Chapter 8
Basic Pricing Theory

8.1 Introduction

This chapter provides an introduction to multi-product monopoly pricing when the variable costs are linear. Profit maximization problems with linear variable costs arise from capacity constraints, where the firm maximizes the expected profit net of the opportunity costs of the capacities used. We argue that under mild assumptions, both the optimal profit function and the expected consumer surplus are convex functions of the variable costs. Consequently, when variable costs are random, both the firm and the representative consumer benefit from prices that dynamically respond to changes in variable costs. Randomness in variable cost is often driven by randomness in demand in conjunction with capacity constraints, and this accounts for some of the benefits of dynamic pricing. We explore conditions for the existence and uniqueness of maximizers of the expected profit and analyze in detail problems with capacity constraints both when prices are set for the entire sales horizon a priori, and when prices are allowed to change during the sales horizon. The firm's problem is discussed in Sect. 8.2, while the representative consumer's problem is presented in Sect. 8.3. The case with finite capacity is discussed in Sect. 8.4. Details about existence and uniqueness for single product problems are discussed in Sect. 8.5. This section also includes applications to priority pricing, social planning, multiple market segments, and peak-load pricing. Multi-product pricing problems are discussed in Sect. 8.6.

G. Gallego, H. Topaloglu, *Revenue Management and Pricing Analytics*,
International Series in Operations Research & Management Science 279,
https://doi.org/10.1007/978-1-4939-9606-3_8

8.2 The Firm's Problem

Consider a firm with variable cost vector $z = (z_1, \ldots, z_n)$ for n products. The firm's profit function is given by

$$R(p, z) := (p - z)'d(p) = \sum_{i=1}^{n}(p_i - z_i)\, d_i(p_1, \ldots, p_n), \tag{8.1}$$

where p and $d(p)$ are the vector of prices and expected demands as a function of prices, all of dimension $n \geq 1$. The goal is to find a prices, say $p(z)$, that maximizes $R(p, z)$. The profit function (8.1) models situations with linear variable costs. Linear costs arise as dual variables of capacity constraints, and this is our primary motivation for the study of this model. The maximum profit, as a function of z, is given by

$$\mathcal{R}(z) := \max_{p \in X} R(p, z), \tag{8.2}$$

where X is the set of allowable prices.

The set of allowable prices $X = X_1 \times \ldots X_n$ defines different type of optimization problems. The *assortment* optimization problem arises when $X_i = \{r_i, \bar{r}_i\}$, where r_i is the regular price of product i and $\bar{r}_i$ is the choke-off price, also known as the null-price for product i, so demand for product i is zero whenever it is priced at or above $\bar{r}_i$. The choke price $\bar{r}_i$ may be finite or infinity. On occasions, we will find it convenient to write ∞ instead of $\bar{r}_i$, and this should be interpreted as not offering product i. The *joint assortment and pricing* problem can be modeled by setting $X_i = \{r_{i1}, \ldots, r_{in_i}, \bar{r}_i\}$, where there is a finite price menu that includes the option of not offering product i. As an example, a product may be offered at the regular price, at a discounted price, or not offered at all. When the set X is finite, a maximizer $p(z)$ is guaranteed to exist. These are combinatorial problems and even relatively simple versions can be NP-hard. Fortunately, as we have seen in the chapter on assortment optimization, there are instances of practical importance that can be solved efficiently, sometimes by linear programming.

In this chapter, we are mainly concerned with the *continuous pricing* problem where $X_i = \Re_+ = [0, \infty)$ for all $i = 1, \ldots, n$. For this problem, we should more formally write $\mathcal{R}(z) = \sup_{p \in X} R(p, z)$ as the maximum may not be attained. We will later investigate conditions that guarantee the existence and uniqueness of a finite maximizer $p(z)$, as well as comparative statics that inform us of how $p(z)$ changes with z. To facilitate the exposition, we would use max instead of sup, but all the arguments except where noted would continue to hold for sup. Regardless of whether an optimizer exists or is unique, we can show that the profit function $\mathcal{R}(z)$ is a decreasing convex function of the cost vector z.[1]

Theorem 8.1 *$\mathcal{R}(z)$ is decreasing convex in z.*

8.2.1 Random Costs

We now investigate the impact of randomness in the variable cost vector. The motivation for this is that in dynamic pricing, the variable cost vector depends on the remaining capacity and time-to-go, and since the remaining capacity depends on random realizations of arrivals and sales, it follows that the variable costs change randomly over time. Let Z denote a vector of random unit costs. Then, by Jensen's inequality $\mathbb{E}[\mathcal{R}(Z)] \geq \mathcal{R}(\mathbb{E}[Z])$. The difference between $\mathbb{E}[\mathcal{R}(Z)]$ and $\mathcal{R}(\mathbb{E}[Z])$ can be interpreted as the difference between a dynamic pricing policy $p(Z)$ that responds to changes in Z and a static pricing policy $p(\mathbb{E}[Z])$ that does not. The following proposition allows us to assess the difference between $\mathbb{E}[\mathcal{R}(Z)]$ and $\mathcal{R}(\mathbb{E}[Z])$.

Proposition 8.2 *If the function $\mathcal{R} : \Re^n \to \Re$ is twice continuously differentiable, and H is the Hessian of $\mathcal{R}(z)$ evaluated at $\mathbb{E}[Z]$, then*

$$\mathbb{E}[\mathcal{R}(Z)] - \mathcal{R}(\mathbb{E}[Z]) \simeq \frac{1}{2}\mathbb{E}[(Z - \mathbb{E}[Z])' H (Z - \mathbb{E}[Z])] \geq 0.$$

This suggests that the difference between static and dynamic pricing is large when the variance of Z is large and $\mathcal{R}$ has significant curvature at $\mathbb{E}[Z]$. On the other hand, in situations where there is little variance in Z or little curvature in $\mathcal{R}$ around $\mathbb{E}[Z]$, we expect dynamic pricing to be of little help in improving profits. Proposition 8.2 follows by taking a second-order Taylor expansion of $\mathcal{R}(z)$ and using convexity.

Example 8.3 Suppose $d(p) = 100(1 - p)$ for $p \in [0, 1]$ and $d(p) = 0$ for $p > 1$. Then $R(p, z) = 100(p - z)(1 - p)$ is maximized at $p(z) = 0.5(1 + z)$ for $z \in [0, 1)$, $z \leq p(z) \leq 1$, and $\mathcal{R}(z) = 25(1 - z)^2$. If Z is random, with $P(Z = 1/3) = P(Z = 2/3) = 0.5$, then by Jensen's inequality \$6.94 $= \mathbb{E}[\mathcal{R}(Z)] \geq \mathcal{R}(\mathbb{E}[Z]) = R(p(\mathbb{E}[Z]), \mathbb{E}[Z]) = \mathbb{E}[R(p(\mathbb{E}[Z]), Z)] =$ \$6.25, so a firm that responds to changes in Z makes 11.1% more profits than one who prices based on $\mathbb{E}[Z]$. Furthermore, a firm who can respond to changes in variable costs benefits from randomness in costs.

[1] We use the terms increasing, decreasing, concave and convex in the weak sense unless stated otherwise.

The following corollary pushes the idea a bit further.

Corollary 8.4 *If* $g(z) : \Re^m \to \Re^n_+$ *is increasing in* z*, then* $\mathcal{R}(g(z))$ *is decreasing in* z*. If* $g(z)$ *is also concave, then* $\mathcal{R}(g(z))$ *is convex in* z*. Moreover, if* $Z \in \Re^m$ *is random, then* $\mathbb{E}[\mathcal{R}(g(Z))] \geq \mathcal{R}(g(\mathbb{E}[Z]))$.

We can interpret $g(z)$ as the vector of unit costs for the products and z as the vector of unit costs of the resources that are used to build the products. As an example, if $g(z) = A'z$ and $A \geq 0$ is an $m \times n$ matrix, with A_{ij} the number of units of resource i required for product j, then $\mathcal{R}(A'z)$ is convex in z. This shows that a risk-neutral firm is better off with random component costs Z than with deterministic component costs equal to $\mathbb{E}[Z]$, provided it can charge prices $p(A'Z)$. The case $A = I$ represents the case where product i only uses component i, and results in the price vector $p(Z)$.

8.3 The Representative Consumer's Problem

While the firm is better off using dynamic pricing $p(Z)$, the reader may wonder whether consumers are better off with $p(Z)$ or with $p(\mathbb{E}[Z])$. In other words, do consumers prefer dynamic or static prices?

To answer this question, we will use the framework of utility theory, where we frame the question in terms of the surplus of the representative consumer. Suppose that a representative consumer derives utility $U(q)$ from purchasing a non-negative vector $q = (q_1, \ldots, q_n)$ of products. It is typically assumed that $U(q)$ is an increasing concave function of q. The *net utility* or consumers' surplus at (q, p) is given by

$$S(q, p) := U(q) - q'p,$$

which is simply the utility $U(q)$ minus the cost of purchasing the bundle q at prices $p = (p_1, \ldots, p_n)$. The *optimal surplus*, also known as the net indirect utility, in the absence of a budget constraint is given by

$$\mathcal{S}(p) := \max_{q \geq 0} S(q, p).$$

The solution, say $q = d(p)$, if it exists, gives us the bundle demanded at p. The first-order condition for optimality (ignoring the non-negativity constraints) is $\nabla U(q) - p = 0$, which is also sufficient given the assumed concavity. If there exists an inverse function, say $\nabla^{-1}U(p)$, then $d(p) = \nabla^{-1}U(p)$ for all p in the set $\mathcal{P} = \{p : \nabla^{-1}U(p) \geq 0\}$. The demand function $d(p)$ can be extended to a set larger than $\mathcal{P}$, but then at least one of the non-negativity constraints on q will be binding, so the problem needs to be projected into a subspace of non-negative demands. The following theorem shows that $\mathcal{S}(p)$ is decreasing convex in p and how $\mathcal{S}(p)$ changes with $p \in \mathcal{P}$.

Theorem 8.5 *$\mathcal{S}(p)$ is decreasing convex in p. Moreover, if $d(p)$ is differentiable in $p \in \mathcal{P}$, then $\nabla\mathcal{S}(p) = -d(p) \leq 0$ for all p in the interior of the set $\mathcal{P}$.*

One implication from Theorem 8.5 is that if p is one dimensional, then $\mathcal{S}(p) = \int_p^\infty d(x)dx$. In particular, if Ω is a random variable with finite mean, and $d(p) = P(\Omega \geq p)$ then

$$\mathcal{S}(p) = \int_p^\infty P(\Omega \geq x)dx = \mathbb{E}[(\Omega - p)^+],$$

so the expected surplus for a consumer with willingness to pay Ω is the expectation of the gain, $(\Omega - p)^+$, from the transaction if it happens.

Returning to the n-dimensional case, Jensen's inequality and the convexity of $\mathcal{S}(p)$ imply that if the price vector P is random with finite expectation $\mathbb{E}[P]$, then $\mathbb{E}[\mathcal{S}(P)] \geq \mathcal{S}(\mathbb{E}[P])$, so consumers prefer random prices P over expected prices $\mathbb{E}[P]$. Notice that this result does not assume that the representative consumer is risk-neutral, as we have already taken into account risk preference through the utility function. This result gives us hope that customers may prefer dynamic prices $p(Z)$ over static prices $p(\mathbb{E}[Z])$. The next result gives sufficient conditions for this to be true.

Corollary 8.6 *If $p(z) : \Re^n \to \Re^n_+$ is increasing in z, then $\mathcal{S}(p(z))$ is decreasing in z. If $p(z)$ is increasing concave in z, then $\mathcal{S}(p(z))$ is decreasing convex in z. Moreover, if $Z \in \Re^m$ is random and $\mathcal{S}(p(z))$ is convex in z, then*

$$\mathbb{E}[\mathcal{S}(p(Z))] \geq \mathcal{S}(p(\mathbb{E}[Z])).$$

This result follows directly from Corollary 8.4. Intuitively, if $p(Z)$ is increasing concave, then $\mathbb{E}[p(Z)] \leq p(\mathbb{E}[Z])$, so prices are lower on average under dynamic pricing. If we combine this with the fact that consumers prefer random prices, we obtain a weaker sufficient condition, namely that $\mathbb{E}[p(Z)] \leq p(\mathbb{E}[Z])$ implies $\mathbb{E}[\mathcal{S}(p(Z)) \geq \mathcal{S}(p(\mathbb{E}[Z]))]$. To verify this, notice that

$$\mathbb{E}[\mathcal{S}(p(Z))] \geq \mathcal{S}(\mathbb{E}[p(Z)]) \geq \mathcal{S}(p(\mathbb{E}[Z])),$$

where the first inequality follows from the convexity of $\mathcal{S}(p)$, and the second from the assumption that $\mathbb{E}[p(Z)] \leq p(\mathbb{E}[Z])$, and the fact that $\mathcal{S}(p)$ is decreasing in p.

We will now argue that $\mathcal{S}(p(z))$ is convex up to a quadratic approximation of an increasing concave utility function.

Theorem 8.7 *The function $\mathcal{S}(p(z))$ is convex in z up to a quadratic approximation of any increasing concave utility function.*

For the single product case, it is possible to show that $p(z)$ is linear if and only if the demand function belongs to one of the following three classes:

- $d(p) = \lambda \exp(-p/\theta)$ for $\lambda, \theta > 0$: the exponential demand.

- $d(p) = ((a - bp)^+)^c$ for $a, b, c > 0$: root-linear demand (linear for $c = 1$).
- $d(p) = (a+bp)^{-c}$ for $a, b > 0$, $c > 1$: constant elasticity of substitution demand ($a = 0$).

The class of single product demand functions for which $p(z)$ is linear contains many of the demand functions that appear commonly in the literature. For any of these demand functions, $\mathcal{S}(p(z))$ is convex in z, and consequently $\mathbb{E}[\mathcal{S}(p(Z))] \geq \mathcal{S}(p(\mathbb{E}[Z]))$. There are cases, where $p(z)$ is increasing convex and yet $\mathcal{S}(p(z))$ is still convex in z provided that $p(z)$ is not "too" convex.

Example 8.8 Suppose that $n = 1$, and $U(q) = q - q^2/200$, then $d(p) = 100(1-p)$ over $p \in [0, 1]$, $p(z) = 0.5(1 + z)$ for $z \in [0, 1]$ and $\mathcal{S}(p(z)) = 12.5(1 - z)^2$. Assume again, as in Example 8.3, that $\mathbb{P}(Z = 1/3) = \mathbb{P}(Z = 2/3) = 1/2$. By Jensen's inequality, we have $\$3.47 = \mathbb{E}[\mathcal{S}(p(z))] \geq \mathcal{S}(p(\mathbb{E}[Z])) = \$3.125$, so consumers are better off by 11.1% when prices are dynamic and driven by Z compared to static prices $p(\mathbb{E}[Z])$. From Example 8.3, we see that the firm is also 11.1% better off using dynamic pricing, resulting in a win-win situation.

8.4 Finite Capacity

Finite capacity is a central theme for both revenue management and dynamic pricing. We will assume that $d(p) : \Re^n_+ \to \Re^n_+$ is continuous function in p and that variable costs are zero. Let c be an m-dimensional vector of resources available for the n products. Let A be an $m \times n$ matrix where the j-th column, say A_j, is the vector of resources consumed by each unit of product j. For $X = \Re^n_+$, the optimal revenue as a function of c is given by

$$\bar{V}(c) := \max_{p \in X} R(p, 0) \text{ subject to } A\,d(p) \leq c. \tag{8.3}$$

The interpretation is that there is a sunk investment in capacity c, the firm wants to maximize the revenue that can be obtained from this capacity, and no variable costs are incurred. The objective function of problem (8.3) may not be concave and the constraint set may not be convex, which makes solving problem (8.3) potentially difficult. There are two techniques that we can use to try to solve problem (8.3). First, we can work with the inverse demand function $p(q)$ assuming it exists. In this case, problem (8.3) can be written as maximizing $p(q)'q$ subject to $A\,q \leq c, q \geq 0$. Now the constraint set is convex and if the objective function is concave, or quasi-concave, then standard techniques such as the KKT conditions and its extensions can be used to solve the problem. The second technique is based on Lagrangian relaxation, as outlined next.

8.4.1 Lagrangian Relaxation

We will explore the Lagrangian relaxation approach on problem (8.3) under the assumption that $R(p,w) := (p-w)'d(p)$ has a finite maximizer $p(w)$ for all $w \in \Re^n_+$ with $d(p(w))$ continuous in w. Although these are strong conditions, they turn out to hold for many important applications. A sufficient condition for the existence of a finite maximizer is that $R(p,w)$ is upper semi-continuous (USC)[2] in p, and that for some $\alpha(w) \in \Re_+$, the upper contour set $\{p : R(p,w) \geq \alpha(w)\}$ is non-empty and compact. In this case, by the extreme value theorem (EVT),[3] the function $R(p,w)$ achieves its maximum over the compact upper contour set. Sharper conditions for the existence of $p(w)$ will be presented later.

Let

$$L(p,z) := R(p,0) + z'(c - Ad(p)) = R(p, A'z) + z'c$$

be the Lagrangian corresponding to the dual vector $z \in \Re^m_+$. The Lagrangian program is $\min_{z\geq 0}\max_{p\geq 0} L(p,z)$. The inner optimization yields

$$L(z) := \max_{p\geq 0} L(p,z) = \mathcal{R}(A'z) + z'c,$$

where we have taken advantage of the assumption that there exists a price, say $p(A'z)$, that maximizes $R(p, A'z)$, and $\mathcal{R}(A'z) = R(p(A'z), A'z)$.

By weak duality $L(z) \geq \bar{V}(c)$ for all $z \geq 0$. The Lagrangian dual problem is

$$\Gamma(c) := \min_{z\geq 0} L(z) = \min_{z\geq 0}\left[\mathcal{R}(A'z) + z'c\right] \geq \bar{V}(c),$$

which is a convex minimization problem in z subject to non-negativity constraints $z \in \Re^m_+$.

If $c \geq A\,d(p(0))$, then $z(c) = 0 \in \Re^m$ is optimal. This follows because $p(0)$ is feasible and maximizes $R(p,0)$. Otherwise $z(c)$ has at least one positive component. We next investigate conditions under which there is no duality gap and we can assert that $\bar{V}(c) = \Gamma(c)$. If $p(A'z(c))$ satisfies the capacity constraint, and the complementary slackness condition $z(c)'(c - A\,d(p(A'z(c)))) = 0$ holds, then there is no duality gap, since then $L(z(c)) = R(p(A'z(c)), 0) = \bar{V}(c)$, and consequently $p(A'z(c))$ is an optimal solution to problem (8.3). We summarize this result in the following proposition.

---

[2]A function $f : X \to [-\infty, \infty]$ is upper semi-continuous if and only if $\{x \in X : f(x) \geq a\}$ is closed for every $a \in \Re$.

[3]The EVT is also known as the Bolzano-Weierstrass theorem.

Proposition 8.9 *Assume that $R(p, A'z)$ has a solution $p(A'z)$ for any $z \geq 0$, and let $z(c) \geq 0$ solve the convex optimization problem* $\min_{z \geq 0}[\mathcal{R}(A'z) + z'c]$. *If $A\,d(p(A'z(c))) \leq c$ and $z(c)'[c - A\,d(p(A'z(c)))] = 0$, then $p(A'z(c))$ is a solution to problem* (8.3).

8.4.2 Finite Capacity and Finite Sales Horizon

Another central theme for dynamic pricing and revenue management is the existence of a finite sales horizon over which the products can be sold. Let t be the time-to-go, and assume a sales horizon of length T. At the end of the sales horizon, no further sales are possible. Let c be the initial inventory, and assume that inventory replenishments are not possible. This situation is typical in fashion retailing and in revenue management applications. We will assume that the demand *rate* $d_t(p)$ at price p at time-to-go t is continuous in p for all $t \in [0, T]$. The profit contribution over the sales horizon from using price path $p_t, t \in [0, T]$ is given by $\int_0^T R_t(p_t, 0)dt := \int_0^T p_t' d_t(p_t)dt$. The optimal revenue as a function of T and c is given by

$$\bar{V}(T, c) := \int_0^T \max_{p_t \in X} R_t(p_t, 0)dt \tag{8.4}$$
$$\text{s.t.} \quad \int_0^T A d_t(p_t)dt \leq c,$$

where as before $X = \Re_+^n$ and $A \geq 0$ is an $m \times n$ matrix representing the consumption of resources by products.

The Lagrangian penalizes component shortfalls at rate $z \in \Re_+^m$. We will assume that there exists a price $p_t(A'z)$ that maximizes $R_t(p, A'z)$ for every $t \in [0, T]$. The inner optimization of the Lagrangian function yields $\int_0^T \mathcal{R}_t(A'z) + z'c$, so the outer optimization is given by

$$\Gamma(T, c) := \min_{z \geq 0} \left[\int_0^T \mathcal{R}_t(A'z)dt + z'c \right] \geq \bar{V}(T, c) \tag{8.5}$$

whose objective function is convex in z. Let $z(T, c)$ be the optimal solution to this convex optimization problem. If the price path $p_t(z(T, c)), t \in [0, T]$ is feasible and the complementary slackness condition

$$z(T, c)' \left[c - A \int_0^T d_t(A' p_t(z(T, c))dt \right] = 0$$

holds, then there is no duality gap and the price path $p_t(z(T, c)), t \in [0, T]$ is an optimal solution to problem (8.4), turning the inequality in (8.5) into an equality.

It is instructive to compare formulation (8.4) to a formulation based on aggregate demand $D(p) = \int_0^T d_t(p)dt$ that yields

$$\bar{V}_f(T, c) := \max_{p \in X} p' D(p) \quad \text{s.t.} \quad A\, D(p) \leq c, \tag{8.6}$$

Notice that in formulation (8.6) we seek a single vector of prices that maximizes revenue over the entire horizon subject to an aggregate capacity constraint. Clearly $\bar{V}(T, c) \geq \bar{V}_f(T, c)$, as the ability to respond to changes in demand $d_t(p)$ over the sales horizon $t \in [0, T]$ gives formulation (8.4) an important advantage over (8.6). Of course, this advantage can materialize only if consumers that arrive at time-to-go t either purchase at $p_t(z(T, c))$ or leave the system. This model assumes that the firm can do price discrimination over time. The model may break down when consumers are strategic and they face no disutility from waiting for a lower price, except when prices $p_t(z(T, c))$ are monotone so there is no incentive for waiting. In retailing, for example, some consumers are strategic and prefer to wait for lower prices, but they are exposed to rationing risks and the disutility of waiting.

8.5 Single Product Pricing Problems

In this section, we investigate issues of existence and uniqueness for single product pricing problems. We next study real options and bargaining as mechanisms to improve profits and reduce the dead weight loss. We end this section with a look at multiple market segments and direct price discrimination.

8.5.1 Existence and Uniqueness

For the single product case with $n = 1$, $d(p)$ is the demand for the single product at price $p \geq 0$. We seek sufficient conditions for the existence of a finite maximizer of $R(p, z) = (p-z)\, d(p)$ over $p \in \Re_+$. Let $\bar{d}(p) := \sup_{\tilde{p} \geq p} d(\tilde{p})$. Notice that $\bar{d}(p) \geq d(p)$ is a decreasing function even if $d(p)$ is not. Let $\bar{R}(p, z) := (p - z)\, \bar{d}(p)$. We next show that if $d(p)$ is USC and $p\bar{d}(p) \to 0$ as $p \to \infty$ (so $\bar{d}(p) = o(1/p)$), then $\bar{R}(p, z)$ has a finite maximizer $p(z)$ that also maximizes $R(p, z)$.

Theorem 8.10 *If $d(p)$ is USC in $p \geq 0$, and $\bar{d}(p) = o(1/p)$, then there exists a finite maximizer $p(z)$, increasing in $z \geq 0$, that simultaneously maximizes $R(p, z)$ and $\bar{R}(p, z)$, so $\mathcal{R}(z) = \bar{\mathcal{R}}(z)$.*

A formal proof of Theorem 8.10 is in the appendix. Notice that Theorem 8.10 does not require $d(p)$ to be decreasing or eventually decreasing in p. While the conditions of Theorem 8.10 may seem technical, they imply the existence of finite maximizers for pricing problems that are typically encountered in practice. For example, if we have a finite population of λ potential consumers with independent and identically distributed (IID) willingness to pay Ω, then the expected demand at price p is $d(p) = \lambda P(\Omega \geq p)$. Then $d(p)$ is USC, and if $E[\Omega^+] < \infty$, then $d(p) = o(1/p)$, so there exist a maximizers $p(z)$ of $R(p, z)$. As an example, assume that Ω is exponential with mean θ. Then $d(p) = \lambda e^{-p/\theta}$, and $p(z) = z + \theta$ maximizes $R(p, z)$, so $\mathcal{R}(z) = \theta\lambda e^{-1-z/\theta}$.

We now turn to conditions on the demand function $d(p)$ that guarantee that $R(p, z)$ does not have local, non-global, maximizers or more succinctly that $R(p, z)$ is unimodal in $p \geq z$. This is equivalent to $R(p, z)$ being quasi-concave in $p \geq z$ and to $R(p, z)$ having convex upper level sets: $\{p \geq z : R(p, z) \geq \alpha\}$ for all $\alpha \geq 0$. If $d(p)$ is continuous and differentiable, then we define the *hazard rate* at p to be $h(p) := -d'(p)/d(p)$ where $d'(p)$ is the derivative of $d(p)$ at p. The hazard rate function $h(p)$ is defined for all $p < \bar{r}$, where $\bar{r}$ is the choke-off price. The hazard rate is the event rate at price p, conditional on $\Omega \geq p$. Taking the derivative of $R(p, z)$ with respect to p leads to first-order condition for optimality:

$$f(p, z) = 1 - (p - z)h(p) = 0.$$

Let $p(z)$ be a root of $f(p, z) = 0$. Then, $p(z)$ is a maximizer of $R(p, z)$, and $R(p, z)$ is quasi-concave if $f(p, z)$ is non-negative for all $p < p(z)$ and non-positive for all $p > p(z)$. The following result provides conditions on the hazard rate that guarantee the existence and uniqueness of a finite maximizer $p(z)$, as well as some results about the optimal mark-up $\Delta(z) := p(z) - z$.

Theorem 8.11

(a) *If $h(p)$ is continuous and increasing in p and $h(z) > 0$, then there is a unique optimal price $p(z)$, strictly increasing in z, satisfying $z < p(z) \leq z + 1/h(z)$, with $\Delta(z) = p(z) - z$ decreasing in z. The upper bound is attained by the exponential demand function.*

(b) *If $ph(p)$ is continuous and strictly increasing in p and there exists a finite $\tilde{z} \geq z > 0$ such that $1 < \tilde{z}h(\tilde{z})$, then there is a unique optimal price $p(z)$, strictly increasing in z, satisfying $z < p(z) \leq z/(1 - 1/\tilde{z}h(\tilde{z}))$. The upper bound is attained by the constant elasticity of demand function. Moreover, if $1/h(z)$ is concave in z, then $p(z)$ is concave in z.*

(c) *If $\tilde{d}(p)$ is a demand function with hazard rate $\tilde{h}(p)$ and $\tilde{h}(p) \geq h(p)$ for all p, then $\tilde{p}(z) \leq p(z)$.*

The condition $ph(p)$ increasing in p is weaker than $h(p)$ increasing in p, and leads to weaker results as we cannot claim that $\Delta(z)$ is decreasing in z. As an example, for the constant elasticity of demand model, $d(p) = \lambda p^{-b}$, $b > 1$, we have $\Delta(z) = z/(b - 1)$, which is increasing in z.

Economists often write the solution to the first-order condition $f(p, z) = 0$ in terms of the (absolute) price *elasticity of demand* $e(p) := -pd'(p)/d(p) = ph(p)$ resulting in

$$p(z) = \frac{e(p(z))}{e(p(z)) - 1} z.$$

This formula suggests that the mark-up on marginal cost should be equal to $e(p(z))/(e(p(z)) - 1)$. Notice that both the left and the right hand sides depend on $p(z)$ except for the constant elasticity demand model, so the mark-up interpretation needs to be taken with a grain of salt. Nevertheless, this mark-up formula provides some guidelines that link elasticities to prices via the mark-up on marginal costs.

The solution to the first-order condition is sometimes written as

$$\frac{\Delta(z)}{p(z)} = \frac{1}{e(p(z))},$$

with the left hand side known as the Lerner index, so the Lerner index is equal to one over the elasticity of demand. If $z = 0$, then $\Delta(z) = p(z)$, so $e(p(0)) = 1$. The last equation is often written as

$$p(z) = z + 1/h(p(z)).$$

It can be shown that if $1/h(p)$ is concave (respectively, convex) in p then $p(z)$ is increasing concave (respectively, convex) in z.

The problem of maximizing $R(p, z)$ can sometimes be transformed so that demand rather than price is the decision variable. This can be done if there is an inverse demand function, say $\tilde{p}(q)$, that yields demand $q \leq d(0)$ at price $\tilde{p}(q)$. The problem is to maximize $(\tilde{p}(q) - z)q$ over $q \geq 0$. It can be shown that the concavity of $\tilde{p}(q)q$ in q is equivalent to the convexity of $1/d(p)$ in p, so from this we surmise that another sufficient condition for $R(p, z)$ to be quasi-concave in p is that $1/d(p)$ is convex. A weaker condition for the quasi-concavity of $(\tilde{p}(q) - z)q$ is that $\tilde{p}(q)$ is log-concave in q. It is interesting to note that there are demand functions for which $R(p, z)$ is concave in p without $(\tilde{p}(q) - z)q$ being concave in q.

8.5.2 Priority Pricing

Consider the finite capacity problem for the single product case. We will assume that there is a unique solution, say $p(z(c))$, to the problem of maximizing $R(p, 0)$ subject to $d(p) \leq c$, where $z(c)$ is the dual variable associated with the capacity constraint and that $d(p)$ is continuous in p. Let $\bar{c}$ be the smallest integer at which the dual variable is zero. Then $p(z(c)) = p(0)$ for all $c \geq \bar{c}$.

Suppose that capacity is a random variable, say C, and that the firm prices at $p(z(C))$. Since the price is the same for all $C \geq \bar{c}$, it is convenient to redefine C to be $\min(C, \bar{c})$, so its support is in $\{0, 1, \ldots, \bar{c}\}$. With this notation, the expected profit to the firm is equal to

$$\mathbb{E}[Cp(z(C))] = \sum_{c=1}^{\bar{c}} cp(z(c))\mathbb{P}(C = c).$$

Changing the order of summation, we see that the average price paid for the cth unit of capacity is equal to $\mathbb{E}[p(z(C))|C \geq c]$. This pricing policy applies to situations where yields are random, and the firm can pass the price signal $p(z(C))$ to consumers who select whether or not they want to buy at that price. The policy is somewhat controversial as it calls for the disposal of capacity when yields are high and can be perceived as price gouging when yields are low. In some instances, such as the consumption of power, consumers cannot react to the changes in capacity in real time. Therefore, the application of this scheme requires a priority matching to consumers who value the service the most, and this is why this is called a priority pricing schedule.

8.5.3 Social Planning and Dead Weight Loss

A social planner is interested in selecting p to maximize the sum of the consumers' surplus $\mathcal{S}(p)$ and the firm's profit $R(p, z)$. The sum of these two quantities is known as the *social welfare* function, given by

$$W(p, z) := \mathcal{S}(p) + R(p, z).$$

Optimizing over p, we obtain the *optimal welfare* function

$$\mathcal{W}(z) := \max_{p \geq z} W(p, z).$$

Proposition 8.12 *If $d(p)$ is differentiable and decreasing in p, then $\mathcal{W}(z) = \mathcal{S}(z)$ is decreasing convex in z.*

The result follows because under the stated conditions $\frac{\partial W(p,z)}{\partial p} = (p-z)d'(p) \leq 0$, so social welfare is decreasing in p and its maximum is attained at $p = z$.

The difference $\mathcal{W}(z) - W(p(z), z)$ is known as the *dead weight loss*. It reflects the difference between the optimal social welfare and the social welfare that results when the firm maximizes its profits. As an example, if $d(p) = \lambda e^{-p/\theta}$, then $p(z) = z + \theta$, $\mathcal{W}(z) = \mathcal{S}(z) = \lambda\theta e^{-z/\theta}$, while $W(p(z), z) = 2\lambda\theta e^{-1}e^{-z/\theta}$, so the dead weight loss is equal to $[1 - 2e^{-1}]\mathcal{W}(z)$ or 26% of the maximum social welfare.

Trying to reduce the dead weight loss is difficult because the optimal solution to the social planner's problem is to set $p = z$ and this results in zero profits for the firm with all of the benefits going to the consumers. We will next explore two cases where the dead weight loss can either be eliminated or reduced. The first case requires the use of real options on services when the booking and the consumption of the service are separated by time and consumers are uncertain about their valuations at the time of booking. The second case corresponds to the situation where the consumers and the firm negotiate instead of using a take it or leave it price.

Call Options on Capacity

Consider first the case of a homogeneous group of consumers booking capacity in advance of consumption. Suppose there are λ consumers, each with independent and identically distributed random willingness-to-pay for the service at the time of consumption. We assume that the distribution $H(p) = \mathbb{P}(\Omega \geq p)$ is common knowledge to consumers and the firm, so the aggregate demand function is $d(p) = \lambda\,\mathbb{P}(\Omega \geq p)$. For this model, the surplus function is $\mathcal{S}(p) = \lambda\mathbb{E}[(\Omega - p)^+]$. We further assume that consumers do not learn the realization of demand until the time of consumption. Under these conditions, the firm can benefit from offering call options to consumers. A call option requires an upfront, non-refundable, payment x that gives the customer the non-transferable right to buy one unit of the service at price p at the time of consumption. The special case where $p = 0$ is called advanced selling, and the case $x = 0$ is called spot selling.

Consumers evaluate call options by the surplus they provide. A customer who buys an (x, p) option will exercise his right to purchase one unit of the service at the time of consumption if and only if $\Omega \geq p$. By doing this, an individual consumer obtains expected surplus $s(p) = \mathbb{E}[(\Omega - p)^+] = \mathcal{S}(p)/\lambda$. Since consumers pay x for this right, the consumer receives expected surplus $s(p) - x$ and would find the (x, p) option attractive only if $s(p) - x \geq 0$.

Consider the problem of maximizing the expected profit from selling (x, p) options subject to the participating constraint $s(p) - x \geq \tilde{s}$, where $\tilde{s} \geq 0$ is a lower bound on the individual surplus that needs to be given to consumers to induce them to buy the option. In practice, the firm may set $\tilde{s} = 0$ to extract as much surplus from consumers. Here we will analyze the problem for other values of $\tilde{s}$ to show that it is possible to eliminate the dead weight loss and use $\tilde{s}$ as a mechanism to distribute profits and surplus between the firm and the consumers.

Since the expected profit from selling (x, p) options that satisfy the participating constraint is $x + (p - z)\,\mathbb{P}(\Omega \geq p)$ and there are λ consumers, the expected profits are equal to $\lambda x + (p - z)\lambda\,\mathbb{P}(\Omega \geq p) = \lambda x + R(p, z)$. This is a function of x, and it is optimal to set $x^* = s(p) - \tilde{s}$. This reduces the problem to that of maximizing $\lambda(s(p) - \tilde{s}) + R(p, z) = \mathcal{S}(p) + R(p, z) - \lambda\tilde{s} = W(p, z) - \lambda\tilde{s}$ with respect to p. We already know that $W(p, z)$ is maximized at $p = z$. Thus, the solution to the provider's problem is to set $p = z$ and $x = s(z) - \tilde{s}$, so the provider obtains profits equal to $\lambda x^* = \lambda(s(z) - \tilde{s})$, while consumers receive surplus $\lambda\tilde{s}$. Since the sum of these two quantities is $\mathcal{S}(z) = \lambda s(z)$, the selling of call options eliminates the dead weight loss and $\tilde{s}$ can be used as a mechanism to distribute the dead weight loss.

We now explore the range of values of $\tilde{s}$ that guarantees that both the firm and the consumers are at least as well off as the solution $(x, p) = (0, p(z))$, where price $p(z)$ is offered to consumers after they know their valuations. Under this scheme, the firm makes $\mathcal{R}(z)$ and consumers receive surplus $\mathcal{S}(p(z)) = \lambda s(p(z))$. As a result, consumers are better off whenever $\lambda\tilde{s} \geq \mathcal{S}(p(z))$, while the firm is better off whenever $\mathcal{S}(z) - \lambda\tilde{s} \geq \mathcal{R}(z)$, so a win-win is achieved for any value of $\tilde{s}$ such that $\mathcal{S}(p(z)) \leq \lambda\tilde{s} \leq \mathcal{S}(z) - \mathcal{R}(z)$. Since $\mathcal{S}(z) \geq \mathcal{R}(z) + \mathcal{S}(p(z))$, the win-win interval is non-empty. In practice, absent competition or an external regulator, the provider may simply select $\tilde{s} = 0$, to improve his profits from $\mathcal{R}(z)$ to $\mathcal{W}(z)$ extracting all consumer surplus while also capturing the dead weight loss. The improvement in profits from options can be very significant. Indeed, in the exponential case, $(\mathcal{W}(z) - \mathcal{R}(z))/\mathcal{R}(z) = (e - 1) = 172\%$.

The idea of using call options can be extended to the case where the variable cost Z of providing the service at the time of consumption is random. In this case, the option is designed by setting $x = \mathbb{E}[s(Z)] - \tilde{s}$ and $p = Z$, so that by paying x in advance the option bearer has the right to purchase one unit of the service at the random variable cost Z.

Bargaining Power

Assume again that demand comes from λ homogeneous consumers with willingness to pay Ω, so $d(p) = \lambda\,\mathbb{P}(\Omega \geq p)$. Without negotiation, the firm sets the price at $p(z)$ and consumers make a purchase if $\Omega \geq p(z)$ and leave the system otherwise. In this section, we will show that the dead weight loss can be reduced if the firm and the consumers negotiate instead of using non-negotiable prices. Suppose that consumers know the realization of their willingness to pay, but the firm knows only the distribution of Ω. We will assume that the firm has a reservation price, say p, under which it is not willing to sell. We will assume the firm or an agent for the firm negotiates with each customer. If $\Omega < p$, then no sale takes place, but if $\Omega \geq p$, we will assume that a sale takes place at the Bargaining Nash Equilibrium (BNE) price $\beta\Omega + (1 - \beta)p$, where $\beta \in [0, 1]$ is the negotiating power of the firm and $1 - \beta$ is the negotiating power of the buyers. Notice that if $\Omega \geq p$, then transaction takes place at the reservation price p, when $\beta = 0$, and at Ω, when $\beta = 1$.

The problem for the firm is to select the reservation price, say $p_\beta(z)$, to maximize expected profits taking into account both the unit cost z and the negotiating power β. Let $\delta(\Omega - p)$ be a random variable taking value 1 if $\Omega \geq p$ and 0 otherwise. The firm wants to select the reservation price p to maximize

$$\begin{aligned} R_\beta(p, z) &:= \lambda\,\mathbb{E}[(\beta\Omega + (1 - \beta)p - z)\delta(\Omega - p)] \\ &= \lambda\,\mathbb{E}[\beta(\Omega - p)\,\delta(\Omega - p)] + \lambda(p - z)d(p) \\ &= \beta\lambda\,\mathbb{E}[(\Omega - p)^+] + R(p, z) \\ &= \beta\mathcal{S}(p) + R(p, z). \end{aligned} \tag{8.7}$$

Let

$$\mathcal{R}_\beta(z) := \max_p R_\beta(p, z).$$

Let $h(p)$ be the hazard rate of $d(p) = \lambda\,\mathbb{P}(\Omega \geq p)$. If $ph(p)$ is increasing in p, then the maximizer of $R_\beta(p, z)$, say $p_\beta(z)$, is the unique root of the equation $(p - z)h(p) = 1 - \beta$. It is easy to see that $p_\beta(z)$ is decreasing in β and increasing in z, while $\mathcal{R}_\beta(z)$ is increasing in β and decreasing in z. By substituting $p_\beta(z)$ into the formula for $R_\beta(p, z)$, we obtain

$$\mathcal{R}_\beta(z) = \beta\mathcal{S}(p_\beta(z)) + R(p_\beta(z), z).$$

At $\beta = 0$, we have $p_0(z) = p(z)$ and $\mathcal{R}_0(z) = \mathcal{R}(z)$. Consequently, pricing at $p(z)$ is tantamount to assuming that the firm has no negotiating power, or equivalently relinquishing the negotiating power. This may be done for expediency for relatively inexpensive goods that are sold in high volumes. At $\beta = 1$, $p_1(z) = z$, so $\mathcal{R}_1(z) = \mathcal{W}(z)$, eliminating all of the dead weight loss, with the firm extracting all of the consumers' surplus. In most cases, $\beta \in (0, 1)$, so it makes sense for the firm to negotiate with consumers for goods that are expensive and sold in relatively low volumes. Indeed, prices for real estate, cars, art, and high-end services are often negotiated, while those of groceries are typically not except in economies where people have more time than money.

The consumers' expected surplus is given by

$$\begin{aligned}
\mathcal{S}_\beta(p_\beta(z)) &= \lambda\,\mathbb{E}[(\Omega - \beta\Omega - (1-\beta)p_\beta(z))\,\delta(\Omega - p_\beta(z))] \\
&= \lambda(1-\beta)\,\mathbb{E}[(\Omega - p_\beta(z))^+] \\
&= (1-\beta)\,\mathcal{S}(p_\beta(z)). \qquad (8.8)
\end{aligned}$$

It is easy to see that the consumers' surplus is decreasing in β, so some of the benefits that the firm derives from negotiation comes from smaller surplus for consumers.

If we now add (8.7) and (8.8), and evaluate it at $p_\beta(z)$, we see that the social welfare that results from negotiation is equal to

$$W_\beta(p_\beta(z), z) := \mathcal{S}_\beta(p_\beta(z)) + \mathcal{R}_\beta(z) = \mathcal{S}(p_\beta(z)) + R(p_\beta(z), z).$$

This quantity is increasing in β. This follows because $W(p, z)$ is decreasing in p and $p_\beta(z)$ is decreasing in β. This implies that the firm makes more than the loss to the consumers when it has negotiating power.

8.5.4 Multiple Market Segments

Suppose that there are multiple market segments with independent demands $d_m(p), m \in \mathcal{M} := \{1, \ldots, M\}$ for $p \in \Re_+$ for a product. We will assume throughout this section that $d_m(p)$ satisfies the conditions of Theorem 8.10 for every $m \in \mathcal{M}$. This guarantees that there exists a $p_m(z)$ increasing in z that maximizes $R_m(p, z) := (p - z)d_m(p)$. If the firm can use direct price discrimination (also known as third degree price discrimination or personalized pricing), then it would use price $p_m(z)$ for market segment $m \in \mathcal{M}$. The possibility to use direct price discrimination arises when it is possible to vary price by time, location, or customer attributes. This is often true for services and less so for physical products as there may be a gray market which creates demand dependencies.

In some cases, we may need to offer the same price for a subset $S \subset \mathcal{M}$ of market segments. This may be due to regulations or if the markets are not sufficiently different. Let $d_S(p) := \sum_{m \in S} d_m(p)$ denote the aggregate demand over market segments in S at price $p \in \Re_+$, and let $R_S(p, z) := (p - z)d_S(p)$ denote the profit function for market segments in S when the variable cost is z. We seek conditions for the existence of a maximizer $p_S(z)$ of $R_S(p, z)$ that is in the convex hull of the set $\{p_m(z) : m \in S\}$.

The following result shows that $d_S(p)$ inherits some desirable properties from the individual market demand functions $d_m(p), m \in S$.

Proposition 8.13 *If $d_m(p)$ satisfies the conditions of Theorem 8.10 for every $m \in \mathcal{M}$, then so does $d_S(p)$. Moreover, there exists a finite price $p_S(z)$, increasing in z, such that $\mathcal{R}_S(z) = R_S(p_S(z), z)$ is decreasing convex in z.*

It may be tempting to conclude that, under the conditions of Proposition 8.13, $p_S(z)$ would lie in the convex hull of $\{p_m(z), m \in S\}$. Example 8.14 shows that this is not true.

Example 8.14 Suppose that $d_1(p) = 1$ for $p \leq 10$ and $d_1(p) = 0$ for $p > 10$. Then $R_1(p, 0)$ is maximized at $p_1(0) = 10$ and $\mathcal{R}_1(0) = 10$. Suppose that $d_2(p) = 1$ for $p \leq 9$, $d_2(p) = 0.1$ for $9 < p \leq 99$ and $d_2(p) = 0$ for $p > 99$. Then $R_2(p, 0)$ is maximized at $p_2(0) = 99$ resulting in $\mathcal{R}_2(0) = 9.9$. The total profit is equal to 19.9 if each segment is allowed to be priced separately. Let $S = \{1, 2\}$, then $R_S(p, 0) = R_1(p, 0) + R_2(p, 0)$ is maximized at $p_S(0) = 9 < \min_{i \in S} p_i(0)$ resulting in $\mathcal{R}_S(0) = 18$.

Since the sum of quasi-concave functions is not, in general, quasi-concave, it should not be surprising that properties of $d_m(p)$ that imply quasi-concavity of $R_m(p, z)$, for each $m \in \mathcal{M}$ are not, in general, inherited by $d_S(p) = \sum_{m \in S} d_m(p)$. Example 8.15 illustrates this.

Example 8.15

(a) Suppose that $d_m(p) = \exp(-p/b_m)$ for $m = 1, 2$ with $b_1 < b_2$. Then the hazard rate $h_m(p) = 1/b_m$, is constant, and there is a unique price $p_m(z) =$

$z + b_m$ that maximizes $R_m(p, z)$ for $m = 1, 2$. Let $S = \{1, 2\}$. The hazard rate $h_S(p)$ of $d_S(p)$ is decreasing in p.

(b) Suppose that $d_m(p) = 1/p^{b_m}$ for some $b_m > 1$, then $ph_m(p) = b_m$ and there is a unique price $p_m(z) = b_m z/(b_m - 1)$ that maximizes $R_m(p, z)$ for $m = 1, 2$. Let $S = \{1, 2\}$. The proportional hazard rate $ph_S(p)$ of $d_S(p)$ is decreasing in p.

In both cases in Example 8.15, the profit function $R_S(p, z)$ is actually quasi-concave, even if the aggregate demand function $d_S(p)$ has decreasing hazard rate (Part a) or decreasing proportional hazard rate (Part b). The next result provides sufficient conditions to bound the maximizer of $R_S(p, z)$ to be within the convex hull of $p_m(z), m \in S$.

Proposition 8.16 *Assume that $d_m(p)$ satisfies the conditions of Theorem 8.10 for each $m \in M$, that the hazard rate $h_m(p)$ is continuous in p, and that $ph_m(p)$ is increasing in p for each $m \in M$. Then, $R_S(p, z)$ has a maximizer in the convex hull of $\{p_m(z), m \in S\}$ for all $S \subset \mathcal{M}$.*

Corollary 8.17 *Proposition 8.16 holds if $h_m(p)$ is increasing in p for all $m \in S$.*

We next consider the problem where we are allowed a price menu that consist of at most $J \le M$ different prices. The limitation to J prices may be managerial in nature, or it may be due to the lack of precise knowledge of the demand parameters for some of the market segments. The extreme cases are $J = 1$, where a single price is used for all the segments (so there is no price discrimination) and $J = M$, where each segment is priced independently (full direct price discrimination). Let $Q_J(z)$ be the maximum profit from using J distinct prices for the M market segments when the marginal cost is z. For $J = 1$, we have $Q_1(z) = R_{\mathcal{M}}(z)$, and for $J = M$, we have $Q_M(z) = \sum_{m\in\mathcal{M}} \mathcal{R}_m(z)$. Clearly $Q_J(z)$ is increasing in J. For $1 < J < M$, the problem is combinatorial in nature, as we need to assign M market segments into J market clusters, with all market segments in a cluster using the same price.

Our aim in this section is to develop a heuristic and a lower bound on the profitability of using J prices. More precisely, we will develop a heuristic with profit $Q^h_J(z)$ such that

$$\frac{Q_J(z)}{Q_M(z)} \ge \frac{Q^h_J(z)}{Q_M(z)} \ge \gamma_J(z),$$

for some function $\gamma_J(z)$ for situations where all of the demand functions $d_m(p), m \in \mathcal{M}$ belong to the same family. As we shall see, it is often possible to obtain most of the potential profits with a relatively small J even if we do not have detailed knowledge of the demand functions.

We will assume that the demand functions $d_m(p), m \in \mathcal{M}$ belong to the same family. By this we mean that $d_m(p) = \lambda_m H_m(p), m \in \mathcal{M}$ and the tail distributions $H_m(p) = P(\Omega_m \ge p), m \in \mathcal{M}$ differ only on their parameters. Examples of families of demand functions include linear, log-linear, CES, logit, among others. We will assume that the profit function $R_m(p, z) = (p - z)d_m(p)$ is quasi-concave for each m and that there is a unique finite maximizer $p_m(z)$ for each $m \in \mathcal{M}$. We will assume that the market segments are ordered so that $p_1(z) \leq \ldots \leq p_M(z)$. Finally, we will assume that for any $S \subset \mathcal{M}$, the profit function $R_S(p, z) = \sum_{m \in S} R_m(p, z)$ has a finite maximizer $p_S(z)$ in the interval $[\min_{m \in S} p_m(z), \max_{m \in S} p_m(z)]$, as guaranteed under the conditions of Proposition 8.16.

Since we will be using heuristic prices, it is convenient to have a measure of how efficient it is to use price p instead of $p_m(z)$ for market segment m. This motivates defining the relative efficiency of using price p instead of price $p_m(z)$ for market segment m as the ratio

$$e_m(p, p_m(z), z) := \frac{R_m(p, z)}{\mathcal{R}_m(z)} \leq 1. \tag{8.9}$$

Notice that $e_m(p, p_m(z), z)$ reaches maximum efficiency at $p = p_m(z)$ and decays on both directions as a result of our quasi-concavity assumption. We will be particularly interested in families of demands for which $e_m(p, p_m(z), z)$ is independent of m. This is true for the linear, the log-linear, and the logit demand functions, among others. It is possible to find closed-form formulas for $e(p, p_m(z), z)$ for many families of demand functions including linear, log-linear, and CES. However, there are distributions that do not admit closed-form expressions for $e(p, p_m(z), z)$ but the results that we will derive here can also be applied, numerically, to distributions that do not admit closed-form expressions. The relative efficiencies of prices will help us deal with situations where we may not know the exact parameters of some of the market segments. On occasions, we will write $e(q, s, z)$ to mean the efficiency of price $q \neq s$ for a (possibly fictitious) market segment for which price s is optimal for z.

We now show how to construct a heuristic that uses $1 < J < M$ prices. The idea is to break down the interval $[p_1(z), p_M(z)]$ into J sub-intervals, which in turn determine market clusters and then to use a common price for all the market segments within a cluster. The precise price used within a cluster will depend on the detailed knowledge of the market segments in a cluster. If only limited information is known, then a *robust* price that maximizes the minimum efficiency will be used, otherwise an *optimal* price all market segments in the cluster will be used.

We start by describing the procedure by showing how to select the break-points and robust prices and later explain how the heuristic can be improved with optimal prices in each cluster.

Consider arbitrary break-points $p_1(z) = s_0 < s_1 < s_2 \ldots < s_{J-1} < s_J = p_M(z)$ and define market clusters $M_j = \{m : p_m(z) \in [s_{j-1}, s_j)\}$ for $j = 1, \ldots, J-1$ and $M_J = \{m : p_m(z) \in [s_{J-1}, s_J]\}$. Let $q_j \in (s_{j-1}, s_j)$ be a common price to be used for all markets in cluster M_j, $j = 1, \ldots, J$. The break-points $s_1, \ldots, s_{J-1}$ and the prices q_j are designed to maximize the minimum efficiency among all of the market segments. More precisely, the s_j's and q_j's are selected so that

$$e(q_j, s_{j-1}, z) = e(q_j, s_j, z) \quad \text{for all} \quad j = 1, \ldots, J \tag{8.10}$$

and

$$e(q_1, s_1, z) = e(q_2, s_2, z) = \ldots = e(q_J, s_J, z). \tag{8.11}$$

Equation (8.10) guarantees that price q_j is just as efficient for s_{j-1} as it is for s_j. Equation (8.11) guarantees that the efficiency of q_j relative to s_j is the same for each market segment. This implies that for any market segment $m \in M_j$, $e(q_j, p_m(z), z) \geq e(q_j, s_j, z)$ for all $j = 1, \ldots, J$.

It is often possible to find the s_j's and the q_j's with very limited information about the market prices $p_m(z)$. Usually, it is sufficient to know the smallest $p_1(z)$ and the largest $p_M(z)$ prices.

Let $Q^h_J(z)$ be the profit obtained by pricing market all market segments in M_j at q_j for all $j = 1, \ldots, J$. Notice that we assign market m to price q_j if j maximizes $e(q_j, p_m(z), z)$, or equivalently $p_m(z)$ is in the interval defined by s_{j-1} and s_j. Thus, relatively little knowledge about the markets is need to implement the heuristic. However, if detailed knowledge is available, then we can improve on the heuristic by using optimal prices $p_{M_j}(z)$ for each market segment m in cluster M_j, $j = 1, \ldots, J$.

By the choice of the break-points s_j and prices q_j, we have

$$\gamma_J(S) := e(q_1, s_1, z) = e(q_2, s_2, z) = \ldots = e(q_J, s_J, z) \leq 1.$$

The next result shows that $Q^h_J(z)/Q_M(z) \geq \gamma_J(z)$. As we shall see $\gamma_J(z)$ can be quite close to one for relatively small values of J. This indicates that we do not need full price discrimination ($J = M$) to obtain most of the potential profits from price discrimination. Put another way, there may be no need to dice the market into tiny segments if the optimal prices for the different segments are not too far apart.

Theorem 8.18 *Assume that the functions $R_m(p, z)$ are quasi-concave and each has a unique finite maximizer $p_m(z)$. Suppose that the market segments are indexed so that $p_m(z)$ is increasing in $m \in \mathcal{M}$. Assume that $e_m(p, p_m(z), z), m \in \mathcal{M}$ is independent of $m \in \mathcal{M}$. Then offering price q_j to all market segment in M_j for $j = 1, \ldots, J$ results in*

$$\frac{Q_J(z)}{Q_M(z)} \geq \frac{Q^h_J(z)}{Q_M(z)} \geq \gamma_J(z).$$

We now illustrate the lower bounds for a variety of demand functions leaving the proofs as exercises. It is important to recall for this purpose that the market segments are ordered so that $p_m(z)$ is increasing in $m \in \mathcal{M}$, so $p_1(z)$ is the lowest price and $p_M(z)$ is the largest price. Let $\Delta_m(z) := p_m(z) - z$ represent the mark-up for market segment m. Clearly $\Delta_1(z) \leq \Delta_m(z) \leq \Delta_M(z)$.

Proposition 8.19 *Consider linear demand functions $d_m(p) = (a_m - b_m p)$, $m \in \mathcal{M}$. Then*

$$e(p, p_m(z), z) = \frac{p - z}{\Delta_m(z)}\left(2 - \frac{p - z}{\Delta_m(z)}\right) \quad \forall \quad m \in \mathcal{M},$$

and

$$\gamma_J(z) = \frac{4\Delta_1(z)^{1/J}\Delta_M(z)^{1/J}}{\left(\Delta_1(z)^{1/J} + \Delta_M(z)^{1/J}\right)^2}. \tag{8.12}$$

To get a feel for this result, suppose that there are M market segments, and the mark-up for market segment M is 4 times the optimal mark-up of segment 1, so $\Delta_M(z) = 4\Delta_1(z)$. Then $\gamma_1(z) = 64.00\%$, $\gamma_2(z) = 88.89\%$, and $\gamma_4(z) = 97.06\%$. These results are independent of the number of market segments. Recall that these are lower bounds assuming a robust price q_j is used for every cluster, so even better results attain if we use optimal prices within each market segment.

We next consider the exponential demand family.

Proposition 8.20 *Consider the exponential demand functions $d_m(p) = a_m \exp(-p/b_m)$, $m \in \mathcal{M}$. Then*

$$e(p, p_m(z), z) = \frac{p - z}{\Delta_m(z)} \exp\left(1 - \frac{p - z}{\Delta_m(z)}\right) \quad \forall \quad m \in \mathcal{M}.$$

Let $u = b_M/b_1$, and let $U_J = \frac{\ln(u)}{J(u^{1/J} - 1)}$. Then

$$\gamma_J(z) = U_J e^{1-U_J}.$$

To get a feel for this result, suppose that there are M market segments and $u = b_M/b_1 = 4$, then $\gamma_1(z) = 79.13\%$, $\gamma_2(z) = 94.21\%$, and $\gamma_4(z) = 98.51\%$. Again, these numbers are independent of the number of market segments. Recall that these are lower bounds assuming a robust price q_j is used for every cluster, so even better results attain if we use optimal prices within each market segment.

In addition to the linear and log-linear demand functions, efficiency functions can be computed for the CES model and for the multinomial logit model. Consequently, pricing heuristics can be computed for those demand functions as well.

So far we have avoided the issue of consumer surplus and total welfare under direct price discrimination. Most of the insights can be obtained from studying what happens with two market segments. As we move from a common optimal price, say $p(z)$, to two prices, say $p_1(z) < p_2(z)$, we typically have $p(z) \in (p_1(z), p_2(z))$ under mild conditions (Example 8.14 shows that this is not always true). In this case, there is a Robin Hood effect that favors the firm and market segment 1 at the expense of market segment 2. The change in total welfare can be either positive or negative. A necessary condition for an increase in total welfare is that the total output increases under direct price discrimination.

8.5.5 Peak Load Pricing

Suppose that a product with variable cost $\alpha > 0$ is sold in different markets or time periods $m \in \mathcal{M}$. We will assume that $d_m(p)$ is continuous in p, and there is a unique price $p_m(\alpha + z_m)$ that maximizes $R_m(p, \alpha + z_m)$ for all $z_m \geq 0$.

Consider the problem of selecting prices to maximize

$$\sum_{m \in \mathcal{M}} R_m(p_m, \alpha) - \beta \max_{m \in \mathcal{M}} d_m(p).$$

We can think of β as the unit cost of serving the peak demand. To tackle this problem, we will assume that the firm will select the prices $p_m, m \in \mathcal{M}$ as well as the installed capacity, say c. The goal is to maximize

$$\sum_{m \in \mathcal{M}} R_m(p_m, \alpha) - \beta c$$

subject to $d_m(p) \leq c$ for all $m \in \mathcal{M}$.

Let z_m be the dual variable associated with the constraint $d_m(p) \leq c$. Then the Lagrangian problem for fixed c is given by

$$L(p, z) := \sum_{m \in \mathcal{M}} R_m(p_m, \alpha + z_m) + \left[\sum_{m \in \mathcal{M}} z_m - \beta\right] c.$$

Maximizing over $p_m, m \in \mathcal{M}$ yields

$$L(z) := \max_p L(p, z) = \sum_{m \in \mathcal{M}} \mathcal{R}_m(\alpha + z_m) + \left[\sum_{m \in \mathcal{M}} z_m - \beta\right] c.$$

Next we consider the convex problem of minimizing $L(z)$ over $z \geq 0$. The solution is to set $z_m = 0$ if $d_m(p_m(\alpha)) \leq c$. If $d_m(p_m(\alpha)) > c$, we select $z_m > 0$ so $d_m(p_m(\alpha + z_m)) = c$. In summary, for fixed c, the solution is given by $z_m(c)$ and $p_m(\alpha + z_m(c))$ for all $m \in \mathcal{M}$ such that $d_m(p_m(\alpha + z_m(c))) \leq c$ is complementary slack with $z_m(c) \geq 0$.

Let

$$L(z(c)) = \sum_{m \in \mathcal{M}} \mathcal{R}_m(\alpha + z_m(c)) + \left[\sum_{m \in \mathcal{M}} z_m(c) - \beta\right] c.$$

At optimality c^* must be selected so that $\sum_{m \in \mathcal{M}} z_m(c^*) = \beta$, as otherwise the objective can be improved by either increasing or decreasing c. Since $\beta > 0$, at least one period has demand equal to capacity. The variable capacity cost β is allocated to the markets in the set $\{m \in \mathcal{M} : z_m(c^*) > 0\}$ with other markets not contributing to the cost of capacity. Peak load pricing has generated its share of controversy, as it is difficult to understand why two markets consuming the peak capacity should pay different prices, and why those consuming less should get a free ride.

8.6 Multi-Product Pricing Problems

For the multiple product cases with $n > 1$, the known conditions for the existence of a finite maximizer $p(z)$ of $R(p, z) = (p - z)'d(p)$ are seldom useful, as they typically require $R(p, z)$ to be concave or quasi-concave over a compact set. The problem is that for $n > 1$, we need to worry about the possibility that at optimality one or more products are priced at infinity. This is equivalent to not offering all of the products, and this makes it is difficult to reduce the domain to a compact set without loss of optimality. Here, we provide some results for substitute products that sometimes allow for the reduction of the optimization problem to a compact set. Let $d(p) = (d_1(p), \ldots, d_n(p))$. We assume that $d_i(p)$ is increasing in p_j, $j \neq i$ to capture the substitution effect (the demand for chicken goes up as the price for beef increases). For convenience, we will write $p = (p_i, p_{-i})$, where p_{-i} represents the price vector of products other than i. By (p_i, ∞) we imply that products $j \neq i$ are not offered. This allows us to define $d_i(p_i) := d_i(p_i, \infty)$, $R_i(p_i, z_i) := (p_i - z_i)\, d_i(p_i)$, and $\mathcal{R}_i(z_i) := \max_{p_i} R_i(p_i, z_i)$, corresponding to the demand, profit, and optimal profit for product $i \in N := \{1, \ldots, n\}$ that prevail when only product i is offered.

A lower bound on $\mathcal{R}(z)$ can be obtained by selecting the product $i \in N$ with the largest $\mathcal{R}_i(z_i)$, and by setting other prices to infinity. For an upper bound, we have

$$R(p, z) = \sum_{i=1}^{n} (p_i - z_i)\, d_i(p) \leq \sum_{i=1}^{n} (p_i - z_i)\, d_i(p_i) = \sum_{i=1}^{n} R_i(p_i, z_i),$$

so

$$\max_{i \in N} \mathcal{R}_i(z_i) \leq \mathcal{R}(z) \leq \sum_{i \in N} \mathcal{R}_i(z_i). \tag{8.13}$$

We are interested in situations where $p(z)$ is bounded when the optimal individual prices $p_i(z_i)$, $i \in N$ are themselves bounded. For this, we will need the concept of super-modularity. We say that $R(p, z)$ is super-modular in $p \in \Re^n_+$, $p \geq z$, for fixed z, if for any two price vectors $p \geq z$ and $\tilde{p} \geq z$

$$R(\max(p, \tilde{p}), z) + R(\min(p, \tilde{p}), z) \geq R(p, z) + R(\tilde{p}, z).$$

If $R(p, z)$ is twice continuously differentiable in p for fixed z, then R is super-modular in p if and only if

$$\frac{\partial^2 R(p, z)}{\partial p_i \partial p_j} \geq 0 \quad \forall i \neq j.$$

One well-known consequence of super-modularity is that if $R(p_i, p_{-i}, z)$ admits a finite maximizer, say $p_i(z \mid p_{-i}) \geq 0$, for fixed p_{-i} and z, then $p_i(z \mid p_{-i})$ can be selected so that it is increasing in p_j for all $j \neq i$. We are now ready to state our next result.

Theorem 8.21 *If $d_i(p)$ is increasing in p_j, $j \neq i$, then (8.13) holds. Moreover, if $p(z)$ is a maximizer of $R(p, z)$ and $R(p, z)$ is super-modular in p for all $z \geq 0$, and $p_i(z_i)$ is finite for all $i \in N$, then $p(z)$ is finite and*

$$p_i(z) \leq p_i(z_i) \quad \forall i \in N. \tag{8.14}$$

We now provide sufficient conditions for $p_i(z_i)$, $i \in N$ to be finite and for $R(p, z)$ to be super-modular in p for fixed z.

Corollary 8.22 *A sufficient condition for $p_i(z_i) < \infty$ for all $i \in N$ is that $d_i(p_i)$ is USC and $\bar{d}_i(p_i)$ is $o(1/p_i)$ for all $i \in N$. A sufficient condition for the super-modularity of $R(p, z)$ is that for all $i \in N$, $d_i(p + z)$ is increasing in p_j for all $j \neq i$ and super-modular in p for all $z \geq 0$.*

If $d_i(p)$ is decreasing in p_i and increasing in p_j, $j \neq i$, then $R(p, z)$ is super-modular in (p_i, z_i) for fixed p_{-i} and z_{-i}, and sub-modular in (p_j, z_i) for fixed p_{-j} and z_{-i}. As a result, an optimizer $p_i(z \mid p_{-i})$ of $R(p_i, p_{-i}, z)$ can be selected so that $p_i(z \mid p_{-i})$ is increasing in z_i and an optimizer $p_j(z \mid p_{-j})$ of $R(p_j, p_{-j}, z)$ can be selected so that $p_j(z \mid p_{-j})$ is decreasing in z_i. That $p_i(z \mid p_{-i})$ is increasing in z_i is intuitive as some of the higher costs are passed on to consumers. Less intuitive is that $p_j(z \mid p_{-j})$ is decreasing in z_i. The explanation is that an increase in z_i reduces the profits of product i, and an effort is made to shift demand to other products by reducing their prices.

When an inverse demand function exists, it is possible to write the profit function in terms of sales instead of price. In some cases, the profit function is sub-modular as a function of sales for fixed z. Consequently, an increase in sales of one product leads to a decrease in the optimal sales for other products. This makes intuitive sense as products are substitutes. The sub-modularity of the profit function in terms of sales, together with the super-modularity of the profit function in terms of prices, implies that an increase in the price of one product leads to an increase in optimal prices and optimal sales of all other products. Similarly, a decrease in the price of a product leads to a decrease in optimal prices and optimal sales of all other products. This suggests that a change in price in one product should result in a price change of other products in the same direction, but not to the extent that a change in sales goes in the opposite direction of the change in prices.

8.6.1 Linear Demand Model

Demand functions for substitute products are often justified by looking at consumers who are utility maximizers. Given a vector of prices p, consumers purchase the quantity $q \geq 0$ that maximizes $U(q) - q'p$. It is well known that the quadratic utility $U(q) = w'q - \frac{1}{2}q'Qq$, with $w \in \Re^n_{++}$, Q symmetric and positive definite, leads to linear demand function $d(p) = a - Bp$ over the polyhedral set $\mathcal{P} = \{p \geq 0 : Bp \leq a\}$, where $a := Bw$, and $B := Q^{-1}$; see the proof of Theorem 8.7 for details.

We are interested in finding conditions on a and B that guarantee the existence of a unique, non-negative, profit maximizing price vector $p(z)$ such that $\mathcal{R}(z) = R(p(z), z)$ for all $z \geq 0$ such that $d(z) \geq 0$. This last condition limits the costs z to the polyhedral set where demands are non-negative at z. If one or more products have costs so high that $d(z)$ is negative for one or more products, then these products can be eliminated from consideration and it is necessary to work on the projection of the demand model into the space where demand for all products at cost z is non-negative. Given B, we denote the transpose by B' and form the symmetric matrix $S = B + B'$.

Theorem 8.23 *If S is positive definite, $S_{ij} \leq 0$ for all $i \neq j$, and $a \in \Re^n_{++}$, then*

$$p(z) = S^{-1}(a + B'z) \geq 0, \tag{8.15}$$

maximizes $R(p, z) = (p - z)'d(p)$ for all z such that $d(z) \geq 0$. Moreover,

$$\mathcal{R}(z) = R(p(z), z) = d(z)'Nd(z) \tag{8.16}$$

where $N = S^{-1}BS^{-1}$.

Notice that the requirements of Theorem 8.23 are very mild. The theorem does not even require that $B_{ij} \leq 0$ for all $i \neq j$, but rather the more mild assumption that $S_{ij} = B_{ij} + B_{ji} \leq 0$ for all $i \neq j$. Notice also that $R(p, z)$ is super-modular if and only if $S_{ij} \leq 0$ for all $i \neq j$. The requirement that S is positive definite is also very natural in this setting.

The solution $d(p)$ presented above was the solution to the problem of maximizing $U(q) - q'p$ without any constraints on q. The solution $q = d(p)$ satisfies the non-negativity constraint if $p \in \mathcal{P}$. We now address the problem for cases where $p \geq 0$, but $p \notin \mathcal{P}$, so the demand $d(p)$ is negative for at least one product, so $q = d(p)$ is not a feasible solution to the problem of maximizing $U(q) - q'p$ subject to $q \geq 0$. Considering the non-negativity constraints explicitly in the optimization problem can be shown to be equivalent to solving the linear complementarity problem where some of the prices are reduced resulting in an optimal solution to the constrained problem of the form $q = d(p - y) \geq 0$ where $y \geq 0$ and $y'd(p - y) = 0$, so prices are adjusted downward for products with negative demands.

Suppose for some $p \in \mathcal{P}$, the unconstrained solution $q = d(p)$ does not satisfy a capacity constraint of the form $q \leq c$. The problem of maximizing $U(q) - q'p$ subject to $q \leq c$ can be shown to be equivalent to solving a linear complementarity problem where prices are adjusted upwards by y, so that $q = d(p + y) \leq c$ is an optimal solution with $y \geq 0$ and $y'(c - d(p + y)) = 0$.

A natural extension to the linear demand model is $D(p) = A - Bp$, where the potential demand A is random with $\mathbb{E}[A] = a$. Are profits higher when A is random? The answer is yes if we can observe A before deciding the price $p(z|A) = S^{-1}(A + B'z)$ to offer. From (8.16), we can write the optimal profit function as $\mathcal{R}(z) = (A-Bz)'N(A-Bz)$ which is a convex function of A given that N is positive definitive. By Jensen's inequality $\mathbb{E}_A(A - Bz)'N(A - Bz) \geq (a - Bz)'N(a - Bz)$, which is the revenue if we price at $p(z) = S^{-1}(a + B'z)$. The implication here is that dynamic pricing can also be driven by randomness in the potential demand A even if the variable value of capacity is unchanged.

The inverse demand function is given by $p = d^{-1}(q) = B^{-1}(a - q)$, so the profit function as a function of q is given by $q'(B^{-1}a - B^{-1}q - z)$. This function is sub-modular in q if and only if $B^{-1}_{ij} \geq 0$ for all $i \neq j$. A sufficient condition for this is that B is an m-matrix, i.e., if $B_{ii} > 0$ for all i, $B_{ij} \leq 0$ for all $i \neq j$, and either $\sum_{i \in N} B_{ij} > 0$ for every $j \in N$ or $\sum_{j \in N} B_{ij} > 0$ for every $i \in N$. If B is an m-matrix, then the profit function is sub-modular in q, and if a finite maximizer exits, then it can be selected so that $q_i(z|q_{-i})$ is decreasing in q_j for all $j \neq i$. This is intuitively consistent with the idea of product substitution. If we want to sell more of product j then it is optimal to sell less of product i.

8.6.2 The Multinomial Logit Model

The multinomial logit (MNL) demand function is normally derived, as we do in an earlier chapter, from a discrete choice model. Here, we show that the MNL function also arises as a special case of the linear random utility model, where the *indirect utility function*[4] $V(p, y)$ obtained from price vector p and income level y is given by

$$V(p, y) := \mathbb{E}[\max_{i \in N}(y - p_i + a_i + \epsilon_i)],$$

where a_i is a measure of the quality of product i and the ϵ_i's are mean-zero random variables. In this model, it is typically assumed that $y \geq p_i$, so if product i is purchased, then $y - p_i$ is the utility derived from the remaining budget and $a_i + \epsilon_i$ is the utility associated with product i. In this case, $V(p, y) = y + \mathbb{E}\max_{i \in N}(a_i - p_i + \epsilon_i)$. A direct application of the Williams-Daly-Zachary theorem, assuming λ statistically identical consumers, results in

$$d_i(p) = -\lambda \frac{\partial V/\partial p_i}{\partial V/\partial y} = \lambda\, \mathbb{P}(a_i - p_i + \epsilon_i = \max_j(a_j - p_j + \epsilon_j)),$$

[4]The consumer's maximal attainable utility when faced with a vector of prices and income.

so the demand for product i is the expected number of customers that prefer product i over all other alternatives. Notice that the demand is independent of the income level as long as $y \geq p_i$ for all i. The so-called Profit demand function arises if the ϵ_i's are IID normal random variables. The MNL model arises if the ϵ_i's are IID Gumbel random variables. The MNL model results in

$$d_i(p) = \lambda \frac{e^{\alpha_i - \beta p_i}}{1 + \sum_{j \in N} e^{\alpha_j - \beta p_j}} \quad \forall i \in N, \quad \text{and} \quad d_0(p) = \lambda \frac{1}{1 + \sum_{j \in N} e^{\alpha_j - \beta p_j}}$$

for some constants α_j, $j \in N$ and $\beta > 0$, after normalizing the attraction of the no-purchase alternative to 1. One can think of α_i as the quality of product i and β as the sensitivity to price.

For convenience, let $\pi_i(p) = d_i(p)/\lambda$ denote the market share of product $i \in N$. Then

$$\frac{\partial d_i(p)}{\partial p_i} = -\beta d_i(p)(1 - \pi_i(p)) \leq 0 \quad \text{and} \quad \frac{\partial d_k(p)}{\partial p_i} = \beta d_i(p) \pi_k(p) \geq 0 \quad \forall k \neq i.$$

Consequently, the (absolute) elasticity of demand for product i is given by $\beta p_i(1 - \pi_i(p))$ and is proportional to the complement of the market share $\pi_i(p)$ of product i. The cross elasticities of the demand for product k relative to the price of product i are given by $\beta p_i \pi_k(p)$, and it is proportional to the market share of product k. The next theorem characterizes the optimal prices under the MNL model.

Theorem 8.24 *There exists a function $\theta(z)$ independent of i such that*

$$p_i(z) = z_i + \frac{1}{\beta} + \theta(z) \quad \forall i \in N$$

and

$$\mathcal{R}(z) = \lambda \theta(z),$$

where $\theta(z)$ is the root of the Lambert equation

$$\beta \theta e^{\beta \theta} = \sum_{j \in N} e^{\alpha_j - \beta z_j - 1}.$$

It is worth noting that Theorem 8.24 implies that all products should be offered with the same mark-up $p_i(z) - z_i = 1/\beta + \theta(z)$. It is easy to see that the optimal mark-up $1/\beta + \theta(z)$ is equal to the reciprocal of $\beta \pi_0(p(z))$, so

$$p_i(z) - z_i = \frac{1}{\beta \pi_0(p(z))} \quad \forall i \in N.$$

The implication in a competitive setting is that the optimal mark-up is the reciprocal of the product of the price sensitivity and the complement of the market share. Consequently, optimal mark-ups are small if customers are price sensitive and the firm has a small market share.

Let $p_i(z_i) = z_i + 1/\beta + \theta_i$ be the optimal price for the set when the set of finite prices is $F = \{i\}$, corresponding to the case $p_j = \infty$ for all $j \neq i$. Then, from the proof of Theorem 8.24, we see that

$$p_i(z_i) = z_i + \frac{1}{\beta} + \theta_i \leq z_i + \frac{1}{\beta} + \theta = p_i(z).$$

This inequality goes in the opposite direction to that of the linear demand model. This may suggest to the reader that $R(p, z)$ may be sub-modular, but this is not the case.

The analysis can be extended to the case where the demand function is of the form

$$d_i(p) = \lambda \frac{e^{\alpha_i - \beta_i p_i}}{1 + \sum_{j \in N} e^{\alpha_j - \beta_j p_j}} \quad \forall i \in N,$$

with $d_0(p) = 1 - \sum_{i \in N} d_i(p)$, so that the sensitivity to price is now product dependent. In this case, it is also optimal to offer all products, and there is a function $\theta(z)$, independent of i, such that

$$p_i(z) = z_i + \frac{1}{\beta_i} + \theta(z) \quad \forall i \in N,$$

and $\mathcal{R}(z) = \lambda\theta(z)$. However, θ is no longer the root of a Lambert equation, but the root of a slightly more complicated function.

8.6.3 The Nested Logit Model

In this section, we consider pricing under the nested logit (NL) model, which is a popular generalization of the standard MNL model. For a certain range of parameters, the NL model is an example of a random utility model where the random component of the utilities of products within a nest are positively correlated and independent of the utilities of products outside the nest. The probability of selecting a product with the largest utility can then be viewed as a sequential decision: At the upper level, customers select a nest of products; at the lower level, they select a product within the nest. Suppose that the substitutable products constitute n nests and nest i has m_i products. Let $p_i = (p_{i1}, p_{i2}, \ldots, p_{i,m_i})$ be the price vector corresponding to nest $i = 1, \ldots, n$, and let $p = (p_1, \ldots, p_n)$ be the price vector for all the products in all the nests. Let $Q_i(p_1, \ldots, p_n)$ be the probability that a customer selects nest i at the upper level; and let $q_{k|i}(p_i)$ denote the probability that product k of nest i is selected at the lower level, given that the customer selects nest i. Under the NL model, the quantities $Q_i(p_1, \ldots, p_n)$ and $q_{k|i}(p_i)$ are given by

$$Q_i(p_1, \ldots, p_n) = \frac{e^{\gamma_i I_i}}{1 + \sum_{l=1}^{n} e^{\gamma_l I_l}}$$

$$q_{j|i}(p_i) = \frac{e^{\alpha_{ij} - \beta_{ij} p_{ij}}}{\sum_{s=1}^{m_i} e^{\alpha_{is} - \beta_{is} p_{is}}},$$

where α_{is} can be interpreted as the "quality" of product s in nest i, $\beta_{is} \geq 0$ is the product-specific price sensitivity for that product, $I_l = \log \sum_{s=1}^{m_l} e^{\alpha_{ls} - \beta_{ls} p_{ls}}$ represents the attractiveness of nest l, which is the expected value of the maximum of the utilities of all the products in nest l, and nest coefficient γ_i can be viewed as the degree of inter-nest heterogeneity and is a measure of the correlation among the utilities of the products in nest i. When $\gamma_i = 1$ for all i, the model reduces to the MNL model. The case $\gamma_i \in (0, 1]$ is consistent with random utility theory.

The probability that a customer will select product j of nest i, which can also be considered as the market share of that product, is

$$\pi_{ij}(p_1, \ldots, p_n) = Q_i(p_1, \ldots, p_n) q_{j|i}(p_i). \tag{8.17}$$

The monopolist's problem is to determine the price vectors $(p_1, \ldots, p_n)$ to maximize the total expected profit

$$R(p, z) := \sum_{i=1}^{n} \sum_{j=1}^{m_i} \lambda(p_{ij} - z_{ij}) \pi_{ij}(p_1, \ldots, p_n), \tag{8.18}$$

where $z = (z_1, \ldots, z_n)$, and z_i is the vector of unit costs for nest i, and λ is the market size. Let $\mathcal{R}(z) := \max_{(p_1, \ldots, p_n)} R(p, z)$. The objective function $R(p, z)$ fails to be quasi-concave in prices. When the objective function is rewritten with market shares as decision variables, then the objective function can be shown to be concave if the price sensitivity parameters $\beta_{ij} = \beta_i$ are product independent in each nest and $\gamma_i \in (0, 1]$ for all i. However, the objective function fails to be concave in the market shares in the more general case where the price sensitivities are product dependent or some of the parameters γ_i are allowed to exceed one.

The results of Theorem 8.24 extend to the NL model, where the optimal price $p_{ij}(z)$ for product j in nest i as a function of the vector of unit costs z is of the form $p_{ij}(z) = z_{ij} + 1/\beta_{ij} + \theta_i$. Also, the nest dependent constants $\theta_i, i = 1, \ldots, n$ are linked to a single parameter as explained in the following theorem.

Theorem 8.25 *If* $\gamma_i \geq 1$ *or* $\frac{\max_s \beta_{is}}{\min_s \beta_{is}} \leq \frac{1}{1-\gamma_i}$, *then there exists a unique constant* ϕ *such that*

$$\theta_i + \left(1 - \frac{1}{\gamma_i}\right) w_i(\theta_i) = \phi,$$

and

$$p_{ij}(z) = z_{ij} + \frac{1}{\beta_{ij}} + \theta_i,$$

where $w_i(\theta) = \sum_{k=1}^{m_i} \frac{1}{\beta_{ik}} \cdot q_{k|i}(\theta_i)$ *and* $q_{k|i}(\theta_i) = \frac{e^{\tilde{\alpha}_{ik} - \beta_{ik}\theta_i}}{\sum_{s=1}^{m_i} e^{\tilde{\alpha}_{is} - \beta_{is}\theta_i}}$, *and* $\tilde{\alpha}_{is} = \alpha_{is} - \beta_{is} z_{is} - 1$ *for all* i *and all* s. *Moreover,*

$$\mathcal{R}(z) = \lambda\phi.$$

Theorem 8.25 is interesting because a non-concave optimization problem over $\sum_{i=1}^{n} m_i$ variables can be reduced, under mild conditions, to a root finding problem over the single variable ϕ. Notice that each value of ϕ gives a set of θ_i's dictated by the first equation in the theorem. For these θ_i's, the second equality in the theorem gives the prices. If $\gamma_i \geq 1$ or $\frac{\max_s \beta_{is}}{\min_s \beta_{is}} \leq \frac{1}{1-\gamma_i}$ fails to hold, then the problem reduces to a single variable maximization problem of a continuous function over a bounded interval, so the problem can be easily solved numerically. Also, if different firms control different nests, then the pricing problem under competition is strictly log-super-modular in the nest mark-up constants, so the equilibrium set is nonempty with the largest equilibrium preferred by all the firms.

8.7 End of Chapter Problems

1. Show that if $d(p) = 1$ for $p \in [0, 10)$ and $d(p) = 0$ for $p \geq 10$, then $\mathcal{R}(z) = (10 - z)^+$ but the maximum is not attained.
2. Show that Theorem 8.10 applies to the demand function $d(p) = a\exp(-bp)\sin^2(p)$ by showing that $\bar{d}(p) \leq a\exp(-bp)$ and $p\bar{d}(p) \to 0$ as $p \to \infty$. Find a formula for $p(z)$.
3. Determine the form of $p(z)$ for $d(p) = \lambda\exp(-p/\theta)$ for $\lambda, \theta > 0$.
4. Determine the form of $p(z)$ for $d(p) = ((a - bp)^+)^c$ for $a, b, c > 0$.
5. Determine the form of $p(z)$ for $d(p) = (a + bp)^{-c}$ for $a, b > 0$, $c > 1$.
6. Show that if $1/h(p)$ is concave in p, then $p(z)$ is increasing concave in z.
7. Consider a single product where c units are available for sale. Let $d_c(p) = \min(d(p), c)$ and consider the following two formulations: (i) $\max_p p\, d_c(p)$, and (ii) $\max p\, d(p)$ subject to $d(p) \leq c$. Solve each problem for $c = 2$, if $d(p) = 3$ for $p \leq 10$ and $d(p) = 0$ for $p > 10$. What happens if $d(p)$ is continuous?
8. Consider a single product problem with a strictly decreasing demand function $q = d(p)$. Let $\tilde{p}(q)$ be the inverse demand function. Assume that $\tilde{r}(q) = q\,\tilde{p}(q)$ is concave with a bounded maximizer, say q^*. Suppose $c < q^*$. Show that $q = c$ solves the problem of maximizing $\tilde{r}(q)$ subject to $q \leq c$ and that there is a $z \geq 0$ such that $q = c$ maximizes $\tilde{r}(q) - qz$. Is z unique? What

if $\tilde{r}(q)$ is differentiable? Is the concavity of $\tilde{r}(q)$ a sufficient condition for the existence of a unique maximizer $p(z)$ of $R(p,z)$? What else may you need? Consider the problem $\Gamma(c) = \min_{z\geq 0}[\mathcal{R}(z) + cz]$. Is the concavity of $\tilde{r}(q)$ a sufficient condition for $\Gamma(c) = \tilde{r}(c)$?

9. Consider a single product, single resource formulation (8.4) with sales horizon $[0, T]$, $T = 1$, $d_t(p) = a_1 - bp$ for $t \in (0, 1/2]$, and $d_t(p) = a_2 - bp$ for $t \in (1/2, 1]$. Find the solution $p_t(0)$ for $t \in [0, T]$. For what values of c is this solution optimal? Show that for such values of c, $\bar{V}(T, c) = (a_1^2 + a_2^2)/8b$. Now solve formulation (8.6) and show that $\bar{V}_f(T, c) = (a_1 + a_2)^2/16b$. Show that the optimal profits are 25% higher under dynamic pricing when $a_1 = 50$ and $a_2 = 150$.
10. Consider the demand function $d(p) = \lambda P(\Omega \geq p)$ and assume that Ω has a gamma distribution with parameters α and β, so $\mu = \mathbb{E}[\Omega] = \alpha\beta$ and $\sigma^2 = \text{Var}[\Omega] = \alpha\beta^2$. We can fit any mean μ and variance σ^2 by setting the parameter $\beta = \sigma^2/\mu$ and $\alpha = \mu^2/\sigma^2$. One may wonder how $p(z)$ and $\mathcal{R}(z)$ behave as a function of σ^2 for fixed μ. Does more variance lead to higher or lower prices and profits? Construct a table of $p(z)$ and $\mathcal{R}(z)$ for $z = 400$, $\lambda = 1$ and $\mathbb{E}[\Omega] = 500$ for values of $\sigma/\mu \in \{k/8 : k = 0, 1, \ldots, 16\}$. What happens to $p(z)$ as σ increases? What happens to $\mathcal{R}(z)$ as σ increases?
11. Consider the problem of maximizing $R(p,z) = (p - z)d(p)$ when $d(p) = \lambda e^{-p/\mu}$ subject to the constraint $d(p) \leq c$. Let $L(p,z) = R(p,z) - \gamma(d(p) - c) = R(p, z + \gamma) + \gamma c$. Argue that $\min_{\gamma\geq 0} L(p,z)$ is equivalent to $\min_{\gamma\geq 0} [\mathcal{R}(z + \gamma) + \gamma c]$. Show that this is a convex program in γ and that if γ_c is an unconstrained minimizer of $R(p + \gamma) + \gamma c$, then $\gamma_c^* = \max(\gamma_c, 0)$ solves the Lagrangian problem, and that $p(z + \gamma_c^*) = \max(p(z), p_{\min}(c))$ is the optimal price, where $p_{min}(c)$ is the root of $d(p) = c$.
12. Consider the multiple market segment problem and show that the total welfare can go up when we move from an optimal common price to direct price discrimination only if the total output goes up. Hint: Use the fact that the surplus function is convex.
13. For the linear function $d(p) = a - Bp$ and for $p \in \mathcal{P} = \{p \geq 0 : d(p) \geq 0\}$, $q = d(p)$ is the solution to the problem $\max_{q\geq 0}[U(q) - p'q]$ as presented in Sect. 8.6.1. Suppose that $p \geq 0$, but $p \notin \mathcal{P}$ and $a - Bp$ has both positive and negative components. To find the demand at p, we need to solve $\max_{q\geq 0}[U(q) - p'q]$ without ignoring the non-negativity constraints. Let $y \geq 0$ be the vector of dual variables to the constraint $q \geq 0$. Show that the optimal solution is given by $q = d(p) + By$, where $y \geq 0$ minimizes $y'[d(p) + By]$. Notice that this is a linear complementarity problem. Solve the linear complementarity problem for the case of $n = 2$, when $d_1(p) > 0$ and $d_2(p) < 0$ to see how the demand for product one is reduced by $b_{12}\, y_2$.
14. Consider again the linear demand model $d(p) = a - Bp$, but assume now that there are n firms with firm i selecting the price of product $i = 1, \ldots, n$. More precisely, assume that firm i maximizes $R_i(p, z) = (p_i - z_i)d_i(p)$ over $p_i \geq z_i$. This results in the best response price $p_i(p_{-i})$ for each firm i, and what we seek

is the equilibrium price vector, so that the prices constitute a Nash Equilibrium. Show that first-order conditions can be written in matrix form as

$$a + \text{diag}(B)z - (B + \text{diag}(B))p = 0.$$

Assume that $M = B + \text{diag}(B)$ is an m-matrix, so that the inverse of M exists and is non-negative and show that the equilibrium prices are given by

$$\tilde{p}(z) = M^{-1}(a + \text{diag}(B)z) = z + M^{-1}d(z).$$

Show also, that $d(\tilde{p}(z)) = \text{diag}(B)M^{-1}d(z)$ and that

$$\sum_{i=1}^{n} R_i(\tilde{p}(z), z) = d(z)'M'^{-1}\text{diag}(B)M^{-1}d(z)$$

$$= (\tilde{p}(z) - z)'\text{diag}(B)(\tilde{p}(z) - z)),$$

and that the monopolist formulas (8.15) and (8.16) coincide with the competition formulas if $\text{diag}(B) = B$, i.e. if there are no cross elasticities. Otherwise, we expect competitive prices to be lower, demand to be higher and aggregate profits to be lower under competition, with more surplus going to the consumers. The results can be extended to the case where each competitor controls the prices of a subset of the products.

8.8 Bibliographical Remarks

Theorems 8.1, 8.5, and 8.7 show that the firm prefers randomness in z, the consumer's prefer randomness in p, and under mild conditions both the firm and the representative consumer prefer prices that respond to randomness in variable costs. Theorem 8.10 allows for demand functions that are not necessarily decreasing or eventually decreasing. Theorem 8.11 provides bounds on optimal prices. The analysis of consumer surplus is due to Chen and Gallego (2019). The reader is directed to van den Berg (2007) and references therein for earlier efforts to characterize the existence or uniqueness of global maximizers. The reader is also referred to Larriviere and Porteus (2001) for an equivalent assumption where the absolute value of the price elasticity of demand $|e(p)| = ph(p)$ is called the generalized hazard rate. Caplin and Nalebuff (1991) have some interesting conditions on the inverse demand function for an optimal price $p(z)$ to exist. Ziya et al. (2004) discuss the relationship between several assumptions used to ensure that the expected revenue function is well behaved. The results on options are due to Gallego and Sahin (2010), Png (1989), Shugan and Xie (2000), and Xie and Shugan (2001). The section on priority pricing is based on the work of McAfee (2004), where he considers the social benefit of coarse matching. See Johnson (1970) for a discussion of positive definite matrices. The pricing results for the NL model can be found in Li and Huh (2011), Gallego and Wang (2014), and Rayfield et al. (2015). The development and discussion of the NL model can be found in McFadden (1974) and Carrasco and de Ortuzar (2002). The MNL and NL models are special cases of a broader class of choice models, called the generalized extreme value models. Zhang et al. (2018) work on pricing problems under generalized extreme value models.

Keller et al. (2014) give mathematical programming formulations for pricing problems under generalizations of the MNL model. Du et al. (2016), Wang and Wang (2017) and Du et al. (2018) study pricing problems under a variant of the MNL model, where the attraction value of a product depends on the market size it garners. Yan et al. (2017) study a joint parameter estimation and pricing problem when the marginal distribution information is available on the utilities. Wang (2018b) studies a pricing problem under the MNL model, where customers form a reference price by using the prices of the offered products and adjust their reactions accordingly. Cui et al. (2018) and Wang et al. (2019) study multi-product pricing problems when products are sold as ancillary to others. Amornpetchkul et al. (2018) examine promotion models, when the amount of promotion depends on the quantity purchased by a customer.

Maglaras and Zeevi (2005) study pricing problems when the firm offers services with different levels of quality using a common pool of capacity. Besbes et al. (2010) design tests to check the validity of a fitted price-demand curve not from the perspective of statistical goodness of fit but from the perspective of operational performance. Eren and Maglaras (2010) consider pricing problems when the price-demand curve is unknown to the firm. Cachon and Feldman (2011) study pricing models to understand the tradeoff between charging on a per-use basis or selling subscriptions. Kostami et al. (2017) give a pricing model when the utility of a customer depends on the presence of the other customers in the system. Cohen et al. (2017b) give performance bounds when only partial information about the price-demand relationship is available. Similarly, Chen et al. (2017a) study pricing problems with only limited information about the price-demand relationship. Cachon et al. (2017) study a stylized pricing model for a two-sided platform where the demand and the supply are both endogenous. Hu and Nasiry (2018) demonstrate that a price-demand model that is obtained by aggregating the behavior of individual customers may not reflect the individual customers anymore. Elmachtoub et al. (2018) bound the relative expected revenue gain when a firm knows the exact willingness to pay of a customer rather than the distribution of willingness to pay, providing insights on the effectiveness of personalized pricing. Boyaci and Akcay (2018) study pricing models when customers cannot fully evaluate the quality of a product. Ho et al. (1998) study a model to understand the reaction of consumers to different pricing strategies. Petruzzi and Dada (1999) and Lu and Simchi-Levi (2013) study incorporating pricing decisions into the newsvendor problem. Tang et al. (2004) analyze the benefits from providing advance booking discounts to reduce demand uncertainty. Tang and Yin (2007) develop a joint procurement and pricing model under deterministic demand.

Rusmevichientong et al. (2006) work on a nonparametric pricing problem, where each customer is represented by a budget and a preference list of products. Caldentey and Wein (2006) give fluid approximations for a joint pricing and admission control problem. Hu et al. (2013a) study a pricing problem with a gray market, which acts as an authorized channel to sell the authentic products of a supplier. Phillips (2013) describes a host of practical issues in pricing credit and gives a mathematical model. In an opaque product, a feature of a product, such as color for a piece of apparel or departure time for a flight, is hidden from the customer until the purchase occurs. Elmachtoub and Wei (2015) and Elmachtoub and Hamilton (2017) study pricing problems for opaque products. Belkaid and Martinez-de-Albeniz (2017) estimate the effect of weather conditions on demand and study the effectiveness of weather-dependent pricing strategies. Courty and Nasiry (2018) observe that certain products with different observable qualities are sold at uniform price and develop a model to resolve this paradox. Ma and Simchi-Levi (2018) develop a model that exploits the information extracted from bundled products to estimate individual price sensitivities.

The reader is referred to Anderson et al. (1992) for more on discrete choice theory of product differentiation and Vives (2001) for comparative static tools and oligopoly pricing.

Appendix

Proof of Theorem 8.1 It is clear that $R(p, z)$ is decreasing in z and that this implies that $\mathcal{R}(z)$ is decreasing in z. To verify convexity, let $\alpha \in (0, 1)$. Then for any $z, \tilde{z}$,

$$
\begin{aligned}
\mathcal{R}(\alpha z + (1-\alpha)\tilde{z}) &= \max_{p \in X} R(p, \alpha z + (1-\alpha)\tilde{z}) \\
&= \max_{p \in X} R(\alpha p + (1-\alpha)p, \alpha z + (1-\alpha)\tilde{z})) \\
&= \max_{p \in X} \left[\alpha (p - z)' + (1-\alpha)(p - \tilde{z})'\right] d(p) \\
&= \max_{p \in X} \left[\alpha R(p, z) + (1-\alpha) R(p, \tilde{z})\right] \\
&\leq \alpha \max_{p \in X} R(p, z) + (1-\alpha) \max_{p \in X} R(p, \tilde{z}) \\
&= \alpha \mathcal{R}(z) + (1-\alpha)\mathcal{R}(\tilde{z}).
\end{aligned}
$$

□

Proof of Proposition 8.2 This follows from a direct application of the Taylor's expansion around $\mathcal{R}(\mathbb{E}[Z])$. □

Proof of Corollary 8.4 The proof of the first part is left as an exercise. From the concavity of g we have $g(\alpha z + (1-\alpha)\tilde{z}) \geq \alpha g(z) + (1-\alpha)g(\tilde{z})$ for any $z, \tilde{z} \in \Re^m$ and any $\alpha \in [0,1]$. Since $\mathcal{R}$ is decreasing, it follows that $\mathcal{R}(g(\alpha z + (1-\alpha)\tilde{z}) \leq \mathcal{R}(\alpha g(z) + (1-\alpha)g(\tilde{z}))$. From the convexity of $\mathcal{R}$, we have $\mathcal{R}(\alpha g(z) + (1-\alpha)g(\tilde{z})) \leq \alpha\mathcal{R}(g(z)) + (1-\alpha)\mathcal{R}(g(\tilde{z}))$. Consequently, $\mathcal{R}(g(\alpha z + (1-\alpha)\tilde{z})) \leq \alpha\mathcal{R}(g(z)) + (1-\alpha)\mathcal{R}(g(\tilde{z}))$, showing that $\mathcal{R}(g(z))$ is convex in z. From Jensen's inequality, it follows that $\mathbb{E}[\mathcal{R}(g(Z))] \geq \mathcal{R}(g(\mathbb{E}[Z]))$. □

Proof of Theorem 8.5 That $\mathcal{S}(p)$ is decreasing follows directly from the fact that $S(q,p)$ is decreasing in p. To verify convexity, let $\alpha \in (0,1)$. Then for any $p, \tilde{p}$

$$\begin{aligned}
\mathcal{S}(\alpha p + (1-\alpha)\tilde{p}) &= \max_{q\geq 0} S(\alpha p + (1-\alpha)\hat{p}, q) \\
&= \max_{q\geq 0} \left[U(q) - (\alpha p + (1-\alpha)\hat{p})'q\right] \\
&= \max_{q\geq 0} \left[\alpha(U(q) - p'q) + (1-\alpha)(U(q) - \tilde{p}'q)\right] \\
&\leq \alpha \max_{q\geq 0} S(p,q) + (1-\alpha)\max_{q\geq 0} S(\tilde{p}, q) \\
&= \alpha\mathcal{S}(p) + (1-\alpha)\mathcal{S}(\tilde{p}).
\end{aligned}$$

Notice that $\mathcal{S}(p) = U(d(p)) - p'd(p)$, so $\nabla\mathcal{S}(p) = \nabla d(p)\nabla U(d(p)) - d(p) - \nabla d(p)p = -d(p)$ on account of $\nabla U(d(p)) = p$ for all $p \in \mathcal{P}$.[5] □

Proof of Theorem 8.7 Let $w'q - \frac{1}{2}q'Qq$ be the quadratic approximation to an increasing concave utility function U, where w is a vector of positive components, and Q is symmetric positive definite matrix.[6] Let $B = Q^{-1}$ and write $d(p) = B(w-p)$ over the set $\mathcal{P} = \{p : p \geq 0, B(w-p) \geq 0\}$. Then

$$\mathcal{S}(p) = U(d(p)) - d(p)'p = \frac{1}{2}(w-p)'B(w-p) \;\text{ over }\; p \in \mathcal{P},$$

which is decreasing convex in $p \in \mathcal{P}$ since B is positive definite. The firm's problem is to find $p = p(z)$ that maximizes

$$R(p,z) = (p-z)'B(w-p).$$

The optimizer is given by $p(z) = (w+z)/2$, which is an increasing linear function of $z \in \mathcal{P}$. The composite function $\mathcal{S}(p(z))$ is therefore convex. □

[5] Notice here that $\nabla d(p)$ is the Jacobian of $d(p)$, i.e., the matrix of partial derivatives $\partial d_i(p)/\partial p_j$.

[6] If Q is not symmetric we can transform $Q \leftarrow (Q+Q')/2$ to make it symmetric without changing the utility function.

Proof of Theorem 8.10 Since $d(p)$ is USC and the product of non-negative USC functions is also USC, it follows that $R(p, z)$ is USC. The USC of $d(p)$ implies the USC of $\bar{d}(p)$ for if $\bar{d}(p)$ is not USC at p_0, then there exist a $p_1 > p_0$ at which $d(p_1) = \bar{d}(p_o)$ fails to be USC. As a result $\bar{R}(p, z)$ is also USC in $p \in [z, \infty)$. If $d(p) = 0$ for all $p \geq z$, then $p(z) = z$ and $\mathcal{R}(z) = R(z, z) = 0$ and there is nothing to prove. Otherwise there exists a price $\hat{p} > z$ such that $0 < \bar{d}(\hat{p}) < \infty$, for if $\bar{d}(p) = \infty$ for all $p > z$, then $\bar{d}(p)$ is not $o(1/p)$. Let $\epsilon = \bar{R}(\hat{p}, z) > 0$. We will show that there is a price $\tilde{p} > \hat{p}$ such that $\bar{R}(p, z) \leq \epsilon$ for all $p > \tilde{p}$, for if not, then for any $\tilde{p} > z$, we can find a $p > \tilde{p}$ such that $\bar{R}(p, z) > \epsilon$, or equivalently, $p\bar{d}(p) > p\epsilon/(p - z)$, contradicting the fact that $p\bar{d}(p) \to 0$ as $p \to \infty$. Given that $\bar{R}(p, z) \leq \epsilon$ for all $p \geq \tilde{p}$, we can restrict the optimization of $\bar{R}(p, z)$ without loss of optimality to the compact set $[z, \tilde{p}]$. The extreme value theorem guarantees the existence of a finite price, say $\bar{p}(z) \in [z, \tilde{p}]$, that maximizes $\bar{R}(p, z)$. We will now show that $p(z) = \bar{p}(z)$ also maximizes $R(p, z)$ so $\mathcal{R}(z) = \bar{\mathcal{R}}(z)$. Assume for a contradiction that $\bar{p}(z)$ is not a maximizer of $R(p, z)$. Then

$$(\bar{p}(z) - z)\bar{d}(\bar{p}(z)) = \bar{\mathcal{R}}(z) \geq \mathcal{R}(z) > (\bar{p}(z) - z)d(\bar{p}(z))$$

implies that $d(\bar{p}(z)) < \bar{d}(\bar{p}(z)) = \sup_{p \geq \bar{p}(z)} d(p)$. Then there exists a $p' > \bar{p}(z)$ such that $d(p') = \bar{d}(p(z))$, but then $\bar{R}(p', z) > \bar{R}(\bar{p}(z), z) = \bar{\mathcal{R}}(z)$ contradicting the optimality of $\bar{p}(z)$. □

Proof of Theorem 8.11 First, we show Part a. If $h(p)$ is continuous and increasing in p, then $f(p)$ is continuous and strictly decreasing in $p \geq z$. Equivalently, $(p - z)h(p)$ is continuous and strictly increasing in p. Now $f(z) = 1 > 0$ implies that $p(z) > z$, while $f(z + 1/h(z)) = 1 - h(z + 1/h(z))/h(z) \leq 0$ on account of $h(z + 1/h(z)) \geq h(z) > 0$ implies that $p(z) \leq z + 1/h(z)$. Because $(p - z)h(p)$ is continuous and strictly increasing in p, there exist a unique $p(z)$ satisfying $p(z) = \sup\{p : f(p) \geq 0\}$ that is bounded below by z and above by $z + 1/h(z)$. Suppose that $z' > z$, then $(p(z) - z')h(p(z)) < 1$, so $p(z') > p(z)$ showing that $p(z)$ is strictly increasing in z. To show that $\Delta(z) = p(z) - z$ is decreasing in z, let $p' = z' + \Delta(z)$ and notice that $(p' - z')h(p') = \Delta(z)h(p') \geq \Delta(z)h(p(z)) = 1$, so $p(z') = z' + q(z') \leq p' = z' + \Delta(z)$ implying that $\Delta(z') \leq \Delta(z)$. For the exponential demand function $d(p) = \lambda e^{-p/\theta}$, we have $h(z) = 1/\theta$ and $p(z) = z + \theta = z + 1/h(z)$, so the upper bound is attained.

Next, we show Part b. If $ph(p)$ is continuous and strictly increasing in p and $\tilde{z}h(\tilde{z}) > 1$, then $f(p)$ is continuous in $p > z$ and the equation $f(p) = 0$ can be written as $ph(p) = p/(p - z)$ with the left hand side increasing in p and the right hand side strictly decreasing to one for $p > z$. Since $zh(z) < \infty$ it follows that $p(z) > z$. Notice that $z/(1 - \tilde{z}h(\tilde{z}))$ is the root of $\tilde{z}h(\tilde{z}) = p/(p - z)$. Since $ph(p) \geq \tilde{z}h(\tilde{z}) \geq p/(p - z)$ for all $p \geq z/(1 - \tilde{z}h(\tilde{z}))$, it follows that $p(z)$ is unique and bounded above by $z/(1 - 1/\tilde{z}h(\tilde{z}))$. Suppose that $z' > z$, then $p(z') > z'$, so if $z' \geq p(z)$ it follows immediately that $p(z') \geq p(z)$. Suppose now that $z < z' < p(z)$, then at $p = p(z)$ we have $ph(p) < p/(p - z')$ implying that $p(z') > p(z)$. For $d(p) = \lambda p^{-b}$, with $b > 1$, we have $ph(p) = b$ for all p, and

$p(z) = bz/(b-1) = z/(1-1/b) = z/(1-1/\tilde{z}h(\tilde{z}))$, so the upper bound is attained. Let $m(z) := 1/h(z)$. Then, using the implicit function theorem on $f(p, z) = 0$, we can find the first and second derivatives of $p(z)$ in terms of $m(z)$. It is easy to see that the first derivative is given by $p'(z) = (1-m'(p(z))^{-1}$, so the second derivative is given by

$$p''(z) = \frac{m''(p(z)p'(z))}{(1 - m'(p(z)))^2} \leq 0,$$

since $m''(z) \leq 0$ and $p'(z) > 0$.

Finally, we show Part c. Clearly $\tilde{f}(p) \leq f(p)$ so $\tilde{p}(z) \leq p(z)$. □

Proof of Proposition 8.13 Since the sum of USC is USC it follows that $d_S(p)$ is USC. Moreover $\bar{d}_m(p) = o(1/p)$ for all $m \in \mathcal{M}$ implies that $\bar{d}_S(p) = o(1/p)$. As a result $d_S(p)$ satisfies the conditions of Theorem 8.10 so there exists a finite price $p_S(z)$, increasing in z, such that $\mathcal{R}_S(z) = R_S(p_S(z), z)$ is decreasing convex in z. □

Proof of Proposition 8.16 It is easy to see that $p_m(z) > z$ is the root of $p/(p-z) = ph_m(p)$. Since the left hand side is decreasing in p and $ph_m(p)$ is increasing in p, it follows that there is a unique root $p > z$. This observation implies that $f_m(p) > 0$ on $p < p_m(z)$ and $f_m(p) < 0$ on $p > p_m(z)$. Let $f_S(p) = 1 - (p - z)h_S(p)$ where $h_S(p)$ is the hazard rate of $d_S(p)$. Since $f_S(p)$ is a convex combination of $f_m(p) = 1 - (p - z)h_m(p)$ with weights $\theta_m(p) = d_m(p)/d_S(p)$, it follows that $f_S(p) > 0$ for all $p < \min_{m\in S} p_m(z)$ because over that interval $f_m(p) > 0$ for all $m \in S$. Also $f_S(p) < 0$ for all $p > \max_{m\in S} p_m(z)$ because over that interval $f_m(p) < 0$ for all $m \in S$. Since the derivative of $R_S(p, z)$ is proportional to $f_S(p)$ it follows that $R_S(p, z)$ is increasing over $p < \min_{m\in S} p_m(z)$ and decreasing over $p > \max_{m\in S} p_m(z)$. Moreover, since $R_S(p, z)$ is continuous over the closed and bounded interval $[\min_{m\in S} p_m(z), \max_{m\in S} p_m(z)]$, appealing to the EVT yields the existence of a global maximizer $p_S(z)$ of $R_S(p, z)$. □

Proof of Theorem 8.18 Clearly

$$\begin{aligned}
\frac{Q_J(z)}{Q_M(z)} \geq \frac{Q_J^h(z)}{Q_M(z)} &= \frac{\sum_{j=1}^J \sum_{m\in M_j} R_m(q_j, z)}{Q_M(z)} \\
&= \sum_{j=1}^J \sum_{m\in M_j} e(q_j, p_m(z), z) \frac{\mathcal{R}_m(z)}{Q_M(z)} \\
&\geq \gamma_J(z) \frac{\sum_{m\in\mathcal{M}} \mathcal{R}_m(z)}{Q_M(z)} \\
&= \gamma_J(z).
\end{aligned}$$

□

Proof of Theorem 8.21 We have already shown that inequalities in (8.13). To show (8.14), notice that the super-modularity of $R(p, z)$ in p for fixed z, allows us to select $p_i(z|p_{-i})$ so that it is increasing in p_{-i}. Consequently, $p_i(z|p_{-i}) \leq p_i(z_i|\infty) = p_i(z_i)$ for all i. In particular, $p_i(z) = p_i(z|p_{-i}(z)) \leq p_i(z_i)$ for all $i \in N$. □

Proof of Theorem 8.23 Maximizing $R(p, z) = (p - z)'d(p)$ with respect to p is equivalent to minimizing $\frac{1}{2}p'Sp - (a + B'z)'p + a'z$ which is quadratic function. A sufficient condition for this function to be convex is that S is positive definitive. It is known that S is positive definitive, if and only if B is, see Johnson (1970). If B is positive definitive then S is invertible and since S is symmetric, so it is inverse S^{-1}. If B is positive definitive then the maximizer of $R(p, z)$ is given by (8.15). A sufficient condition for $p(0) = S^{-1}a \geq 0$ is for $S^{-1} \geq 0$, since $a > 0$. However, this is true because S is an s-matrix, i.e. a real symmetric, positive definitive matrix with non-positive off-diagonal elements. It is known that an s-matrix has a non-negative inverse implying that $S^{-1} \geq 0$, and consequently that $p(0) = S^{-1}a \geq 0$. Since $p(z)$ is non-decreasing in z by Theorem 8.1, it follows that $p(z) \geq p(0) \geq 0$ for all $z \geq 0$ such that $d(z) \geq 0$.

By adding and subtracting Bz to the expression in parenthesis on the right hand side of (8.15) we can write $p(z) - z = S^{-1}d(z)$, where $d(z)$ is the demand at $p = z$. It is also possible to write $d(p(z)) = a - Bp(z) = a - B(p(z) + z - z) = a - Bz - B(p(z) - z) = (I - BS^{-1})d(z)$ and then use the fact that $I - BS^{-1} = B'S^{-1}$ to obtain $d(p(z)) = B'S^{-1}d(z)$. This allows us to write $\mathcal{R}(z) = (p(z) - z)'d(p(z)) = d(z)'S^{-1}B'S^{-1}d(z) = d(z)'S^{-1}B'S^{-1}d(z)$ resulting in (8.16). □

Proof of Theorem 8.24 The first-order conditions are of the form

$$\frac{\partial R(p, z)}{\partial p_i} = d_i(p)[1 + \beta R(p, z)/\lambda - \beta(p_i - z_i)] = 0 \quad \forall i \in N.$$

For every subset $F \subseteq N$, let $p^F(z)$ be the solution to the first-order conditions obtained by setting the expression in brackets equal to zero for all $i \in F$ and by setting $d_i(p) = 0$ for all $i \notin F$. Then,

$$p_i = z_i + 1/\beta + R(p, z)/\lambda \quad \forall i \in F \quad \text{and} \quad p_i = \infty \quad \forall i \notin F.$$

For each $F \subseteq N$, there exists a constant, $\theta_F = R(p^F(z), z)/\lambda$, given by the root of the Lambert type equation

$$\beta\theta e^{\beta\theta} = \sum_{j \in F} e^{\alpha_j - \beta z_j - 1},$$

such that $p_i^F(z) = z_i + 1/\beta + \theta_F$ for all $i \in F$, and $R(p^F(z), z) = \lambda\theta_F$, so θ_F represents the optimal profit per customer when we are allowed to offer only products in F. Since the root θ_F is increasing in F, it follows that among all the 2^n

solutions to the first-order conditions, the one with the highest profit corresponds to $F = N$. Thus, at optimality, we have

$$p_i(z) = z_i + \frac{1}{\beta} + \theta,$$

where θ is the root of the Lambert equation for $F = N$. Moreover, $\mathcal{R}(z) = \lambda\theta$ is the optimal profit. □