The Impact of Strategic Consumer Behavior on the Value of Operational Flexibility

Gérard P. Cachon
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104
cachon@wharton.upenn.edu

Robert Swinney
Graduate School of Business
Stanford University
Stanford, CA 94305-5015
swinney@stanford.edu

October 1, 2008

Chapter to Appear in *Operations Management Models with Consumer-Driven Demand,* Serguei Netessine and Christopher Tang, Editors

Abstract

Increasingly sophisticated consumers have learned to anticipate future price reductions and forego purchasing products until markdowns occur. Such forward-looking or strategic behavior on the part of consumers can have a significant impact on retail margins by shifting a large number of sales from higher, "full" prices to lower, "clearance" prices. Some firms, however, have become adept at dealing with the strategic consumer problem by implementing various forms of operational flexibility (for example, investing in faster supply chains capable of rapidly responding to changing demand conditions). A firm famous for this strategy is the Spanish fashion retailer Zara. In this chapter, we explore the strategic consumer purchasing phenomenon, and in particular address how the Zara model of operational flexibility impacts consumer behavior (and, conversely, how consumer behavior impacts the value of operational flexibility). We examine in detail the consequences of *volume flexibility*—the ability of a firm to adjust production or procurement levels to meet stochastic demand—and demonstrate that this type of flexibility can be highly effective at reducing the extent of strategic behavior. Indeed, we show that in many cases, the value of volume flexibility is greater when consumers are strategic than when they are not. We also show that volume flexibility is always socially optimal (i.e., it increases the total welfare of the firm and consumers) and may also improve consumer welfare (i.e., it can be a Pareto improving strategy). We also discuss the impact of other types of operational flexibility—*design flexibility*, in which a product's design can be modified to suit changing consumer tastes, and *mix flexibility*, in which production capacity can be dynamically allocated amongst several similar product variants—and argue that these types of flexibility are also effective at mitigating strategic customer purchasing behavior.

1 Introduction

Accustomed to rigid seasonality and trained by years of predictable sale patterns, consumers have come to expect frequent and significant price reductions in the retail sector. As a result, many retailers suffer from eroded margins generated by customers intentionally waiting for markdowns

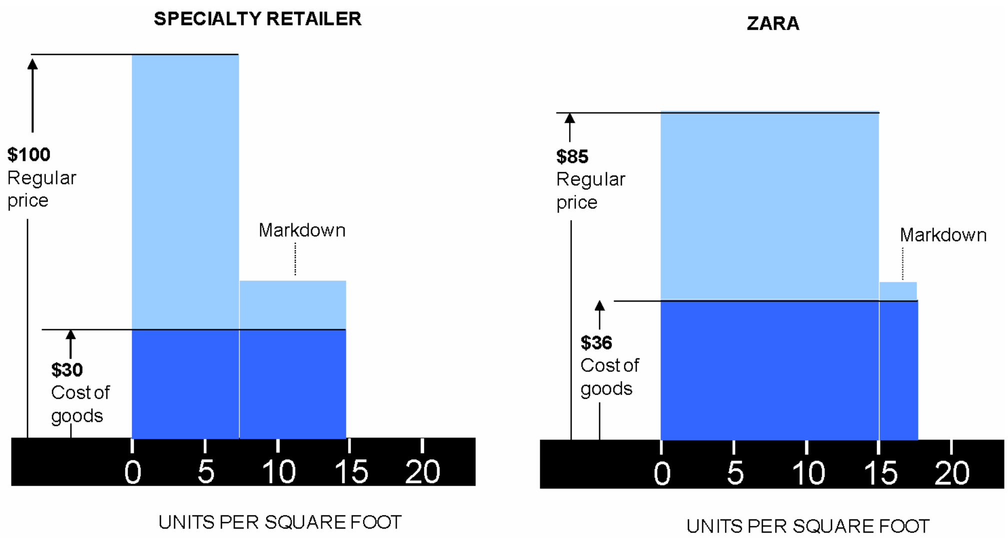

Figure 1. Pricing patterns at Zara versus competing specialty retailers. Adapted from Grichnik et al. (2008).

before purchasing (Hurlbut 2004). Consumers expect deep end-of-season clearance sales, and firms, anxious to clear space for newer products, often oblige them. There is, however, at least one firm that has achieved success at managing and even preventing strategic customer purchasing behavior: the Spanish fashion retailer Zara. There are two key components to their strategy. First, they produce in small batches with fast replenishment lead times to their stores. Second, their initial price for an item at the start of a selling season is not outrageously high. Consequently, they rarely need to markdown merchandise (because they do not stock too much inventory) and when they do offer a discount, it is not particularly deep (because their initial price is reasonable). These factors combine to train Zara's customers to avoid the "wait for the discount" strategy - if a customer sees an item that she likes, she should purchase it now either because it will be sold at the same price later on or it will not even be available. As a result, compared to its chief competitors, consumers are much more likely to purchase an item at the full price at Zara (Ghemawat and Nueno 2003).

To achieve its operational flexibility, Zara produces locally (e.g., Spain, Eastern Europe or North Africa). As a result, Zara's leadtimes are typically less than five weeks for new designs and two weeks for the replenishment of existing designs (Ghemawat and Nueno 2003). This contrasts with their competitors who can incur average design and production leadtimes of nine months. But Zara's operational flexibility comes with a cost (from, for example, higher labor costs and expedited shipping). Combined with Zara's lower initial prices, one might naturally be concerned that the company enjoys smaller gross margins per unit. However, as illustrated by Figure 1, Zara makes up for this deficit with volume: it typically sells a much higher percentage of its inventory at its full price than other retailers, which can result in superior overall financial performance.

The goal of this article is to study the Zara model to better understand its success. We begin with a model of consumer behavior first developed by Su and Zhang (2005). As in their model, we consider a single retailer who sets an initial price and makes a production decision before the realization of stochastic consumer demand. Each consumer decides whether to purchase at the initial (i.e., full) price or to wait for the discount period. The discount period offers a better deal (i.e., a lower price), but there may not be any inventory left to purchase. Hence, the scarcity of product at the discount price may make a consumer purchase at the full price.

We depart from Su and Zhang (2005) by introducing operational flexibility. With operational flexibility the firm can make a second production decision after observing demand, a system that is often called quick response (see, e.g., Fisher and Raman 1996). Of course, this second production opportunity is more expensive. However, in the absence of strategic consumer behavior, this operational flexibility is well known to benefit firms by allowing them to better match their supply to their demand. We want to assess the value of operational flexibility in the presence of strategic consumer behavior. In particular, relative to the value of matching supply with demand, does operational flexibility provide more or less value when consumers are strategic?

We find that operational flexibility is generally more valuable (but not always) when the retailer must sell to strategic consumers. Put another way, even though operational flexibility is known to increase profits considerably with non-strategic consumers (i.e., consumers that never wait for the discount no matter what prices are chosen) we show that it can be even more valuable when the firm must sell to strategic consumers, often substantially more valuable. Cachon and Swinney (2008a) arrive at a similar conclusion, but with a significantly different model. Thus, here we provide further support for the conclusion that the presence of strategic consumers enhances the value of operational flexibility.

The remainder of this chapter is organized as follows. In §2 we describe our approach to modeling production flexibility, while in §3 we discuss modeling details of strategic customer purchasing. We then solve models of non-flexible and flexible supply chains with strategic customers in §4, and discuss the incremental value of flexibility in §5. §6 presents a discussion of complications and extensions to the basic setup, and §7 concludes the chapter with a discussion of the results.

2 Modeling Traditional and Flexible Production

We refer to our base model with non-flexible production as the *traditional replenishment* model. It is also known as a newsvendor model - a canonical model in operations management that is well suited to capture the supply-demand mismatch issues inherent in fashion retailing. This model consists of a single firm selling a single product with the following key features:

1. **Constant Selling Price During the Season:** The firm sells the product at a constant ("full") price p throughout a short selling season.

2. **Demand Uncertainty:** The size of the market D (the number of consumers) is stochastic and initially unknown to the firm. The firm has prior beliefs that the market size follows distribution $F(\cdot)$.

3. **Inventory Production or Procurement:** Prior to learning the size of the market, the firm orders q units that will arrive, ready for sale, by the start of the selling season. Each unit in this order costs c, where $c < p$, and so the total purchase cost is cq.

4. **Supply-Demand Mismatches and End-of-Season Salvaging:** The firm sells the minimum of demand D and inventory q at the full price p, and all remaining inventory is salvaged at the markdown price $s < c$ at the end of the season.

We assume that a large salvage market is available at the end of the season, in which the firm may sell all remaining units at an exogenous price $s < c$ per unit. While such a market is commonly assumed in newsvendor models without further justification, in our model it may be useful to think of this market as representing a second consumer segment (beyond the initial D consumers), e.g., a large number of "bargain hunting" customers who possess very low valuations for the product. Cachon and Swinney (2008a) also incorporate a bargain hunting segment into their model.

This traditional replenishment model mimics the production environment of the majority of Zara's competitors: long design and production leadtimes lead to inventory commitment far in advance of the selling season, when precise demand is still quite uncertain. As typically presented, the newsvendor model consists of an exogenous selling price p. However, we make this price endogenous - the firm sets the price p at the start of the selling season (after demand information is revealed but before any sales occur–see Figure 2). The newsvendor model with pricing is explored by, for example, Dada and Petruzzi (1999). The nature of the pricing decision in our context is discussed further in the next section.

In contrast to the traditional replenishment system, a *flexible replenishment* model represents the system employed by Zara: greatly reduced leadtimes resulting in some inventory decisions being made very near (or during) the selling season, when demand information is far more accurate. Typically, the model employed to analyze this sort of production flexibility is a *quick response* or *reactive capacity* model (see, e.g., Cachon and Terwiesch 2005). This model is identical to the traditional replenishment model described above, with one exception: an additional procurement opportunity is available after precise market size (D) is revealed to the firm. As with the firm's initial order, units in this second order arrive by the start of the selling season. Because this second order is placed much closer to the start of the selling season, each unit in this second order costs the firm c_f, where it is natural to assume that $c_f > c$ - it is cheaper to order units in advance of learning demand.[1] (The f subscript denotes the "flexible" replenishment model.) Furthermore, like the first order, there is no capacity constraint imposed on the quantity in this second order. The sequence of events in the two models is presented in Figure 2.

The quick response framework frequently assumes that demand uncertainty is completely eliminated by the time of the second procurement–a simplification, to be sure, but one that leads to clean analytical results. In reality, the firm may receive a series of forecast updates with each reducing (but not entirely eradicating) error in the forecasting process. For the sake of simplicity, we adopt the traditional assumption that uncertainty is completely resolved.[2]

[1] We describe this model as if the second order is placed before the season starts but after some demand information is learned. In some cases, demand information is learned only at the start of the selling season and so the second order can only arrive at some point during the season. As long as initial season sales are highly informative, and the lead time to receive the second order is sufficiently short, our model can approximately represent that situation as well - the first order covers sales at the start of the season and the second order should arrive before inventory is depleted.

[2] We suspect that our results continue to hold even in a more complex setting with imperfect demand signals. In particular, even in that setting the optimal second order quantity does not depend on the full price - it is a function of c, c_f and s. Thus, our analysis would not require significant modification.

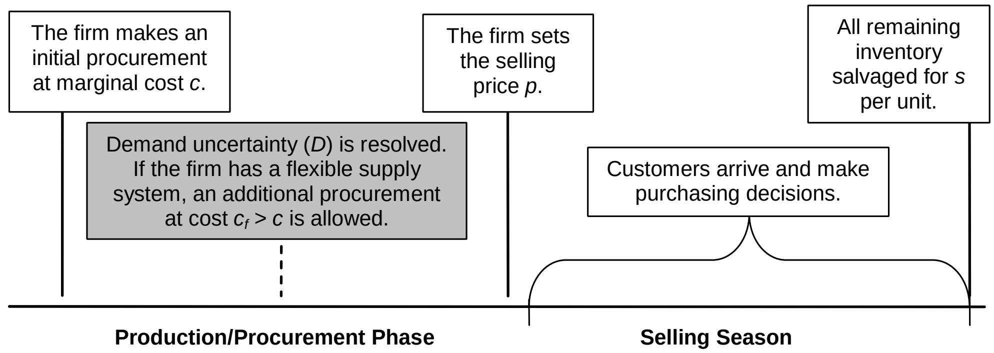

Figure 2. The sequence of events.

3 Modeling Strategic Consumer Purchasing

To address the issue of strategic customer purchasing behavior, we modify the classic newsvendor and quick response settings by enriching the consumer demand model. Suppose that consumers are risk-neutral surplus maximizers and have homogeneous valuations for the product equal to v (constant over the entire season).[3] When customers arrive at the firm, they observe the selling price p and whether the product is currently in-stock. We consider two types of customers: myopic (or non-strategic) customers and forward-looking (or strategic) customers. Myopic consumers do not consider purchasing the product at the end of the season at the markdown price, s, possibly because they no longer value the product at the end of the season or because they do not anticipate the price reduction. Myopic consumers have zero reservation utility and hence purchase if their surplus is non-negative; in other words, myopic customers purchase at price p if the product is in-stock and if

$$v - p \geq 0. \tag{1}$$

Strategic customers, on the other hand, are forward-looking in the sense that they anticipate the opportunity to purchase the product at the sale price s.[4] (Implicit in this statement is the

[3] We model risk-neutral consumers for simplicity. Risk-averse consumers behave similarly; see Liu and van Ryzin (2008).

[4] A wide variety of recent models address operational issues related to such forward-looking strategic consumers, including: intertemporal pricing with no capacity constraints in Besanko and Winston (1990); pricing policies with finite inventory in Aviv and Pazgal (2008); pricing policies for consumable goods with voluntary customer stockpiling in Su (2007); product display formats in Yin et al. (2007); and restaurant reservations in Alexandrov and Lariviere (2006); in addition to the previously cited papers.

assumption that consumers are *capable* of obtaining a unit at the salvage price–i.e., excess inventory is cleared in a way that makes it available to the general population, as with end-of-season clearances at fashion retailers, rather than alternative methods of salvaging such as material recycling or industrial disposal.) Thus, these customers compare the surplus of an immediate purchase at the full price $(v - p)$ with the expected surplus of waiting for the sale. The value of waiting depends on the discount price, s, as well as the chance there will be inventory remaining to purchase at the discount price. Let ϕ be a consumer's expectation for the probability of being able to procure a unit at the clearance price (more on the nature of this expectation will be discussed momentarily). If consumers do not obtain the product at the sale price, they receive zero surplus. Expected surplus from waiting for the sale is thus $\phi(v - s)$. We assume that strategic customers purchase at the full price if they are indifferent between the two options; hence, strategic customers purchase at price p if the product is in-stock and

$$v - p \geq \phi(v - s). \tag{2}$$

Given these assumptions, with strategic consumers there are only two candidates for equilibrium purchasing behavior: either all consumers purchase early (at the higher price) or all purchase late (at the salvage price).[5] Because the salvage price is less than the production cost, it follows that an equilibrium with all consumers strategically waiting results in market failure–the firm does not produce at all. For the remainder of the chapter, we focus on the more interesting case in which positive production occurs, i.e., equilibria in which all strategic consumers attempt to purchase at the full price. This implies that (2) is satisfied in any relevant equilibrium, and ϕ is the consumer belief of the probability of obtaining a unit at price s conditional on all other consumers purchasing at price p (i.e., $\phi(v-s)$ is the expected surplus resulting from a unilateral deviation from equilibrium by a single consumer).

Now we turn to the issue of how the ϕ expectation is set. Assuming arbitrary beliefs can lead to problems of consistency; consumers could expect ϕ to be the probability of obtaining a unit at the

[5] All consumers must have the same expectation, ϕ, value, v, and opportuntity to purchase early at the full price, p, or late at the discount price, s. Therefore either they all prefer to purchase early, $v - p \geq \phi(v - s)$, or they all prefer to wait. Here, we assume that a consumer indifferent between the two options chooses to purchase early. If the indifferent consumer chooses a mixed strategy, then the firm could shave its full price by an infinitesimal amount to make consumers strictly prefer purchasing early while not reducing revenue by a material amount.

sale price but the firm may act in a way that leads to an entirely different probability of being able to purchase at the markdown price. While it may not always be desirable to rule out inconsistency on axiomatic grounds – inconsistent beliefs may be an entirely real phenomenon with important implications – such irregularities do not appear to be the norm in the sort of predictable, seasonal industries (such as fashion apparel) that provide our prime motivation.

It is natural, then, to seek models of customer purchasing in which beliefs are consistent with reality: in other words, to specify that consumer *expectations* of firm behavior are *rational*. The idea of rational expectations – discussed in the context of financial markets by Muth (1961) – were first formally integrated in a game theoretic framework by Stokey (1981) and Bulow (1982) to explain the strategic consumer purchasing problem. In short, rational expectations imply that (a) consumers have expectations of future firm decisions, and (b) these expectations are rational and consistent with actual firm decisions. Such correct anticipation of firm actions may be thought of as the outcome of a series of repeated interactions in which consumers learn about firm policies, for instance, with regard to inventory availability ("my size is never in-stock at the end of the season") or sale pricing patterns ("this store never has deep discounts"). Our analysis in this chapter – and a great deal of the literature on strategic consumer purchasing – makes use of the rational expectations paradigm.

We note here that while the term *rational expectations* is used to highlight the fact that consumers correctly anticipate firm actions and hence behave optimally given firm actions, this concept is inherent in the definition of a Nash equilibrium with full information. For example, in a two-player game a Nash equilibrium represents a pair of actions such that each player expects the other player to choose the equilibrium actions and in equilibrium it is optimal for each person to choose their equilibrium actions given their expectations. The same applies in our game - the firm chooses an optimal q given its belief regarding consumer actions and consumers choose optimal actions (buy now or later) given their expectations. The subtle difference has to do with what is assumed regarding what the players know. Suppose we were to define the game between the firm and consumers such that the firm chooses q and consumers choose whether to purchase at the full price or to wait for the discount. Let $G(q)$ be the probability that inventory is available for a consumer to purchase in the discount period conditional that all other consumers purchase at the full price. Furthermore, assume consumers know the $G(q)$ function. A consumer's surplus from waiting to purchase at the discount price, assuming all other consumers purchase at the full price, is $G(q)(v-s)$. The consumer purchases at the full price if $v-p \geq G(q)(v-s)$. Note, to evaluate an equilibrium the consumer does not need to observe the actual q choice. Instead, the consumer infers q will be chosen because it is optimal for the firm given the consumers' equilibrium actions. In other words, consumers and the firm choose their actions simultaneously. In our model we merely replace $G(q)$ with ϕ, i.e., we assume consumers have an expectation for the probability inventory is available in the discount period without necessarily knowing the mapping between the firm's action, q, and that probability. To maintain consistency, we then require that in equilibrium $G(q) = \phi$. Therefore, the equilibrium in these two games are identical even if they make different assumptions regarding what consumers know. Put another way, the rational expectations terminology is technically unnecessary but we invoke this terminology for ease of exposition. (In addition, one may argue that it makes less stringent assumptions regarding consumer knowledge and analytical capabilities.)

Returning to our model, recall that rational expectations requires ϕ is the *actual probability* that a strategic customer successfully obtains a unit if she waits for the sale. To calculate ϕ, we must provide some sort of *rationing rule* that specifies how inventory is allocated should demand exceed supply at the salvage price. We employ the same rule as Su and Zhang (2005): strategic consumers are "first in line" at the salvage price, followed by customers from the infinite pool that makes up the salvage market. This is an appealing choice for several reasons. First, customers who arrive early in the season and intentionally choose to delay purchasing until a price reduction may closely monitor the price of the product and "pounce" once a sale occurs. Second, if we consider the infinite salvage pool to be a large group of consumers with lower valuations (i.e., with valuations equal to s), then this allocation rule maximizes consumer welfare. Third, the rule is particularly amenable to analysis, yielding closed form equilibrium solutions to the game.[6]

Given this allocation rule, what is the resulting probability that a strategic consumer obtains a unit at price s if she unilaterally deviates from an equilibrium in which all consumers purchase at price p? Such a consumer will receive a unit at the lower price if and only if the firm has enough inventory (q) to satisfy all demand (D). In other words, the probability is $\Pr(q \geq D) = F(q)$. To

[6] Alternative allocation mechanisms do exist and do not appear to substantially alter qualitative results: see Cachon and Swinney (2008a) and Swinney (2008) for random rationing rules.

connect this result with our earlier discussion, $F(q) = G(q)$, but as already mentioned, in a rational expectations framework, consumers need not be aware of the actual demand distribution function, $F(q)$.

Having described our firm and consumer models, we may now proceed to analyze the game between consumers and the firm in each of the two systems: traditional replenishment and flexible replenishment.

4 Equilibrium analysis

This section evaluates the equilibrium choices and profits for the four models constructed from the two types of replenishment modes (traditional or flexible) and the two types of consumers (myopic or strategic). (We do not consider models with a mixture of consumer types.) To help keep track of notation, we use a "t" subscript to denote "traditional replenishment" (i.e., a single order), analogous to c_f, an "f" subscript to denote "flexible replenishment" (i.e., a second order opportunity), an "m" subscript to denote "myopic consumers" , and an "s" subscript to denote "strategic consumers". For example, p_{mf} will be the firm's optimal full price with myopic consumer and flexible replenishment.

4.1 Traditional Replenishment

With myopic consumers, the traditional model resembles a newsvendor model with endogenous pricing. With strategic consumers, the traditional model mirrors the model in Su and Zhang (2005). We replicate some of their results here to ease the comparison with the flexible replenishment model in the next section.

We first observe that given the sequence of events depicted in Figure 2, and because price is directly observed by consumers, the game is essentially one of two stages. In the first stage, the firm is a Stackelberg leader in price, and in the second stage consumers and the firm play a simultaneous game in inventory and purchasing. Exploiting its status as a price leader, the firm sets the price that yields the greatest expected profit; since we focus on equilibria in which all strategic consumers purchase at the full price, this is clearly the greatest price such that either condition (1) or (2) holds, depending on whether consumers are myopic or strategic, respectively.

In other words, the optimal price with myopic consumers is

$$p_{mt} = v, \tag{3}$$

while the optimal price with strategic customers is

$$p_{st} = v - \phi(v - s). \tag{4}$$

Note that $p_{st} \leq p_{mt}$, i.e., the firm must choose a more moderate full price when selling to strategic consumers because they will purchase early only if they enjoy a positive surplus with the full price.[7]

We are concerned only with equilibria in which all strategic customers purchase at the full price, so the firm's expected profit as a function of inventory (q) and the full price (p) is

$$\pi_t(q) = \mathbb{E}\left[p \min(q, D) - cq + s(q - D)^+\right],$$

where the expectation operator, $\mathbb{E}[\cdot]$, is taken over demand D, and $(x)^+ = \max(x, 0)$. This expression provides the profit of the firm in both the myopic and strategic customer cases, with the only difference between the two being the optimal selling prices given by (3) and (4). For fixed p, this function is concave in q and possesses a unique optimum. Thus, we may immediately deduce that a firm maximizing profit in the inventory-purchasing subgame invests in an inventory level q that satisfies

$$1 - F(q) = \frac{c - s}{p - s}. \tag{5}$$

Combining (3) with (5) yields the optimal inventory level with myopic consumers, q_{mt}, which satisfies

$$1 - F(q_{mt}) = \frac{c - s}{v - s}. \tag{6}$$

In the case of strategic consumers, we note that a Nash equilibrium with rational expectations to the game between the firm and strategic consumers satisfies:

[7] Recall, we assume that strategic consumers earn value v no matter when they make a purchase. Therefore, because the discount price, s, is lower than the full price, p, the strategic consumer strictly prefers to purchase at the discount price if the item is available. She will purchase at the full price when she earns some surplus from doing so and there is a sufficiently high risk that the item will not be available in the discount period.

1. The firm prices optimally, $p_{st} = v - \phi(v-s)$;

2. The firm chooses an inventory level that maximizes expected profit, $1 - F(q_{st}) = \frac{c-s}{p-s}$;

3. Consumer expectations are rational, $\phi = F(q_{st})$.

Combining these three conditions, we see that the unique equilibrium inventory and price satisfy

$$1 - F(q_{st}) = \sqrt{\frac{c-s}{v-s}} \text{ and } p_{st} = s + \sqrt{(v-s)(c-s)}. \tag{7}$$

4.2 Flexible Replenishment

In this section we address the case of flexible replenishment. Recall that the flexible replenishment model is identical to the traditional replenishment model analyzed in the previous section, with the following exception: after learning perfect demand information (i.e., market size D), the firm has the opportunity to procure additional inventory before the season begins at a higher marginal cost $c_f > c$. As in the previous section, q represents the quantity purchased or produced in the early stocking opportunity.

It remains true that the optimal prices with myopic and strategic consumers satisfy (3) and (4), respectively: acting as a Stackelberg leader in the price game, the firm sets the greatest possible price supported by the equilibrium. The optimal procurement at the second order point is clear: if, after learning D, the firm has sufficient inventory to cover all demand ($q > D$), then the firm orders no additional inventory. On the other hand, if the firm has insufficient inventory to cover demand ($q < D$), then the firm orders precisely enough supply to perfectly match demand ($D - q$) as long as $c_f \leq p$, otherwise the firm again orders no additional inventory. Consequently, the firm's expected profit function – with either myopic or strategic consumers – is:

$$\pi_f(q) = \mathbb{E}\left[p \min(q, D) - cq + s(q-D)^+ + (p - c_f)(D-q)^+\right].$$

Or, rearranging the terms of this expression,

$$\pi_f(q) = \mathbb{E}\left[pD - cq - c_f(D-q)^+ + s(q-D)^+\right].$$

As in the traditional replenishment case, the profit function is concave in q and possesses a unique optimal inventory quantity, given by the solution q to

$$1 - F(q) = \frac{c - s}{c_f - s}. \tag{8}$$

Note that this is independent of the full price p – consequently, the initial inventory procurement is *independent* of whether consumers are strategic or myopic. This is our first important result concerning the value of a flexible replenishment system. Flexibility simplifies the firm's inventory planning duties by proving to be robust to the presence of strategic customers: misjudging or ignoring the extent of strategic customer behavior can be far less costly in a flexible replenishment system.

From (8), we immediately deduce that with myopic customers, the optimal inventory level and price in an flexible replenishment system satisfy

$$1 - F(q_{mf}) = \frac{c - s}{c_f - s}$$

and $p_{mf} = v$, respectively. Recall that with strategic consumers, a Nash equilibrium satisfies:

1. The firm prices optimally, $p_{sf} = v - \phi(v - s)$;
2. The firm chooses an inventory level that maximizes expected profit, $1 - F(q_{sf}) = \frac{c-s}{c_f-s}$;
3. Consumer expectations are rational, $\phi = F(q_{sf})$.

The second condition becomes trivial in the flexible replenishment system, as the initial inventory procurement is independent of the full price. Thus, combining the first and third conditions yields an equilibrium full price

$$p_{sf} = v - \frac{c_f - c}{c_f - s}(v - s). \tag{9}$$

Again, the second term in (9) indicates that there is a "strategic consumer penalty." If consumers are non-strategic, the firm merely charges v; due to forward-looking behavior, the firm must reduce the price to induce early purchasing.

Note that p_{sf} is decreasing in c_f - as c_f increases, the firm purchases more in advance and so the firm needs to offer a lower full price to induce strategic consumers to purchase at the full price.

In fact, when $c_f = s + \sqrt{(c-s)(v-s)}$, it follows that $p_{st} = p_{sf} = c_f$. For any greater c_f, the firm finds itself in a situation in which the second procurement opportunity is of no value because the full price, p_{sf}, is then *less* than the cost of procuring additional units. Therefore, in the strategic consumer model the firm uses flexible replenishments only when $c_f < s + \sqrt{(c-s)(v-s)}$ and does not use flexible replenishments when

$$s + \sqrt{(c-s)(v-s)} \leq c_f \leq v.$$

5 The Value of Flexibility

Armed with equilibrium results for both the traditional and flexible replenishment systems, we may now address the *value of flexibility* – that is, the increase (or decrease) in expected profit that a firm experiences when moving from a traditional to a flexible replenishment system. When discussing the value of flexibility, we have two choices for the unit of analysis: the *absolute* or the *relative* value. The absolute value refers to the incremental change in firm profit with either myopic consumers, Δ_m, or strategic consumers, Δ_s:

$$\begin{aligned}\Delta_m &= \pi_{mf} - \pi_{mt} \\ \Delta_s &= \pi_{sf} - \pi_{st}\end{aligned}$$

The relative value of flexibility, δ, on the other hand, refers to the percentage change in firm profit:

$$\begin{aligned}\delta_m &= \frac{\pi_{mf} - \pi_{mt}}{\pi_{mt}} \\ \delta_s &= \frac{\pi_{sf} - \pi_{st}}{\pi_{st}}\end{aligned}$$

Both measures can be important to a firm exploring the value of flexibility. In this section, we address both quantities, beginning with the relative value. Of primary interest are the following key questions: (1) how does the value of flexibility change when consumers are strategic, rather than myopic, and (2) what are the drivers of this change in value?

5.1 The Relative Value of Flexibility

To analyze the relative value of flexibility, we examine the behavior δ_m and δ_s as a function of c_f – the cost of a flexible replenishment. Focusing on the marginal cost of flexibility provides a natural starting point for the analysis; we intuitively expect that, either with myopic or strategic consumers, if c_f is very high, flexibility is not very valuable, while if c_f is very low, flexibility should hold more value. It is far less intuitive how the difference in relative value, $\delta_s - \delta_m$, changes with c_f.

From differentiation of δ_m we obtain

$$\frac{d\delta_m}{dc_f} = \frac{1}{\pi_{mt}} \frac{d\pi_{mf}}{dc_f} < 0 \qquad \text{and} \qquad \frac{d\delta_s}{dc_f} = \frac{1}{\pi_{st}} \frac{d\pi_{sf}}{dc_f} < 0,$$

where the inequality follows from

$$\frac{d\pi_{mf}}{dc_f} < 0 \qquad \text{and} \qquad \frac{d\pi_{sf}}{dc_f} < 0.$$

As we would expect, the relative value of flexibility decreases as the cost of flexibility increases – a natural result. Comparing this expression with strategic and myopic consumers, we have

$$\frac{d\delta_s}{dc_f} - \frac{d\delta_m}{dc_f} = \frac{1}{\pi_{st}} \frac{d\pi_{sf}}{dc_f} - \frac{1}{\pi_{mt}} \frac{d\pi_{mf}}{dc_f}.$$

Note that $\pi_{st} \le \pi_{mt}$ implies

$$\frac{1}{\pi_{st}} \ge \frac{1}{\pi_{mt}} > 0.$$

Furthermore, from the Envelope Theorem,

$$\begin{aligned}
\frac{d\pi_{mf}}{dc_f} &= -\frac{c_f - c}{c_f - s} < 0, \\
\frac{d\pi_{sf}}{dc_f} &= -\frac{c_f - c}{c_f - s} - \mu\,(v - s)\,\frac{c - s}{(c_f - s)^2} < \frac{d\pi_{mf}}{dc_f}.
\end{aligned}$$

Therefore, the relative value of flexibility decreases faster with strategic consumers than with myopic consumers:

$$\frac{d\delta_s}{dc_f} - \frac{d\delta_m}{dc_f} < 0. \tag{10}$$

The difference in the relative value of flexibility is

$$\delta_s - \delta_m = \left(\frac{\pi_{sf}}{\pi_{st}} - 1\right) - \left(\frac{\pi_{mf}}{\pi_{mt}} - 1\right).$$

Now consider a particular point, $c_f = c$, in which case $\pi_{sf} = \pi_{mf}$. Hence, for $c_f = c$, the difference in the relative value of flexibility can be written as

$$\delta_s - \delta_m = \left(\frac{\pi_{mf}}{\pi_{st}} - 1\right) - \left(\frac{\pi_{mf}}{\pi_{mt}} - 1\right) = \pi_{mf}\left(\frac{1}{\pi_{st}} - \frac{1}{\pi_{mt}}\right) > 0.$$

Therefore, when $c_f = c$, the relative value of flexibility is greater with strategic consumers than with myopic consumers ($\delta_s > \delta_m$), but (10) indicates that δ_s decreases faster with c_f than $\delta_m$does. This raises the possibility that for a large enough c_f, δ_s decreases to the point that it is less than δ_m. In fact, this occurs. Recall that flexible replenishment provides no value with strategic consumers when $c_f \geq s + \sqrt{(c-s)(v-s)}$. In that regime $\delta_s = 0$. On the other hand, $\delta_m > 0$ for all $c_f < v$. Consequently, there exists some $\hat{c}_f < s + \sqrt{(c-s)(v-s)}$ such that $\delta_s > \delta_m$ for all $c_f \in [c, \hat{c}_f)$ and $\delta_s \leq \delta_m$ for all $c_f \in [\hat{c}_f, v]$. In words, as long as the cost of the second replenishment is not too high (less than $\hat{c}_f$), flexible replenishment provides greater value with strategic consumers than it does with myopic consumers.

This result is depicted graphically in Figure 3. As the figure demonstrates, the relative value of flexibility with myopic consumers is rather flat, whereas the value with strategic consumers is strongly dependent on c_f. When c_f is small (in the figure, $c = 3$) then flexibility can offer an enormous advantage with strategic consumers, resulting in a profit increase of over 250% in the example.

The potentially large increase in the relative value of flexibility under strategic customer behavior has significant implications for how a firm evaluates a flexible supply chain. It is well established in the literature that flexible replenishment can provide significant value when consumers are myopic (see, e.g., Fisher and Raman 1996, Eppen and Iyer 1997, Iyer and Bergen 1997, and Fisher et al. 2001). Here, we find that flexible replenishment can provide substantially more value when consumers are strategic as long as the marginal cost of the second replenishment is not too high, $c_f < \hat{c}_f$. However, if the marginal cost is high ($\hat{c}_f \leq c_f$), then flexible replenishment provides little

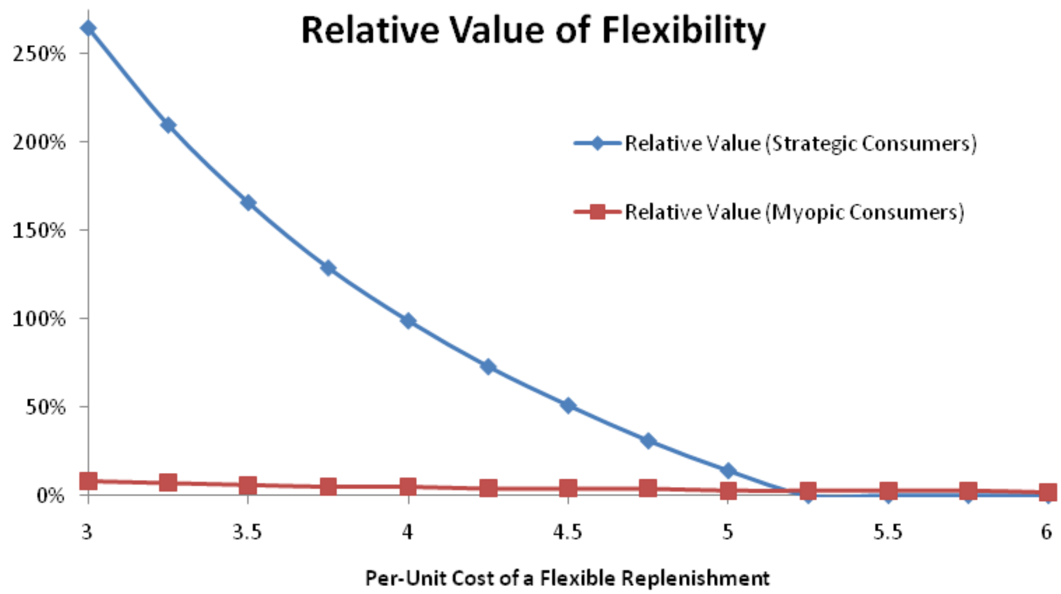

Figure 3. The relative value of flexibility as a function of the unit cost of a flexible replenishment (c_f). In this example, demand is normally distributed with mean 50 and standard deviation 10, and $v = 10$, $c = 3$, and $s = 1$.

value in a market with strategic consumers. This occurs because strategic consumers require a lower full price than myopic consumers to induce them to purchase at the full price. If the flexible replenishment system cannot deliver goods at a cost lower than the full-price, it provides no value. Referring to Figure 1, if the speciality retailer sells to myopic consumers at a price of \$100, then flexible replenishment is valuable to that retailer for any $c_f <$ \$100. In contrast, Zara sells to strategic consumers for \$85, so its flexible replenishment system provides value only if $c_f <$ \$85.

5.2 The Absolute Value of Flexibility

Despite the fact that we have shown that flexibility possesses greater relative value if consumers are strategic and c_f is not too high, it need not be the case that the analogous result holds with absolute values. To see this, suppose $\pi_{st} = 10$ and $\pi_{sf} = 20$, yielding a relative value of 100% and an absolute value of 10 under strategic behavior. If, for example, $\pi_{mt} = 100$ and $\pi_{mf} = 150$, then with myopic customers the relative value is 50% (less than with strategic customers) while the absolute value is 50 (more than with strategic customers). Hence, in the following subsection we explicitly address the *absolute value* of flexibility.

To calculate this value, note that in general, the firm's expected profit equals the expected maximum profit (i.e., the profit if the firm produces exactly at the demand level and incurs no lost sales or excess inventory) minus the expected mismatch cost, i.e., the opportunity cost of lost sales ($p-c$ per unit) plus the cost of inventory that must be sold at the discount price ($c-s$ per unit). (See Cachon and Terwiesch 2005.) Let M_Ω be the expected mismatch cost in one of our four models, $\Omega \in \{mt, mf, st, sf\}$:

$$M_\Omega = (c-s)\,\mathbb{E}\,(q_\Omega - D)^+ + \begin{cases} (p_\Omega - c)\,\mathbb{E}\,(D-q_\Omega)^+ & \Omega \in \{mt, st\} \\ (\min\{c_f, p_\Omega\} - c)\,\mathbb{E}\,(D-q_\Omega)^+ & \Omega \in \{mf, sf\} \end{cases},$$

where the first term is the cost of discounted inventory and the second term is the cost of lost sales with traditional replenishment and the cost of satisfying demand above the initial order quantity with flexible replenishment. Thus, expected profit in model Ω may be written as

$$\pi_\Omega = (p_\Omega - c)\,\mathbb{E}\,(D) - M_\Omega.$$

Using the expression for profit, the absolute value of flexibility with myopic consumers (Δ_m) is

$$\begin{aligned} \Delta_m &= (p_{mf} - c)\,\mathbb{E}\,(D) - M_{mf} - (p_{mt} - c)\,\mathbb{E}\,(D) + M_{mt} \\ &= (p_{mf} - p_{mt})\,\mathbb{E}\,(D) - M_{mf} + M_{mt} \\ &= M_{mt} - M_{mf}, \end{aligned}$$

where the latter follows from $p_{mf} = p_{mt} = v$. The absolute value of flexible replenishment with strategic consumers is

$$\begin{aligned} \Delta_s &= (p_{sf} - c)\,\mathbb{E}\,(D) - M_{sf} - (p_{st} - c)\,\mathbb{E}\,(D) + M_{st} \\ &= (p_{sf} - p_{st})\,\mathbb{E}\,(D) - M_{sf} + M_{st}. \end{aligned}$$

The difference in the absolute values of flexibility can now be expressed as

$$\begin{aligned}\Delta_s - \Delta_m &= (p_{sf} - p_{st})\,\mathbb{E}(D) - M_{sf} + M_{st} - M_{mt} + M_{mf} \\ &= (p_{sf} - p_{st})\,\mathbb{E}(D) + M_{st} - M_{mt},\end{aligned}$$

where the latter follows from $q_{sf} = q_{mf}$, which in turn implies $M_{sf} = M_{mf}$. The first term reflects the use of flexible replenishment to increase its per unit revenue: $p_{sf} \geq p_{st}$. This occurs because flexible replenishment lowers the initial order quantity, thereby lowering the availability of inventory in the discount period, thereby allowing the firm to charge a higher full price (assuming $p_{sf} \geq c_f$). The second term reflects the differences in mismatch costs with traditional replenishment.

Consider the difference in the absolute value of flexibility when $c_f = c$. In this case $p_{sf} = v = p_{mt}$, which implies

$$\begin{aligned}\Delta_s - \Delta_m &= (p_{mt} - p_{st})\,\mathbb{E}(D) + M_{st} - M_{mt} \\ &= \pi_{mt} - \pi_{st} > 0.\end{aligned}$$

Hence, the absolute value of flexibility is greater with strategic consumers when flexibility is cheap (when $c_f = c$) and therefore highly effective. Furthermore,

$$\frac{d(\Delta_s - \Delta_m)}{dc_f} = \frac{dp_{sf}}{dc_f}\mathbb{E}(D) < 0,$$

and so the difference in absolute value of flexibility decreases as c_f increases. Thus, we have established the same pattern as with the relative value of flexibility: Δ_s is initially greater than Δ_m (for $c_f = c$) but decreases as flexibility becomes costlier. For large enough c_f, we have established that $\Delta_s = 0$ (because then $c_f > p_{sf}$) while Δ_m remains positive. Thus, for some value $\bar{c}_f$ we have $\Delta_s > \Delta_m$ for all $c_f < \bar{c}_f$ and otherwise $\Delta_s \leq \Delta_m$. This pattern is illustrated in Figure 4. In this example the absolute value of flexibility can be substantially greater with strategic consumers than with myopic consumers, upwards of 10 times more valuable.

Figures 3 and 4 also illustrate that relative and absolute values need not correspond. For example, for the range $5 < c_f < 5.25$ we see that the relative value of flexibility is greater with strategic

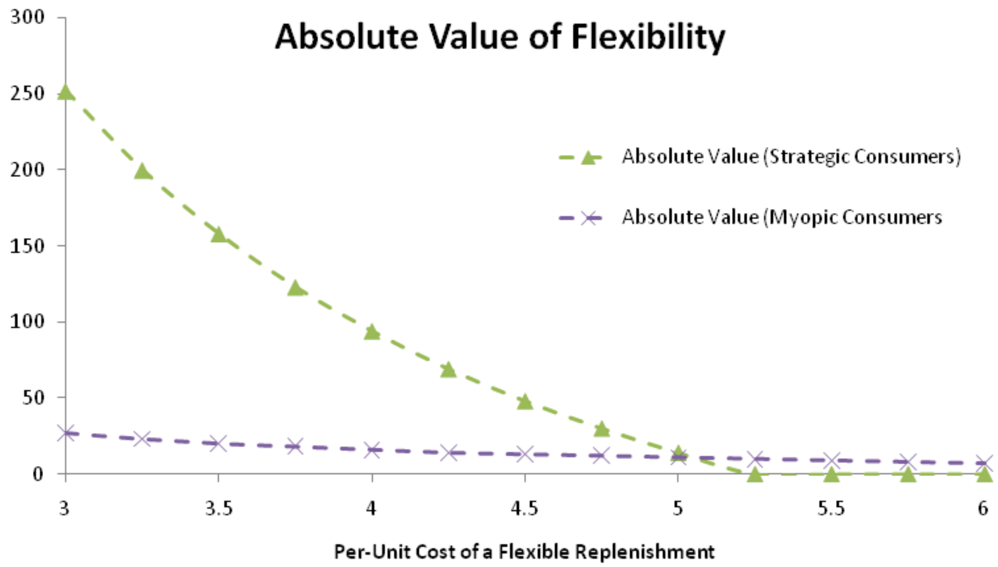

Figure 4. The absolute value of flexibility as a function of the unit cost of a flexible replenishment (c_f). In this example, demand is normally distributed with mean 50 and standard deviation 10, and $v = 10$, $c = 3$, and $s = 1$.

customers but the absolute value of flexibility is greater with myopic consumers. Nevertheless, in the range $c_f < 5$, both the relative and absolute values of flexibility are greater with strategic consumers.

To summarize, we find that not only does flexibility often provide greater *relative* value when consumers are strategic, it can also provide greater *absolute* value. This in an important result for firms considering implementing a flexible supply system: if their customer base is strategic, then flexibility can provide enormous additional benefits over the myopic customer case, giving greater justification to spending the high fixed costs associated with implementing a flexible system.

5.3 Drivers of the Value of Flexibility

What causes the potentially large increase in the value – both relative and absolute – of flexibility under strategic customer behavior? The key lies in two distinct consequences of flexibility: matching supply with demand, and reducing strategic behavior.

The value to the firm of better matching supply to demand is present regardless of the type of customer population it faces. Flexibility in our model eliminates lost sales (assuming $c_f \leq p$). Thus, the firm uses flexibility to lower its initial purchase quantity, which reduces the cost of excess inventory that needs to be marked down to the salvage price. However, the full price with strategic consumers is lower than the full price with myopic consumers. Hence, the value of eliminating lost sales is actually lower with strategic consumers than with myopic consumers. In other words, the value of better matching supply with demand is higher when the full price is higher. If only this effect were present, we would conclude that flexibility is more valuable with myopic consumers than with strategic consumers. But there is a second effect.

The full price with myopic consumers is independent of whether the firm possesses flexibility or not, i.e., $p_{mt} = p_{mf} = v$. However, with strategic consumers, adding flexibility allows the firm to increase the full price, $p_{st} < p_{sf}$, because flexibility reduces the initial order quantity, $q_{sf} < q_{st}$. With less initial inventory, the firm is less likely to sell any inventory at the salvage price and so strategic consumers are willing to pay a higher full price. Therefore, adding operational flexibility allows the firm to earn a higher price on *all* sales. This effect can dominate the former - while flexibility helps to reduce lost sales and excess inventory, it can be more valuable to use flexibility to increase revenue on all regular season sales. Our numerical studies indicate that not only can this effect dominate, it tends to dominate by a considerable amount over a large range of parameters. Therefore, while we cannot conclude that operational flexibility is always more valuable with strategic consumers, we find that operational flexibility is generally more valuable.

5.4 Consumer and Social Welfare

Flexibility results in higher prices when consumers are strategic, so an immediate concern is that operational flexibility results in decreased consumer (and possibly social) welfare. We define consumer welfare to be the total surplus of the customer population, i.e., the surplus of each individual who successfully obtains a unit times the expected number of sales. Observe that with myopic consumers, the firm extracts all surplus in either replenishment system, resulting in zero consumer surplus with either replenishment system.

With strategic consumers, however, this is not the case. The surplus of an individual consumer who obtains a unit is $v - p$, where p is the full price. In the traditional replenishment model, the resulting total equilibrium consumer surplus is

$$\left(v - s - \sqrt{(v-s)(c-s)}\right) \mathbb{E}\left(\min\left(q_{st}, D\right)\right).$$

In this system, consumers pay a low price (so individual surplus is high) but not all consumers are served. In the flexible replenishment model (again, assuming $c_f \leq p_{sf}$), total surplus is

$$\frac{c_f - c}{c_f - s} (v - s) \mathbb{E}(D).$$

With flexibility, consumers pay a higher price (so individual surplus is lower) but all consumers are ultimately served. Therefore, it is not clear whether flexibility increases or decreases consumer surplus.

Let $\eta = \mathbb{E}(\min(q_{st}, D)) / \mathbb{E}(D)$, which is the fill rate (fraction of demand that is fulfilled) with traditional replenishment We can now write an expression for when consumer surplus is greater with flexible replenishment than with traditional replenishment:

$$\frac{c_f - c}{c_f - s} \geq \eta \left(1 - \sqrt{(c - s) / (v - s)}\right).$$

The left hand side is increasing in c_f. If we let c_f equal its maximum feasible value, $c_f = s + \sqrt{(c - s)(v - s)}$, then the above expression can be written as

$$\left(v - s - \sqrt{(v - s)(c - s)}\right) \geq \eta \left(v - s - \sqrt{(v - s)(c - s)}\right),$$

which always holds (given that $\eta < 1$). Therefore, as long as c_f is sufficiently large, flexibility increases consumer surplus when selling to strategic consumers. This pattern is illustrated in Figure 5.

Also of interest is the impact of flexibility on *social* welfare, i.e., the sum of firm and consumer surplus. Combining consumer welfare with expected firm profit, we see that social welfare in the traditional system is

$$W_{st} = \mathbb{E}\left[v \min(q_{st}, D) - c q_{st} + s (q_{st} - D)^+\right],$$

while social welfare in the flexible system is

$$W_{sf} = \mathbb{E}\left[v \min(q_{sf}, D) - c q_{sf} + (v - c_f)(D - q_{sf})^+ + s (q_{sf} - D)^+\right].$$

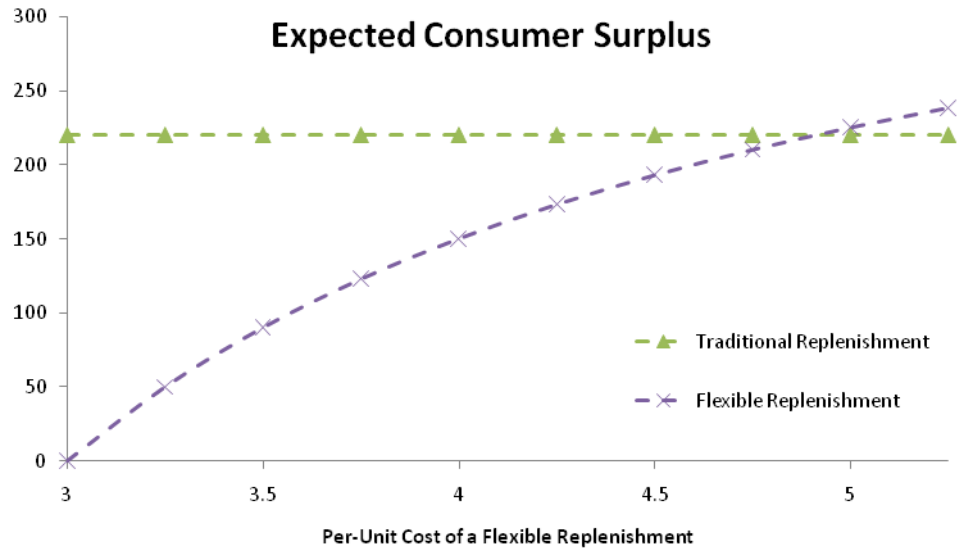

Figure 5. Expected consumer surplus as a function of c_f. In this example, demand is normally distributed with mean 50 and standard deviation 10, and $v = 10$, $c = 3$, and $s = 1$.

Note that if $c_f = c$,

$$W_{sf} = \mathbb{E}\left[(v-c)D\right] \geq \mathbb{E}\left[v\min\left(q_{st}, D\right) - cq_{st} + s\left(q_{st} - D\right)^{+}\right] = W_{st}.$$

Alternatively, if $c_f = v$, $q_{st} = q_{sf}$ and

$$W_{sf} = \mathbb{E}\left[v\min\left(q_{sf}, D\right) - cq_{sf} + s\left(q_{sf} - D\right)^{+}\right] = W_{st}.$$

Since $dW_{sf}/dc_f < 0$, it follows that $W_{sf} \geq W_{st}$, i.e., flexible replenishment is socially optimal for all viable c_f. Thus, flexibility never decreases social welfare–see Figure 6. We conclude that while flexibility may not always be in the best interests of the individual consumer, it is socially optimal (and may even be Pareto optimal).

6 Extensions and Complications

In this section we discuss the impact of a variety of extensions and complications to the simple model analyzed thus far. In each of the three extensions, just as in the preceding model, the interaction of operational flexibility with strategic customer behavior is addressed. The first considers the

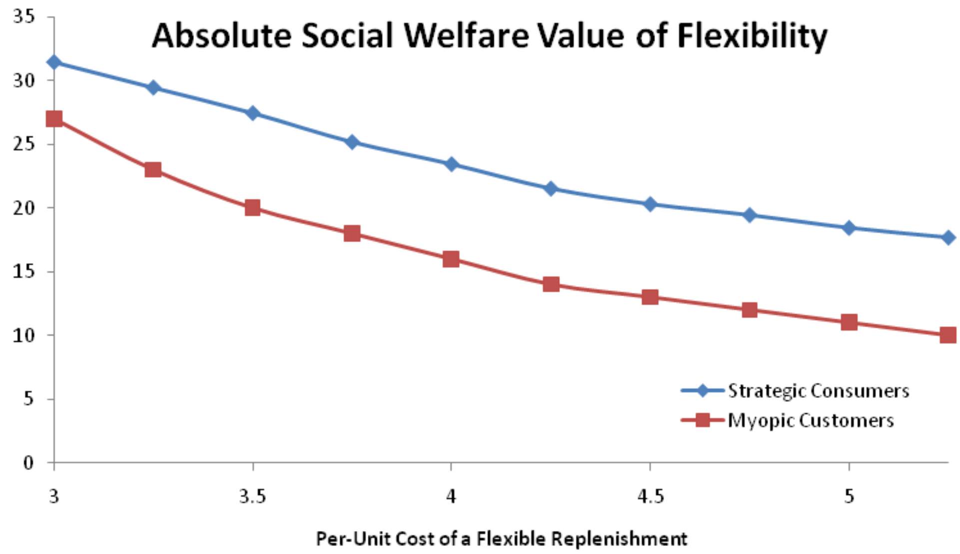

Figure 6. The value of flexibility (in terms of social welfare) as a function of c_f. In this example, demand is normally distributed with mean 50 and standard deviation 10, and $v = 10$, $c = 3$, and $s = 1$.

impact of dynamic markdown pricing (rather than static pricing as we assume in our model) and consumers with heterogeneous valuations. The second explores the consequences of consumers that do not know (but learn) their valuations over time. The third discusses alternative forms of operational flexibility, and considers their impact on strategic customer behavior.

6.1 Dynamic Sale Pricing and Consumer Heterogeneity

There are two key simplifications present in our the model: the end-of-season clearance price is exogenously determined and *ex ante* fixed, and consumers are homogeneous both in the degree to which they are strategic (i.e., they are either all myopic or all strategic) and in their valuation for the good.

Suppose the firm is allowed at the end of the season to choose to keep the full price, p, or to lower the price to s to clear inventory in the infinite salvage market. This does not, in the context of homogeneous consumers, alter the analysis because all strategic consumers purchase early in equilibrium and so the firm always lowers the price to s at the end of the season to clear inventory. Hence, a consumer unilaterally deviating from equilibrium finds the product for sale at price s during the clearance period, precisely as she would in our static pricing model.

Imagine, however, that consumers are heterogenous in their valuation for the product, possessing, for example, uniform valuations in a fixed interval. Because consumers are no longer homogeneous, it need not be the case that all consumers purchase at the full price in a viable equilibrium; indeed, it can be shown that in equilibrium, consumers with high valuations purchase early while consumers with low valuations purchase later. If the firm is free to set a clearance price, the equilibrium division of consumers according to valuations will induce the firm to *price skim*–set a high price for higher value customers that purchase earlier, and set a lower price for lower value customers who purchase later. If a large "bargain hunting" customer segment exists (e.g., the infinite salvage market) then the firm may still lower the price to s if inventory during the clearance phase is significant (i.e., if demand during the full price phase is low). Thus, a dynamic clearance price is a function of stochastic demand and the equilibrium number of consumers who purchase at the full price. When making their purchasing decision consumers must consider all possible future sale prices to calculate their expected surplus of waiting until the clearance sale.

What is the impact of this richer model on our results concerning supply chain flexibility? Because flexibility lowers the amount of excess inventory, it decreases the chances that a firm will have to set a deep discount to clear inventory during the end-of-season sale. It also ensures that the firm has adequate inventory to cover demand if the product is a "hit" and prices are high. Flexibility thus has two effects: it *reduces supply-demand mismatch* (just as in our simple model) and it *increases the clearance price*. By increasing the expected clearance price, the firm is able to encourage more consumers to buy at the full price; why wait for a sale if the savings are not very significant? In turn, this allows the firm to set a higher full price and reap greater demand at the full price. The net result is that flexibility may possess even greater value when consumers have heterogenous valuations and sale prices are endogenously determined–in addition to the benefits discussed in our preceding analysis, the firm gains additional value from higher prices at the end of the season.

We might also imagine other models of consumer heterogeneity, for instance heterogeneous rates of consumption (e.g., different discount rates) or heterogeneous degrees of foresight (some myopic consumers mixed in with strategic consumers). The intuition in such models is similar: with dynamic pricing, flexibility helps to raise the average clearance price and thus encourage more consumers to purchase at the full price – see Cachon and Swinney (2008a). This, in turn, enhances the value of flexibility beyond that which we derived in our simple model. We therefore conjecture that the benefits of flexibility in helping a firm cope with strategic behavior are robust to complications involving heterogeneous customer populations and various pricing schemes.

6.2 Uncertain Consumer Valuations and Learning

In the model analyzed in §§2–5, consumers are fully informed concerning their valuations at the start of the game. While this may be true for generic goods or products with relatively simple attributes that are easily analyzed (e.g., clothing), consumers may not initially know how much they value some products. Examples include innovative or complex products, products with long life cycles reliant on secondary goods of uncertain quality (e.g., video game systems), and even experience goods. With many of these products, information about value is disseminated over time to the customer population as, for example, expert product reviews are published, secondary goods are released, and consumers experience the product's features via units purchased by friends or in-store demonstrations. See Swinney (2008) for an analysis of this problem, the key results of which we summarize here.

With products of this type, consumers have an added benefit to delaying a purchase: to gather more information about the product's value to them. In such a setting, it is possible for a flexible supply chain to actually *decrease* a firm's profit (even in the absence of a fixed costs to implement a flexible system) – by operating with an agile supply system capable of meeting future demand, the firm increases the overall availability of the product, thereby minimizing rationing risk and increasing consumer incentives to learn as much information about product value as possible before purchasing. This effect can be called *demand shifting* – by increasing availability, the firm causes customers to purchase later.

When this occurs, overall demand to the firm can actually decrease. This is because in selling to consumers before valuations are learned – also know as *advance selling*, see Xie and Shugan (2001) – the firm inevitably induces some consumers to purchase the product who ultimately will not value it. If customers are encouraged to delay their purchase, some of these "false positives" are eliminated from the firm's demand, thereby decreasing profit.

There are cases, however, in which shifting demand (and subsequently reducing advance selling) benefits the firm. If, for instance, prices increase over time (e.g., due to promotional discounts during new product introduction) or if unsatisfied customers are allowed to return products for full refunds, selling to many customers early (before value is fully learned) can actually harm the firm; in these cases, supply chain flexibility – which reduces the extent of advance selling – helps the firm by increasing the number of sales at a higher, later price or by decreasing the number of false positive purchases resulting in costly product returns.

These results imply that flexibility can provide both positive and negative value when a firm sells a product for which consumers have uncertain valuations. Thus, firms must carefully consider the nature of their product and the ease with which consumers can judge value when choosing their own supply chain structure. A key implication of this result is that it is critical for the operations side of the firm to work closely with marketing, design, and development groups to properly ascertain the characteristics of the consumer population and their interaction with the product.

6.3 Alternative Forms of Flexibility

This discussion has focused on a particular form of supply chain flexibility: the ability to rapidly procure additional inventory to meet demand, also know as volume flexibility or quick response. There are, however, other forms of operational flexibility that may be analyzed, two of which we discuss here: *design flexibility* and *mix flexibility* (also know as postponement).

Design flexibility refers to the ability to modify or create new product designs close to the start of the selling season in order to capture evolving and uncertain consumer trends. This type of flexibility is a crucial part of Zara's philosophy – by vastly reducing both design and production leadtimes, Zara can create styles that are more suited to consumer tastes *and* produce inventory that more closely matches its demand. Essentially, such practices serve to increase overall consumer value for a product, thereby increasing consumer willingness-to-pay. By giving consumers more valuable products, strategic behavior is lessened; customers are less willing to wait for a sale and risk a stock-out if they highly value the item. Furthermore, design flexibility and supply flexibility are complimentary in nature: higher value products increases the value of matching supply and demand, and less supply-demand mismatch increases the benefit of raising consumer value for a product. Thus, the combination of both types of flexibility–often referred to as a *fast fashion* system by Zara–results in a superadditive increase in firm profit. See Cachon and Swinney (2008b) for more on the effect of design flexibility on strategic customer behavior.

Mix flexibility or postponement refers to the ability of a firm to dynamically allocate capacity between two or more different products. Suppose, for example, a firm sells to a market of fixed size N but with uncertain aggregate preferences between two product variants: a fraction β prefers variant 1 while a complementary fraction $1-\beta$ prefers variant 2, where β is *ex-ante* stochastic to the firm. Consumers know their private preference between variants, and just as in our previous models, the product is sold at a high price during the selling season and cleared at a low price at the end of the season. A firm without mix flexibility must make inventory decisions prior to the revelation of market preferences (β), essentially solving two (correlated) newsvendor problems, resulting in similar consumer incentives and strategic purchasing to the model we analyzed in this chapter.

A firm with mix flexibility, however, may pre-manufacture a common base product, while postponing final assembly into specific variants until after β is learned. If the firm has mix flexibility, then it will clearly be optimal to produce exactly N units of the base product and, after learning β, allocate final assembly such that the supply of each variant perfectly matches demand. Consequently, no sales occur at the salvage price, and there is no chance for consumers to obtain a unit at the clearance sale.[8] Strategic behavior is hence completely eliminated. While this simple model provides overly sharp results, the basic intuition supports our conclusions that operational flexibility in general benefits the firm by reducing strategic behavior.

7 Conclusions

In this chapter, we show how techniques for generating operational flexibility – long thought to be valuable solely by virtue of matching supply with uncertain demand – can have an enormous impact on customer purchasing behavior and pricing. This impact is almost always beneficial to the firm, and indeed can result in the value of flexibility being substantially greater when consumers are strategic relative to when they are non-strategic (i.e., myopic).

These results help to refine and strengthen our understanding of how "fast fashion" firms such

[8] To be precise, this depends on whether consumers are atomistic. If consumers are atomistic (i.e., they do not consider the impact that their own behavior has on quantities like availability), then there is zero availability at the clearance price if they delay purchasing and the firm produces exactly to the level of demand. If consumers do consider their own impact on product availability, then this may not be true–however, in this case, the firm may react by reducing inventory by some small amount (e.g., one unit) thereby restoring zero availability at the clearance price.

as Zara have achieved so much success in an industry facing an increasingly savvy and strategic customer base. By producing inventory much closer to the start of the selling season, Zara is able to generate and utilize more precise demand forecasts than its competitors. Exploiting the increased precision of these demand forecasts, Zara is able to reduce the likelihood of drastically over-producing a given product, which in turn reduces the chance and magnitude of a potential markdown at the end of the season. In short, Zara exploits its "fashion on demand" capabilities to limit the extent of season-ending sales, thereby lowering the incentive for consumers to strategically delay purchases.

The key innovation in this work is to study the interaction between a firm's operational strategy and consumer behavior. This analysis leads to new insights into the value of operational flexibility as well as to insights on a firm's optimal pricing strategy. We feel there are many other opportunities to further explore and develop models that refine the dependency between consumer behavior and operations, not just in procurement and supply chain management but in all aspects of operations. By addressing such models, our hope is that a more complete picture of the impact of operating practices emerges, one that addresses not just firms and their suppliers, but also another crucial member of the supply chain - consumers.

References

Alexandrov, A., M. A. Lariviere. 2006. Are reservations recommended? Working paper, Northwestern University.

Aviv, Y., A. Pazgal. 2008. Optimal pricing of seasonal products in the presence of forward-looking consumers. *Manufacturing Service Oper. Management* **10**(3) 339–359.

Besanko, D., W. L. Winston. 1990. Optimal price skimming by a monopolist facing rational consumers. *Management Sci.* **36**(5) 555–567.

Bulow, J. I. 1982. Durable-goods monopolists. *The Journal of Political Economy* **90**(2) 314–332.

Cachon, G. P., R. Swinney. 2008a. Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Sci.* Forthcoming.

Cachon, G. P., R. Swinney. 2008b. Using a fast fashion system for rapid product design with strategic consumers. Working paper, University of Pennsylvania.

Cachon, G. P., C. Terwiesch. 2005. *Matching Supply with Demand: An Introduction to Operations Management*. McGraw-Hill/Irwin.

Dada, M., N. Petruzzi. 1999. Pricing and the newsvendor model: A review with extensions. *Oper. Res.* **47** 183–194.

Eppen, G. D., A. V. Iyer. 1997. Improved fashion buying with Bayesian updating. *Oper. Res.* **45**(6) 805–819.

Fisher, M., K. Rajaram, A. Raman. 2001. Optimizing inventory replenishment of retail fashion products. *Manufacturing Service Oper. Management* **3**(3) 230–241.

Fisher, M., A. Raman. 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* **44**(1) 87–99.

Ghemawat, P., J. L. Nueno. 2003. ZARA: Fast Fashion. Case Study, Harvard Business School.

Grichnik, K., C. Winkler, J. Rothfeder. 2008. *Make or Break: How Manufacturers Can Leap from Decline to Revitalization*. McGraw-Hill.

Hurlbut, T. 2004. The markdown blues. Inc.com.

Iyer, A. V., M. E. Bergen. 1997. Quick response in manufacturer-retailer channels. *Management Sci.* **43**(4) 559–570.

Liu, Q., G. van Ryzin. 2008. Strategic capacity rationing to induce early purchases. *Management Sci.* **54**(6) 1115–1131.

Muth, J. F. 1961. Rational expectations and the theory of price movements. *Econometrica* **29**(3) 315–335.

Stokey, N. L. 1981. Rational expectations and durable goods pricing. *The Bell Journal of Economics* **12**(1) 112–128.

Su, X. 2007. Inter-temporal pricing and consumer stockpiling. Working paper, University of California, Berkeley.

Su, X., F. Zhang. 2005. Strategic consumer behavior, commitment, and supply chain performance. Forthcoming, Management Science.

Swinney, R. 2008. Selling to strategic consumers when product value is uncertain: The value of matching supply and demand. Working paper, Stanford University.

Xie, J., S. M. Shugan. 2001. Electronic tickets, smart cards, and online prepayments: When and how to advance sell. *Marketing Sci.* **20**(3) 219–243.

Yin, R., Y. Aviv, A. Pazgal, C. S. Tang. 2007. Optimal markdown pricing: Implications of inventory display formats in the presence of strategic customers. Working paper, Arizona State University.