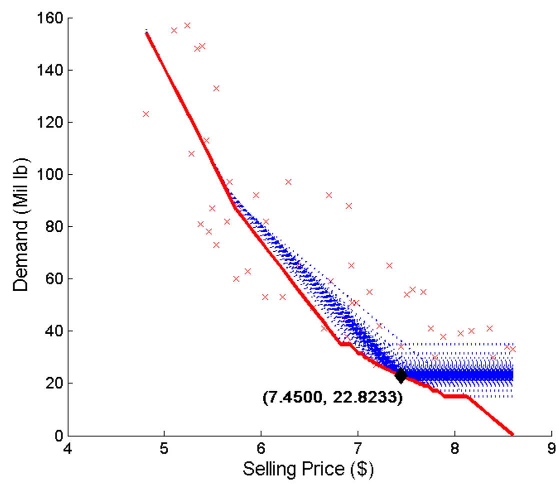
(a) $p = 6.0, \kappa = 1.04$



(b) $p = 6.0, \kappa = 1.08$

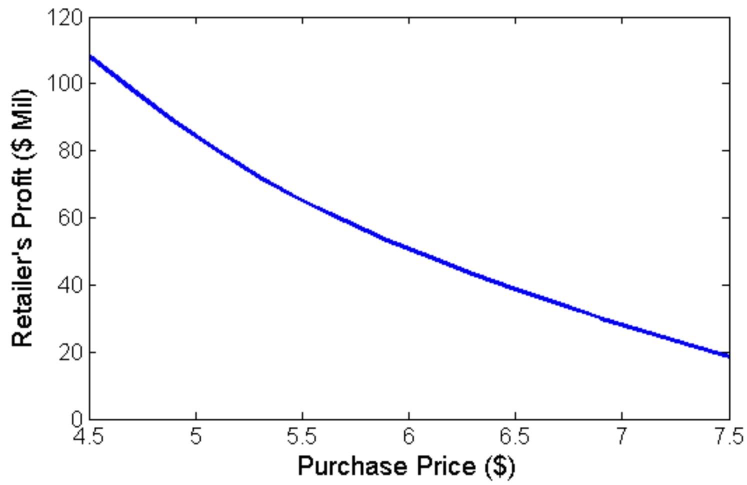


(c) $p = 6.0, \kappa = 1.12$

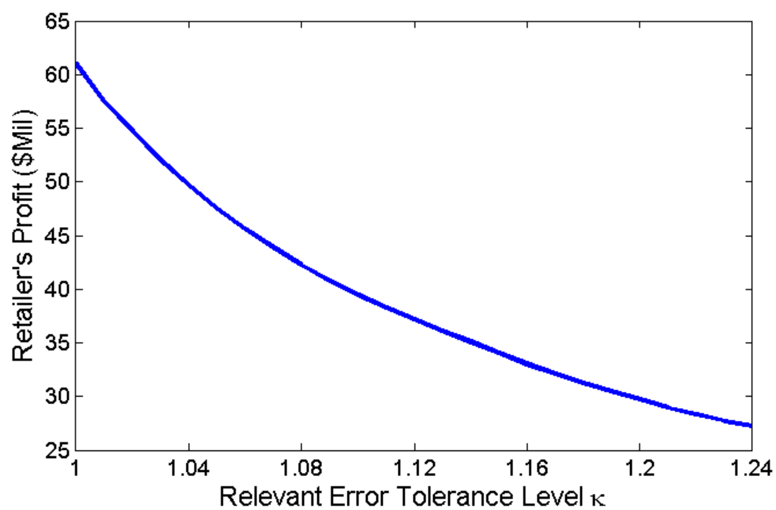


(d) $p = 6.0, \kappa = 1.16$

Figure 3: Generated cuts and lower envelop function for the relative error tolerance level



(a) Purchase price v.s. maximum profit ($\kappa = 1.08$)



(b) Relative error tolerance level v.s. maximum profit ($p = 6.0$)

Figure 4: The maximum profit with respect to the purchase price and the relative error tolerance level

4.2 Model Performance

We now discuss the market performances of (FRP). Tables 1 and 2 report optimal selling prices, order quantities, and objective values for various settings of κ and p. Note that the optimal value is the multiplication of the optimal selling price and the optimal order quantity. We call the derivative of the optimal value $r^*(p, \epsilon)$ with respect to p as the maximum marginal profit (see Theorem 2.4).

A retailer's concern is to know how to react to the changes in the purchase price. We fix the relative error tolerance level $\kappa = 1.08$ and increase the purchase price p from \$4.5/lb to \$7.5/lb. The change increases the optimal selling price from \$6.72/lb to \$8.45/lb, decreases the optimal order quantity from 48.87 Mil lbs to 19.64 Mil lbs, and reduces the maximum profit from \$108.46 million to \$18.68 million. Figure 4a gives retailer's profit as a function of the purchase price. This curve is convex as shown in Theorem 2.4. Moreover, Theorem 2.4 implies that the decrease in the optimal order quantity reduces the maximum marginal profit. It is observed from Table 2 that \$0.5 change in the purchase price p makes a significant drop in the maximum profit when p is small, but incur less change for larger p. The retailer's maximum profit drops by \$25.36 million when p increases from \$4.5/lb to \$5.0/lb. In comparison, when p increases from \$7.0/lb to \$7.5/lb, there is only \$5.83 million reduction.

A retailer's additional concern is uncertainty in market demand function. A large relative error tolerance κ expands the uncertainty set of demand functions. The computational results in Figure 3 show that, as κ increases, the lower envelop function drops down; therefore, there is less possibility of overestimating the market demand. To verify the impact of κ on the maximum profit, we fix the purchase price $p = \$6.0/lb$, and change κ from 1 to 1.24. This generates a decreasing convex curve of the maximum profit shown in Figure 4b. Choosing $\kappa = 1$, we obtain the smallest feasible uncertainty set of demand functions, which is the collection of all best fitted nondecreasing convex demand functions under the least-squares fitting criterion. When κ increases from 1 to 1.12, the optimal selling price changes between \$7.57/lb and \$7.76/lb. However, when we further increase κ to 1.16, the optimal selling price is stable at \$7.4500/lb. On the other hand, the increase in κ reduces the optimal order quantity from 36.08 Mil lbs to 18.88 Mil lbs. This observation suggests a expected behavior that, when a retailer is uncertain about the demand function, it is better for the retailer to order a smaller quantity. We also observe that the maximum profit decrease from

\$61.04 million to \$27.38 million as κ increases. An over-conservative retailer earns a smaller profit.

5 Conclusions

We have proposed a novel functionally robust model, where we have specified an uncertainty set of nonparametric demand functions. This functional robust model does not assume that the form of a model function must be pre-determined.

We have discussed the impact of functional robustness and the purchase price on the maximum profit that a retailer can earn. The performances of the functionally robust model exhibits the principle that a conservative retailer may choose to reduce the order quantity for avoiding risks due to a lack of knowledge of the true demand function, and compromise with less profit. Solution methods have been given for the two cases where demand functions are either convex or concave. We have developed a cut surface algorithm in the former case, and reformulated the model as a second-order conic program in the latter case.

The functional robustness has been applied to the problem of a grocery store's decision on selling price and order quantity of porterhouse beef. Performance of the functionally robust model has been analyzed to numerically verify the impact of the functional robustness and the purchase price on the maximum profit that a retailer can earn. The concept of functional robustness introduced in this paper will have further applications within and outside the domain of retailer pricing and ordering problems.

Acknowledgments This research was supported by NSF Grant CMMI-1362003.

References

R. G. D. Allen. *Mathematical Analysis For Economists.* Brunton Press, 2008.

A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

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

A. Ben-Tal and A. Nemirovski. Robust solutions of uncertain quadratic and conic-quadratic programs. *SIAM Journal on Optimization*, 13:535–560, 2002.

A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. Robustness. In R. Saigal, L. Vandenberghe, and H.Wolkowicz, editors, *Handbook of Semidefinite Programming*, pages 139–162. Kluwer Academic, Dordrecht, The Netherlands, 2000.

D. Bertsimas and I. Popescu. Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal on Optimization*, 15(3):780–804, 2005.

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

D. Bertsimas and M. Sim. Tractable approximations to robust conic optimization problems. *Mathematical Programming*, 107:5–36, 2006.

D. Bertsimas, D. Pachamanova, and M. Sim. Robust linear optimization under general norm. *Operations Research Letters*, 32(6):510–516, 2004.

D. Bertsimas, X. V. Doan, K. Natarajan, and C. Teo. Models for minimax stochastic linear optimization problems with risk aversion. *Mathematics of Operations Research*, 35(3):580–602, 2010.

D. Bertsimas, D. B. Brown, and C. Caramnis. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011.

H. Beyer and B. Sendhoff. Robust optimization - a comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, 196:3190–3218, 2007.

L. M. A. Chan, Z. J. Shen, D. Simchi-Levi, and J. L. Swann. Coordination of pricing and inventory decisions: A survey and classification. In D. Simchi-Levi, S. D. Wu, and Z. J. Shen, editors, *Handbook of Quantitative Supply Chain Analysis: Modeling in the E-Business Era*, pages 335–392. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004.

N. Chehrazi and T. A. Weber. Monotone approximation of decision problems. *Operations Research*, 58(4):1158–1177, 2010.

X. Chen and D. Simchi-Levi. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research*, 52(6):887–896, 2004.

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

J. Dupačová. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20(1):73–88, 1987.

L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18:1035–1064, 1997.

V. Gabrel, C. Murat, and A. Thiele. Recent advances in robust optimization: An overview. 2013. URL `http://www.optimization-online.org/DB_FILE/2012/07/3537.pdf`.

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

M. L. Greenhut, G. Norman, and C. shun Hung. *The Economics of Imperfect Competition, a spatial approach.* Cambridge University Press, 1987.

J. Hu and S. Mehrotra. Robust decision making using a general utility set. `http://www.optimization-online.org/DB_FILE/2012/07/3527.pdf`, 2012.

J. Hu and S. Mehrotra. Robust decision making using a risk averse utility set with an application to portfolio optimization. *IIE Transactions*, 47(4):358–372, 2015.

S. Karlin and C. Carr. Prices and optimal inventory policies. In S. Karlin, K. J. Arrow, and H. Scarf, editors, *Studies in Applied Probability and Management Science*. Stanford University Press, Stanford, CA, 1962.

P. Klemperer and M. Meyer. Price competition vs. quantity competition: The role of uncertainty. *The RAND Journal of Economics*, 17(4):618–638, 1986.

A. Kocabiyikoğlu and I. Popescu. An elasticity approach to the newsvendor with price-sensitive demand. *Operations Research*, 59(2):301–312, 2011.

D. A. Malueg. Monopoly output and welfare: The role of curvature of the demand function. *The Journal of Economic Education*, 25(3):235–250, 1994.

N. G. Mankiw. *Principles of Economics*. South-Western Cengage Learning, 2008.

A. Mas-Colell, M. D. Whinston, and J. R. Green. *Microeconomic Theory*. Oxford University Press, 1995.

S. Mehortra and D. Papp. A cutting surface algorithm for semi-infinite convex programming with an application to moment robust optimization. *SIAM Journal on Optimization*, 24(4):1670–1697, 2014.

S. Mehrotra and H. Zhang. Models and algorithms for distributionally robust least squre problems. *Mathematical Programming*, 146(1-2):123–141, 2014.

E. S. Mills. Uncertainty and price theory. *Quarterly Journal of Economics*, 73(1):117–130, 1959.

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

N. C. Petruzzi and M. Dada. Pricing and the newsvendor problem: A review with extensions. *Operations Research*, 47(2):183–194, 1999.

G. Pflug and D. Wozabal. Ambiguity in portfolio selection. *Quantitative Finance*, 7(4):435–442,, 2007.

A. Prékopa. *Stochastic Programming*. Kluwer Academic, 1995.

G. Raz and E. L. Porteus. A fractiles perspective to the joint price/quantity newsvendor model. *Management Science*, 52(11):1764–1777, 2006.

H. Scarf. A min-max solution of an inventory problem. In *Studies in the Mathematical Theory of Inventory and Production*, chapter 12, pages 201–209. 1958.

A. Shapiro and S. Ahmed. On a class of minimax stochastic programs. *SIAM Journal on Optimization*, 14(4):1237–1249, 2004.

A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming : Modeling and theory.* SIAM, 2009.

S. Tayur, R. Gareshan, and M. Magazine. *Quantitative Models for Supply Chain Management.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.

T. M. Whitin. Inventory control and price theory. *Management Science*, 2(1):61–68, 1955.

C. A. Yano and S. M. Gilbert. Coordinated pricing and production/procurement decisions: A review. In A. K. Chakravarty and J. Eliashberg, editors, *Managing Business Interfaces: Marketing, Engineering, and Manufacturing Perspectives.* Kluwer Academic, Norwell, MA, 2004.

L. Young. Price, inventory, and the structure of uncertain demand. *New Zealand Operational Research*, 6(2):157–177, 1978.

Z. Zhu, J. Zhang, and Y. Ye. Newsvendor optimization with limited distribution information. *Optimization Methods and Software*, 28(3):640–667, 2013.