Selling to Strategic Consumers When Product Value is Uncertain: The Value of Matching Supply and Demand

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

May, 2008. Revised January, 2009.

Abstract

Quick response inventory practices–which combine reduced production leadtimes, sophisticated information systems, and continuous demand forecasting improvement–are often discussed as a potential remedy to the negative aspects of supply and demand mismatches. We address the value of these practices when selling to a forward-looking consumer population with uncertain, heterogeneous values for a product. Consumers have the option of purchasing the product early, before its value has been learned, or delaying the purchase decision until a time at which valuation uncertainty has been resolved, a trade-off frequently made by consumers shopping for new or innovative products. While individual consumer valuations are uncertain *ex ante*, the market size is uncertain to the firm. The firm may either commit to a single production run at a low unit cost prior to learning demand, or commit to a quick response strategy which allows additional production (at a higher unit cost) after learning additional demand information. We find that the value of quick response is generally lower with strategic (forward-looking) customers than with non-strategic (myopic) customers in this setting. Indeed, it is possible for a quick response strategy to decrease the profit of the firm, though whether this occurs depends on various characteristics of the market; specifically, we identify conditions under which quick response increases profit (when prices are increasing, when dissatisfied consumers can return the product) and conditions under which quick response may decrease profit (when prices are constant or when consumer returns are not allowed).

1 Introduction

In October of 2007, Susan faced a dilemma. Her three year-old son, Ryan, had recently developed a strong interest in the Little Einstein line of toys produced by Fisher-Price. Susan knew that one particular toy–the Pat-Pat Rocket–was rumored to be a "hot toy" for toddlers during the 2007 holiday shopping season. By chance, Susan one day came across a store with several Pat-Pat Rockets in stock. She knew that if she purchased the toy immediately, there was no guarantee that Ryan would still enjoy Little Einstein products in three months time; his taste in toys seemed to change weekly. If she did not buy the toy now, however, she believed her chances of finding it in the future may be slim, particularly if the toy turned out to be a hot holiday gift. Susan ultimately chose to purchase the Pat-Pat Rocket; the risk of not obtaining the toy was too great to delay her decision until closer to the holidays.

Parents increasingly participate in this unfortunate holiday ritual (Slatalla 2002), trading off the risk of buying early and facing uncertain value for the product (i.e., possibly buying a toy that turns out to be a "dud" or that their child does not want) with the risk of buying late and facing uncertain availability for the product (i.e., experiencing a stock-out). Recent history provides numerous examples of hot holiday toys for which demand outstrips expectations, from Barbie to Elmo to the Nintendo Wii video game system. Stock-outs during the holidays are assumed to be particularly costly to firms, as consumers shopping for gifts are more likely to switch to a competitor's product rather than wait for an inventory replenishment that occurs after the holidays have concluded. Indeed, the perception is that potential losses due to inventory shortages during the holidays can be enormous–Richtel (2007), for example, reports estimates that Nintendo experienced lost sales in excess of $1 billion due to unsatisfied demand on the Wii video game system during the 2007 holiday season.

Long production and shipping leadtimes are often cited as key causes for holiday gift shortages, particularly on those products manufactured in Asia and exported to the US or Europe. Due to these long leadtimes, demand forecasts must be made far in advance of the selling season, when uncertainty concerning final demand is high. Thus, if leadtimes could be reduced–via, for example, localized production, increased capacity, improved information systems, and expedited shipping methods–allowing for a rapid response to updated demand information closer to (or during) the selling season, supply and demand could be more closely matched, reducing or eliminating costly shortages. Such techniques (often known as *quick response* systems) can be expensive due to expedited production or transportation costs, but are known to provide significant value to firms by better matching supply with uncertain demand (Fisher and Raman 1996). The consensus is that quick response systems are beneficial to a firm: indeed, in the absence of fixed costs related to the implementation of such systems, the opportunity to procure additional inventory after learning updated demand information is an option which always possesses positive value.

In this paper, we consider whether (and under what conditions) quick response inventory practices do in fact benefit a firm. Motivated by our example of holiday gifts, we consider a product characterized by initially uncertain consumer value. This may be the case if, for instance, the product is a new or innovative item (e.g., a complex or innovative product such as a Nintendo Wii, an Apple iPhone, or an automobile), a media item (such as books, movies, music, or video games), or the consumer's requirements for the item are uncertain (e.g., snow skis for a potential weekend trip in two months when weather is unknown, or a gift for a child whose preferences frequently change). Over time, consumers learn more information about the product and gain a better sense of its value; for example, via channels such as professional product reviews from web sites and magazines, the reviews of fellow consumers (e.g., from online retailers such as Amazon.com), the experiences of friends and family who may have purchased the same product, or via the resolution of intrinsic uncertainty in product value (e.g., the weather affecting the value of a pair of skis is known the day of the ski trip). Individual consumers thus make a decision on *when* and *whether* to purchase the product: the later the customer waits to buy, the more information she will have about product value and the greater the risk of a stock-out.

When consumers experience this type of time dependent learning, greater availability resulting from an improved matching of supply and demand encourages consumers to delay purchasing the product: by reducing the likelihood of a stock-out, the firm decreases the riskiness of waiting to learn more information about product value. Thus, there is a clear interaction of consumer learning (of product value) and firm learning (of product demand). We explore this interaction by addressing how the responsiveness of the firm's supply chain–its ability to respond to improved demand information–affects consumer purchasing behavior, and vice-versa. We analyze models with a single firm selling to a rational, forward-looking consumer population. Consumers choose to either purchase early–prior to learning their value for a product–or purchase late, after learning their value. The firm chooses to either commit to a single production run in advance of learning product popularity or to adopt the ability to rapidly produce inventory after stochastic demand is revealed.

Using this stylized framework, we demonstrate that quick response generally possesses less value when the firm faces strategic consumers that learn about value over time than when it faces non-strategic consumers. Furthermore, the basic intuition that quick response provides an option with purely positive value may be incorrect; that is, even without fixed costs it is possible for a firm to be *worse off* if it has an additional procurement opportunity after receiving updated demand information. This occurs when consumers–cognizant of the results of the firm's operating policies, in particular the inventory availability–modify their own purchasing behavior to account for the implementation of a quick response system. In other words, while quick response *does* better match supply and demand, demand itself can be negatively affected once consumers become aware of the increased availability resulting from quick response and optimize their own behavior accordingly. The net effect may decrease firm profits, though we demonstrate that whether this occurs (and to what degree it occurs) depends heavily on several characteristics of the selling environment; specifically, when prices increase over time or when dissatisfied consumers can return the product for a full refund, quick response always increases firm profit, whereas if prices are constant or decline over time or if consumers cannot return the product for a full refund, quick response may decrease firm profit.

The remainder of the paper is organized as follows. §2 provides a review of the literature. §3 introduces the model, and §§4–5 analyze the consumer decision and the equilibrium to the game, respectively. Two extensions are then analyzed: consumer returns in §6 and pricing in §7. §8 concludes the paper with a discussion of the results.

2 Related Literature

There are two areas of the literature that are of particular relation to this analysis: the first concerns forward-looking or "strategic" consumer behavior. Explicitly modeling the intertemporal purchasing decision of rational consumers has received increased attention in recent years; see, for example, Aviv and Pazgal (2008), Liu and van Ryzin (2008), Su and Zhang (2008), and Jerath et al. (2007). The most relevant papers to our own from this stream of research are those addressing uncertain consumer valuations. DeGraba (1995) explains why a firm may intentionally understock to induce consumers to purchase when valuations are uncertain and learned over time. Unlike our model, there is no demand uncertainty to the firm. Xie and Shugan (2001) demonstrate that selling to consumers prior to the determination of value and consumption (e.g., with advance ticket sales) can substantially increase firm profits. Alexandrov and Lariviere (2006) consider the problem of a restaurant choosing whether to offer reservations (guaranteed seats) to customers who may or may not value dining on a given night, demonstrating when reservations increase the profit of the firm. Dana (1998) and Akan et al. (2007) discuss optimal pricing to screen heterogeneous consumers whose values are revealed over time. In these papers, in contrast to our model, inventory (or capacity) is either infinite, exogenously set, or fixed throughout the selling season, and hence issues of inventory replenishment after receiving updated demand information are not considered. An exception is Debo and van Ryzin (2007), who consider a periodic review inventory problem. However, in their model the base-stock level is exogenously given, as is the decision of how often to replenish, whereas in our model, inventory levels and the decision of whether to obtain additional inventory are endogenously determined.

Uncertain consumer valuations are also a hallmark of the herding literature–see, e.g., Bikhchandani et al. (1992). Typically in this literature, the actions of the firm are fixed, while consumers observe the sequence of sales and use Bayesian updating to determine whether to purchase a product with uncertain value. An exception is Debo (2007), which incorporates both the consumer learning characteristics of the herding literature and the firm's decision to price optimally with fixed inventories. In contrast to the herding literature, our model does not allow consumers to observe the sequence of sales of the product and infer value from this information; rather, valuations are exogenously revealed over time, e.g., via external channels such as expert product reviews. We abstract from the consumer learning dynamics of the herding literature to focus instead on the inventory game between the firm and consumers, as well as the decision to adopt quick response inventory practices.

In §6, we address consumer returns policies which allow customers who buy prior to learning their value to return the product should their realized valuation turn out to be low. Such policies have received attention in the literature: Davis et al. (1995) and Moorthy and Srinivasan (1995) analyze the value of money-back guarantees when selling to consumers with uncertain value; Su (2009) provides an analysis of how consumer product returns affect inventory decisions when valuations are learned after the purchase of an item (e.g., experience goods); Liu and Xiao (2008) analyze returns policies with heterogeneous customers; and Coughlan et al. (2007) address the role of returns policies in competitive retail settings. These papers do not consider the impact of consumer returns policies on a firm's incentives to adopt a rapid procurement strategy, however.

The second broad stream of research related to our own is the quick response literature–see, for example, Fisher and Raman (1996), Eppen and Iyer (1997), Iyer and Bergen (1997), and Fisher et al. (2001). In the absence of strategic consumer behavior, these papers demonstrate the value engendered by the ability of a firm to react quickly to updated demand information. In a related paper (Cachon and Swinney 2009), we address the value of quick response systems in a fashion retail setting with forward-looking consumers. The model in Cachon and Swinney (2009) is characterized by markdowns and known consumer valuations, and consumers may strategically delay purchasing in order to "get a good deal" when the item goes on sale. The present model, by contrast, is characterized by constant (or increasing, as discussed in §7) prices and unknown consumer valuations–as a result, consumers delay purchasing to obtain better information about product value. Consequently, while Cachon and Swinney (2009) is applicable to settings in which value is easily judged and markdowns are likely (e.g., fashion), the present analysis is applicable to more complex products in which value is difficult to judge or inherently stochastic.

3 Model

A firm sells a single product at an exogenous price[1] p to a consumer population of size N over a single selling season. There are two potential production opportunities for the firm: early production (far in advance of the selling season) and late production (very close to the start of the season). Early production is far enough in advance of the season that market size is unknown, though the firm does possess some forecast of demand; thus, during the early production opportunity, N is assumed to be a random variable with positive support, distribution function $F(\cdot)$ and density $f(\cdot)$. The late production opportunity is close enough to the start of the selling season that market size is known perfectly.[2] Production during the early opportunity incurs a unit cost c_1, while production during the late opportunity incurs a higher unit cost $c_2 \geq c_1$ due to, e.g., expedited production and shipping costs. Production at either point in time is assumed to be uncapacitated, and production during the late opportunity is assumed to have a short enough leadtime that all units arrive before the start of the selling season.

The firm thus operates in one of two potential regimes: the *single procurement* regime (SP) or the *quick response* regime (QR). In the single procurement regime, all production occurs during

[1] In §7, we relax the assumption of exogenous prices and consider endogenous pricing.

[2] In reality, forecast updating and refinement may be the the result of an endogenous process that may continue even during the selling season, e.g., monitoring early sales and imputing total demand, or performing market research. To avoid issues outside the scope of this analysis–e.g., demand estimation based on stochastic arrivals–we assume that the revelation of N is exogenous and perfect and occurs just prior to the start of the season.

	High Value	Low Value
High Signal	$\theta\alpha$	$(1-\theta)(1-\alpha)$
Low Signal	$\theta(1-\alpha)$	$(1-\theta)\alpha$

Table 1. The four possible combinations of signal and consumer value, and the probability of each for a given signal strength.

the early production opportunity, while in the quick response regime, production may occur at both times. Excess inventory remaining at the end of the selling season has zero value. In both operating regimes, we denote the early production quantity by q (the late production quantity in the QR regime is derived later), and the firm chooses production levels to maximize total expected profit.

While the firm faces market size uncertainty, consumers initially face uncertainty about their own private valuations for the product. Nature moves first (prior to the start of the game) and decides the "type" of each consumer: a fraction θ of the population has positive value $v > p$ for the item, while a fraction $1-\theta$ has zero value. If a consumer possesses value v for the product, we refer to her as a "high type" consumer, whereas if she possesses zero value for the product, we refer to her as a "low type" consumer.

At the start of the selling season, consumers do not know their private valuation for the product (their type). At a random time during the selling season (i.e., uniformly distributed throughout the season), each consumer exogenously learns her value for the product (via, for instance, product reviews from professionals and other consumers, experiences with demonstration units in-store, etc.). While consumers do not know their individual valuations at the start of the selling season, they are not completely ignorant: each consumer receives a noisy private signal that is an indication of her type. We define α to be the quality of the signal, i.e., the probability that the signal is correct. For example, a high type consumer receives a signal of high product value with probability α, and a low type consumer receives a signal of low product value with the same probability. Thus, there are four possible consumer segments (corresponding to pairings of the two possible signals and the two possible values), summarized in Table 1.

We allow consumers to be heterogeneous in the quality of their private signals by letting α be distributed among the population (independently of consumer type) according to the continuous distribution $G(\cdot)$ and density $g(\cdot)$ with support on the interval $(1/2, 1)$. Such heterogeneity in the

Signal	Consumer Type	Fraction of Population
High Value	High Value (Correct Signal)	$\theta \int_{1/2}^{1} \alpha g(\alpha)\, d\alpha$
	Low Value (Incorrect Signal)	$(1-\theta) \int_{1/2}^{1} (1-\alpha)\, g(\alpha)\, d\alpha$
Low Value	High Value (Incorrect Signal)	$\theta \int_{1/2}^{1} (1-\alpha)\, g(\alpha)\, d\alpha$
	Low Value (Correct Signal)	$(1-\theta) \int_{1/2}^{1} \alpha g(\alpha)\, d\alpha$

Table 2. A summary of the distribution of signals and consumer types amongst the population.

quality of the signal may represent, for example, domain expertise of the population in the product category (e.g., some consumers are highly technical and capable of accurately judging the quality of a new, high tech product, while some less sophisticated consumers receive more noisy signals that leave them less sure of product value). Thus, the total number of consumers in each segment depicted in Table 1 is found by integrating the probabilities in that table over the distribution of signal strengths. The resulting distribution of consumer segments arising from this information structure is summarized in Table 2.

After receiving their private signals, consumers arrive at the firm at the start of the selling season. Each consumer updates her beliefs of product value (via Bayes' rule) and calculates the expected utility of purchasing early (before knowing product value) and the expected utility of delaying her purchase until she learns the value of the product, based on her private signal and individual signal strength. In order to evaluate the expected surplus of delaying a purchase, consumers must also form expectations on the probability that she will be able to obtain a unit at some later point in the selling season, which we denote $\widehat{\phi}$.

After consumers learn their value, they purchase if and only if they have positive surplus and the product is in-stock, and any consumer who does not obtain a unit receives zero surplus. Consumers are risk-neutral expected utility maximizers who choose the purchasing strategy (before or after learning product value) that maximizes their total expected surplus (expected product value minus purchase price). We assume that customers who are indifferent between the two strategies purchase before learning product value. To summarize, each consumer knows:

1. Her private signal of product value (high or low) and her individual signal strength α;

2. The common valuation distribution and its parameters (i.e., that a fraction θ of the population has value v);

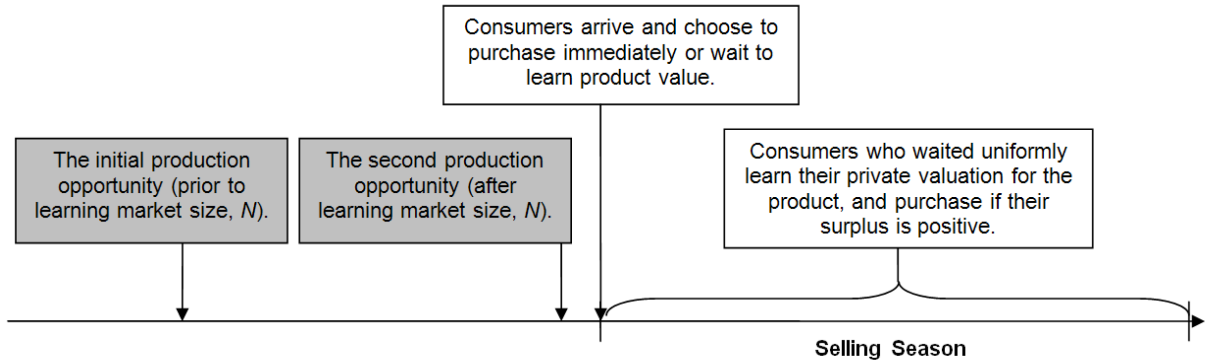

Figure 1. Sequence of events.

3. The purchase price p;
4. An expectation of the future availability of the product, $\widehat{\phi}$.

Note that customers do not observe the firm's inventory level, the realization of demand, or the firm's cost structure, nor are they aware of the firm's demand forecast. To simultaneously model both strategic (forward-looking) and non-strategic (myopic) customers, we introduce a parameter $\delta \in \{0, 1\}$ that is analogous to a discount factor: if $\delta = 0$, customers do not anticipate the opportunity to purchase after learning product value, while if $\delta = 1$, they do. The sequence of events is summarized in Figure 1.

4 The Consumer Decision: Wait or Buy

In this section we analyze the consumer decision: whether to wait or buy. We begin by discussing the nature of consumer expectations of product availability ($\widehat{\phi}$). We assume that consumers form *rational expectations* (see, e.g., Muth 1961, Su and Zhang 2008, and Cachon and Swinney 2009) concerning the availability of the product (i.e., consumers possess beliefs about the chance of obtaining the product that are consistent with the equilibrium availability).[3] Because consumers randomly and uniformly learning their valuations throughout the selling season, the allocation of

[3] While a one-shot model does not provide any details concerning how these expectations may be formed, we might think of rational expectations deriving from repeated interaction with a firm over time; for instance, consumers have come to expect that video game manufacturer Nintendo is incapable of rapid inventory replenishment to meet demand (Richtel 2007) and hence future availability is low. On the other hand, consumers have come to expect that General Motors will satisfy demand on hit products and hence future availability is high, a belief that GM is now actively trying to change (Stoll 2007).

inventory is essentially random (all consumers have an equal chance of procuring a unit). Because expectations are rational and consumers have an equal chance of obtaining a unit, all consumers must have identical expectations of $\widehat{\phi}$. Furthermore, rationality implies that $\widehat{\phi}$ must, in equilibrium, be the probability that a random customer obtains a unit–in other words, the *fill rate*.

In analyzing the consumer decision, the relevant unit of analysis is a consumer who arrives at the start of the selling season, finds a unit in-stock,[4] and considers purchasing the product immediately (which ensures that a unit will be obtained, but not that value will be high) or delaying the purchase decision until she learns her valuation (which ensures that the consumer will only purchase if she has high value for the product, but does not ensure that she will successfully obtain a unit).[5] The expected surplus of an immediate purchase is $\gamma_s(\alpha)v - p$, where $\gamma_s(\alpha)$ is the posterior probability that the consumer has high value for the product, conditional on a signal $s \in \{l, h\}$ (i.e., low or high value) and signal strength α. For a consumer receiving a high value signal, this posterior probability is

$$\gamma_h(\alpha) = \frac{\Pr(\text{High Type and High Signal})}{\Pr(\text{High Signal})} = \frac{\alpha\theta}{\alpha\theta + (1-\alpha)(1-\theta)}. \tag{1}$$

Note that $\gamma_h(\alpha)$ is increasing in α. Similarly, if the consumer receives a signal indicating that the product is low value, the posterior probability is

$$\gamma_l(\alpha) = \frac{(1-\alpha)\theta}{(1-\alpha)\theta + \alpha(1-\theta)}. \tag{2}$$

Note that $\gamma_l(\alpha)$ is decreasing in α. If $\gamma_l(\alpha)v - p > 0$ for some α, consumers receiving a low signal may receive positive surplus from an early purchase, whereas if $\gamma_l(\alpha)v - p < 0$, they always receive negative surplus. In the following analysis, we assume that the latter case holds for all α.[6] Due to this assumption, all consumers receiving a low signal have negative expected surplus

[4] If any consumer finds the firm out-of-stock, the game is essentially over; due to our assumption that the firm's QR order arrives prior to the start of the selling season, if a consumer finds the firm out-of-stock, all subsequent consumers will as well, regardless of the operating regime.

[5] Technically, the consumer chooses between purchases before learning her value and after learning her value, both of which could potentially be at any time during the selling season. However, conditional that a consumer decides to purchase before learning her value, the optimal time to purchase is immediately at the start of the season (as this minimizes the risk of a stock-out). Similarly, conditional on purchasing after learning product value, the optimal purchase time is at the moment she realizes her value for the product, as this too minimizes the risk of a stock-out. Hence, the consumer effectively chooses between an immediate purchase at the start of the season and a purchase at the moment she learns her valuation.

[6] This assumption allows us to ignore customers who a receive a low value signal in all further equilibrium discussion,

from purchasing before learning their valuation. It follows that all such consumers will delay purchasing until after learning their valuations, and only those consumers who receive a high signal will consider a purchase prior to learning their valuations. This fact leads to our first preliminary result, which characterizes consumer actions in any possible equilibrium with rational expectations:

Lemma 1 *Given any rational expectation $\widehat{\phi}$ of the product availability, in any equilibrium there exists a unique critical α^* such that all consumers who receive a high value signal and have $\alpha \geq \alpha^*$ purchase before learning product value while all consumers with $\alpha < \alpha^*$ wait until after learning product value.*

Proof. Consumers who receive a high value signal purchase early if $\gamma_h(\alpha)\, v - p \geq 0$ and if the expected surplus from purchasing early is greater than the expected surplus from delaying, i.e., if

$$\gamma_h(\alpha)\, v - p \geq \delta\widehat{\phi}\gamma_h(\alpha)\,(v-p)\,, \tag{3}$$

where, recall, $\delta \in \{0, 1\}$ determines if consumers are non-strategic or strategic. Because $\delta\widehat{\phi} \leq 1$, it is true that expected surplus from a delay (the right hand side of the above expression) is increasing in α at a slower rate than expected surplus from an immediate purchase (the left hand side). Furthermore, if $\alpha = 1$, then early surplus is $v - p$ while late surplus is $\delta\widehat{\phi}\,(v-p)$, i.e., early surplus is weakly greater than late surplus. If $\alpha = 1/2$, the opposite relationship holds (from our assumption that $\gamma_l(1/2)\, v - p < 0$). Thus, there exists some (unique) critical α^* such that, for all $\alpha > \alpha^*$, the inequality in (3) is strict, while for $\alpha < \alpha^*$, the inequality is violated, which leads to the result. ■

Lemma 1 shows that, in any equilibrium, consumers who receive a signal of high product value and who have high signal quality (accurately judge product value) will purchase before learning their valuations, while consumers who have low signal quality (poorly judge product value) will delay until after learning product value. Thus, we may characterize the equilibrium behavior of the consumer population by a single parameter, the critical signal strength α^*. The firm, which forms beliefs on consumer behavior in order to estimate total demand, thus possesses a belief $\widehat{\alpha}$ on the critical signal strength. We assume that these expectations are consistent with the equilibrium

as their dominant action is to delay purchasing. If we relax this assumption, we must account for low signal customers in each equilibrium, but the qualitative effects of the model remain unchanged.

consumer actions, i.e., the expectations are rational.

5 Equilibrium and the Value of Quick Response

From the analysis in the preceding section, we conclude that we seek an equilibrium to the game that consists of *consumer purchasing behavior* and an *inventory decision* on the part of the firm. Such an equilibrium will be characterized by values of q (the firm's inventory level in the early production opportunity) and α (the critical signal strength of the consumer population, see Lemma 1). Let the superscript $*$ denote a generic equilibrium parameter (replacing $*$ with SP or QR when referring specifically to the single procurement or quick response case). We then formally define the equilibrium as follows:

Definition 1 *An* ***equilibrium*** (q^*, α^*) *with rational expectations to the game between the firm and the consumer population satisfies:*

1. *The firm chooses an initial inventory level* q^* *to maximize total expected profit, conditional on beliefs about consumer behavior,* $\widehat{\alpha}$*;*
2. *The consumer population determines the critical signal strength* α^**, conditional on beliefs about product availability* $\widehat{\phi}$*;*
3. *Beliefs are rational,* $\widehat{\alpha} = \alpha^*$ *and* $\widehat{\phi} = \phi(q^*, \alpha^*)$*, where* $\phi(q, \alpha)$ *is the fill rate given initial inventory* q *and critical signal strength* α*.*

The critical signal strength is determined by calculating the surplus from an immediate purchase by a consumer who arrives at the store and finds a unit in-stock, and equating that surplus with the expected surplus of delaying the purchase until learning product value, yielding

$$\alpha^* = \frac{(1-\theta)\,p}{(1-\theta)\,p + \theta\,(v-p)\left(1-\delta\widehat{\phi}\right)}. \tag{4}$$

Because the actions of all consumers may be summarized by a single variable (α^*, the critical signal strength), there are essentially two actions in the game: the firm chooses an inventory level (which depends upon how many consumers purchase early and how many purchase late), and the consumer population determines the critical signal quality, which depends upon the expected product availability (and hence the inventory level of the firm). The rational expectations hypothesis thus implies the game is one of simultaneous moves with two players.

We must prove that the equilibrium to the game exists (and that such an equilibrium is unique) in order to discuss its properties; the following lemma accomplishes this for the SP regime.

Lemma 2 *When the firm operates in the single procurement regime, an equilibrium (q^{sp}, α^{sp}) exists and is unique. The equilibrium total demand to the firm is*

$$D = N\left(\theta + (1-\theta)\int_{\alpha^{sp}}^{1}(1-x)\,g(x)\,dx\right). \tag{5}$$

Proof. In order to determine the equilibrium to the game, we must first derive the firm's best reply to a belief $\widehat{\alpha}$ concerning customer behavior. All consumers who receive a high value signal with a signal strength greater than $\widehat{\alpha}$ purchase early, thus this portion of demand is composed of two consumer segments: those with high value (probability θ) and correct signals (probability α), and those with low value (probability $1-\theta$) and incorrect signals (probability $1-\alpha$). Let

$$\xi_1(\widehat{\alpha}) = \theta\int_{\widehat{\alpha}}^{1} xg(x)\,dx + (1-\theta)\int_{\widehat{\alpha}}^{1}(1-x)\,g(x)\,dx.$$

The total demand from these consumers is thus $N\xi_1(\widehat{\alpha})$. All consumers with signal strengths less than $\widehat{\alpha}$ delay purchasing, and subsequently only those consumers with high value will purchase the product. Demand from these consumers will thus consist of all consumers who have high value (probability θ) and received a low value signal (probability $1-\alpha$), and consumers who have high value (probability θ), received correct signals (probability α), and chose to delay their purchase (signal strength between 1/2 and $\widehat{\alpha}$). Let

$$\xi_2(\widehat{\alpha}) = \theta\int_{1/2}^{1}(1-x)\,g(x)\,dx + \theta\int_{1/2}^{\widehat{\alpha}} xg(x)\,dx,$$

such that the total demand from these consumer segments is $N\xi_2(\widehat{\alpha})$. The *total* firm demand is thus $D = N\xi(\widehat{\alpha})$, where

$$\xi(\widehat{\alpha}) = \xi_1(\widehat{\alpha}) + \xi_2(\widehat{\alpha}) = \theta + (1-\theta)\int_{\widehat{\alpha}}^{1}(1-x)\,g(x)\,dx.$$

The firm's expected profit is $\pi(q) = \mathbb{E}[p \min(q, D) - c_1 q]$, which is a concave function of q yielding an optimal inventory level satisfying $\Pr(D < q) = (p - c_1)/p$. Substituting for D, we see that the best reply function is

$$q(\widehat{\alpha}) = \left(\theta + (1-\theta) \int_{\widehat{\alpha}}^{1} (1-x)\, g(x)\, dx\right) F^{-1}\left(\frac{p - c_1}{p}\right).$$

We may now derive the equilibrium to the game by imposing the rational expectations hypothesis, which implies $\widehat{\alpha} = \alpha^{sp}$ and $\widehat{\phi} = \phi(q^{sp}, \alpha^{sp})$. Furthermore, this equilibrium is unique (see the technical appendix for a detailed derivation). ■

From (5), the equilibrium demand of the firm is decreasing in α^{sp}. It is apparent, then, that the firm prefers more consumers to purchase early as this increases total demand. This result is sometimes referred to as the advance selling phenomenon–see Xie and Shugan (2001)–in which a firm exploits consumer valuation uncertainty by inducing some consumers to purchase the product before learning their value that will ultimately be dissatisfied (have low valuation).

We next move to the game in which the firm operates in the QR regime. Recall that the firm behaves in a subgame perfect manner; that is, when determining the number of units to produce using quick response, the firm chooses an inventory level that maximizes total profit. As a result, if the firm has quick response capabilities, rational consumers must believe that the fill rate at that firm is equal to 1; after learning the true value of demand, the firm cannot credibly commit to satisfying anything less than the total demand it receives.[7] Consequently, quick response increases the expected surplus of consumers who delay their purchase and strengthens the incentive for consumers to wait. All else being equal, this will shift demand to later times, which will in turn decrease the amount of advance selling that occurs.[8]

[7] It is worthwhile to consider what would happen if the firm did *not* maximize revenue when placing the quick response order–for instance, the firm might plan *ex ante* to fulfill some fraction of the total demand while leaving some residual rationing risk in order to "train" repeat customers to expect limited availability. This case is analyzed in the technical appendix and discussed in §8.

[8] It is important to note that this result is a direct implication of our assumption of rational consumer expectations of product availability, $\widehat{\phi}$. Since, as this discussion demonstrates, the fill rate is generally higher if the firm possesses quick response than if it does not, rational expectations imply that consumers implicitly know whether the firm operates with quick response. (We stress that consumers need not *excplicitly* know details of the firm's operating practices, merely that they are aware of the higher fill rates resulting from these practices.) In §8, we discuss how the results change if consumers are oblivious to the operating regime of the firm, i.e., if consumers cannot deduce (even implicitly) that fill rates are higher at a firm with quick response.

The story does not end with the effect of quick response on consumer behavior, however; QR also offers value by better matching supply and demand under uncertainty. Thus, it remains to be seen how QR affects the profit of the firm in equilibrium. Before we answer this question, we must first demonstrate that an equilibrium exists and is unique when the firm operates in the QR regime. The following lemma does this, in addition to comparing the equilibrium outcomes (critical signal strength and inventory level) to the single procurement regime.

Lemma 3 *When the firm operates in the quick response regime, a subgame perfect equilibrium* (q^{qr}, α^{qr}) *exists and is unique. In equilibrium, more consumers delay their purchases* ($\alpha^{sp} \leq \alpha^{qr}$) *and the firm sets a lower inventory level* ($q^{qr} \leq q^{sp}$) *than in the single procurement regime.*

Proof. Because the firm operates in the QR regime, the only rational belief of the consumer population is that $\widehat{\phi} = 1$; because the quick response procurement is subgame perfect, the firm will satisfy all demand. Hence, the consumer best reply is independent of any firm actions, and is dictated by the solution to (4) with $\widehat{\phi} = 1$, which implies

$$\alpha^{qr} = \frac{(1-\theta)\, p}{(1-\theta)\, p + \theta\,(v-p)\,(1-\delta)}.$$

This is clearly a unique consumer best reply, and it is immediately apparent that $\alpha^{sp} \leq \alpha^{qr}$ for any equilibrium fill rate in the single procurement regime. The firm's profit function is $\pi(q) = \mathbb{E}\left[pD - c_1 q - c_2 (D-q)^+\right]$, where $D = N(\xi_1 + \xi_2)$ and ξ_1 and ξ_2 as are in the proof of Lemma 1. It follows that the firm best reply exists and is unique, given by

$$q(\alpha^{qr}) = \left(\theta + (1-\theta)\int_{\alpha^{qr}}^{1} (1-x)\, g(x)\, dx\right) F^{-1}\left(\frac{c_2 - c_1}{c_2}\right),$$

hence the equilibrium existence and uniqueness results follow. This furthermore implies

$$(1-\theta)\int_{\alpha^{sp}}^{1} (1-x)\, g(x)\, dx \geq (1-\theta)\int_{\alpha^{qr}}^{1} (1-x)\, g(x)\, dx,$$

and hence it follows that total equilibrium demand to the firm is greater in the SP regime than in the QR regime, yielding $q^{qr} \leq q^{sp}$. ■

Having demonstrated that equilibria exist and are unique in both regimes, we may now address the value of quick response: the incremental increase in profit due to the adoption of a quick response system. Our first result demonstrates how the value of quick response is affected by strategic customer behavior:

Theorem 1 *The incremental equilibrium value of quick response ($\pi^{qr} - \pi^{sp}$) is smaller if consumers are strategic ($\delta = 1$) than if they are non-strategic ($\delta = 0$).*

Proof. Let $\Delta = \pi^{qr} - \pi^{sp}$ be the incremental equilibrium value of quick response. Recall that $\xi(\alpha^{qr}) N$ is the equilibrium total demand in the QR regime, while $\xi(\alpha^{sp}) N$ is the demand in the SP regime, where

$$\xi(\alpha^*) = \theta + (1-\theta) \int_{\alpha^*}^{1} (1-x) g(x) \, dx$$

and

$$\alpha^* = \frac{(1-\theta) p}{(1-\theta) p + \theta (v-p)(1-\delta\phi^*)}.$$

From the expression for α^*,

$$\alpha^{sp}_{\delta=0} = \alpha^{qr}_{\delta=0} = \frac{(1-\theta) p}{(1-\theta) p + \theta (v-p)} \leq \alpha^{sp}_{\delta=1} \leq \alpha^{qr}_{\delta=1},$$

where, e.g., $\alpha^{sp}_{\delta=0}$ denotes the equilibrium critical signal strength in the SP regime when $\delta = 0$. Because $\xi'(\alpha^*) < 0$,

$$\xi\left(\alpha^{qr}_{\delta=1}\right) \leq \xi\left(\alpha^{sp}_{\delta=1}\right) \leq \xi\left(\alpha^{qr}_{\delta=0}\right) = \xi\left(\alpha^{sp}_{\delta=0}\right).$$

The equilibrium firm profit is, in the SP regime,

$$\pi^{sp} = \xi(\alpha^{sp}) \left[p \min\left(N, F^{-1}\left(\frac{p-c_1}{p}\right)\right) - c_1 F^{-1}\left(\frac{p-c_1}{p}\right) \right],$$

and in the QR regime,

$$\begin{aligned} \pi^{qr} &= \xi(\alpha^{qr}) \left[p \min\left(N, F^{-1}\left(\frac{c_2-c_1}{c_2}\right)\right) - c_1 F^{-1}\left(\frac{c_2-c_1}{c_2}\right) \right. \\ &\quad \left. + (p-c_2) \mathbb{E}\left(N - F^{-1}\left(\frac{c_2-c_1}{c_2}\right)\right)^+ \right]. \end{aligned}$$

In each expression, the term inside the bracket is the maximum expected profit without strategic customers (i.e., a traditional newsvendor and a newsvendor with quick response, respectively). Denote the bracket term in regime R by B^R, $R = sp, qr$. Note that $B^{qr} \geq B^{sp}$. The incremental value of QR is thus

$$\Delta = \xi\left(\alpha^{qr}\right) B^{qr} - \xi\left(\alpha^{sp}\right) B^{sp}.$$

When $\delta = 0$, this implies

$$\Delta_{\delta=0} = \xi\left(\alpha^{sp}_{\delta=0}\right)\left(B^{qr} - B^{sp}\right),$$

and when $\delta = 1$

$$\Delta_{\delta=1} = \xi\left(\alpha^{qr}_{\delta=1}\right) B^{qr} - \xi\left(\alpha^{sp}_{\delta=1}\right) B^{sp} \leq \xi\left(\alpha^{sp}_{\delta=1}\right)\left(B^{qr} - B^{sp}\right) \leq \Delta_{\delta=0},$$

which proves the result. ■

In other words, Theorem 1 shows that quick response yields *less* value to the firm when consumers are strategic than when they are non-strategic. This is because strategic behavior on the part of consumers reduces the total demand to the firm: when customers are strategic ($\delta = 1$) some individuals intentionally delay their purchase, and inevitably some of these customers will not buy the product once they learn their valuation. As a result, the value of matching supply and demand is lower (there is less potential demand to match).

A natural question to ask is: how *much* is the value of quick response reduced by strategic behavior? Can it ever be negative? Theorem 2 addresses this question.

Theorem 2 *The incremental equilibrium value of quick response ($\pi^{qr} - \pi^{sp}$) is strictly decreasing in the cost of quick response (c_2), and if $c_2 = p$, $\pi^{qr} \leq \pi^{sp}$.*

Proof. Define $\pi^{qr}(q) = \mathbb{E}\left[pD - c_1 q - c_2 (D-q)^+\right]$, where D is the total demand at the firm (a function of α^{qr}). Let π^{qr} be the equilibrium profit of the firm with quick response, and let π^{sp} be the equilibrium profit without QR. Differentiating π^{qr} with respect to c_2, we have, from the Envelope Theorem,

$$\frac{d\pi^{qr}}{dc_2} = \left.\frac{\partial \pi^{qr}(q)}{\partial c_2}\right|_{q=q^{qr}} + \frac{\partial \pi^{qr}(q)}{\partial q}\frac{dq^{qr}}{dc_2} = \left.\frac{\partial \pi^{qr}(q)}{\partial c_2}\right|_{q=q^{qr}}.$$

Note that since α^{qr} contains no dependence on c_2, there is no derivative term with respect to α^{qr}. This implies

$$\frac{d\pi^{qr}}{dc_2} = -\Pr\left(D > q^{qr}\right) = -1 + \frac{c_2 - c_1}{c_2} < 0.$$

Thus, the equilibrium profit of the firm is decreasing in c_2. Note that, in the limit as $c_2 \to p$, the margin on each unit sold that is procured via QR goes to zero. Hence, the firm's profit effectively becomes the same as if it did not have QR capabilities, with one caveat: in equilibrium, more consumers will delay purchasing than if the firm did not have QR. Thus, $\lim_{c_2 \to p} \pi^{qr} = \pi^{sp}|_{\alpha=\alpha^{qr}} \leq \pi^{sp}|_{\alpha=\alpha^{sp}}$, i.e., for large c_2 QR yields lower expected profits than the SP regime. ■

Theorem 2 implies a surprising result: quick response may *reduce* the profit of the firm even if the marginal procurement cost is strictly less than the selling price. This stands in contrast to the existing literature on quick response: with non-strategic consumers (e.g., Fisher and Raman 1996) or with strategic consumers in the absence of learning (Cachon and Swinney 2009), quick response always provides non-negative value if the margin on a unit procured using quick response is weakly positive (i.e., if $c_2 \leq p$). Theorem 2 shows that this need not be the case when consumers learn about their valuations over time: it is possible for quick response to yield a positive margin on each unit sold while simultaneously yielding lower expected profit to the firm than the single procurement regime.

The key to both theorems lies in the dual effects of quick response: *shifting demand* and *matching supply with demand*. These two effects pull the equilibrium profit of the firm in opposite directions. Shifting demand reduces profits by decreasing the amount of advance selling. Matching supply with demand increases profits by eliminating lost sales–all demand is captured, albeit at a higher unit procurement cost–and reducing the chance of overstock. Hence, the firm only values quick response so long as the cost of shifting demand is exceeded by the gain from better matching supply with demand; see Figure 2.[9]

Theorem 1, on the other hand, demonstrates that value of both effects is higher when consumers are non-strategic ($\delta = 0$) than when they are strategic ($\delta = 1$). When consumers are non-strategic, the demand shifting effect is eliminated. Furthermore, total demand to the firm is higher, so the

[9] In Figure 2 and all other graphical examples, $v = 18$, $p = 10$, $c_1 = 5$, $\theta = 0.75$, N is gamma distributed with mean 10 and standard deviation 5, and α follows a beta distribution with both parameters equal to 5 condensed to lie in the interval $(1/2, 1)$.

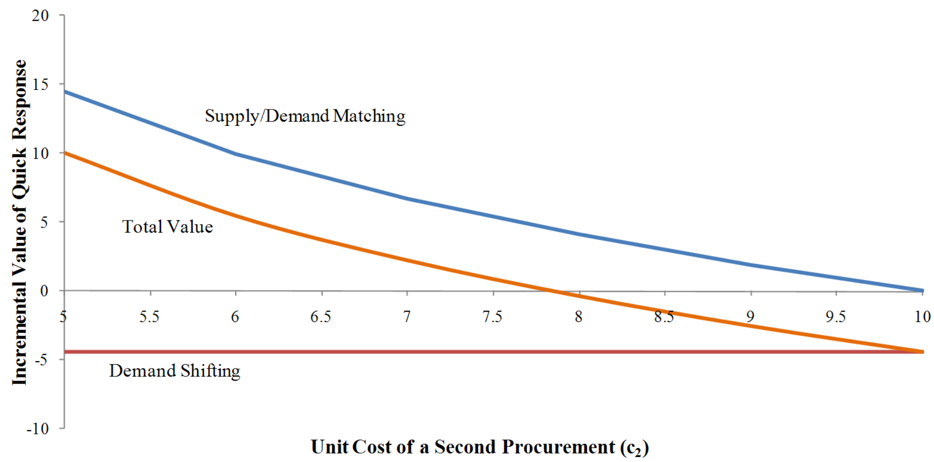

Figure 2. The incremental value of quick response ($\pi^{qr} - \pi^{sp}$) as a function of the cost of an expedited procurement (c_2) when $\delta = 1$, separated into component factors. Matching supply and demand provides positive value while shifting demand provides negative value.

value of matching supply and demand–for any given c_2–is higher than when consumers are strategic. Thus, when $\delta = 0$, all three curves depicted in Figure 2 are higher, as Figure 3 demonstrates.

While we have shown that the value of quick response is lower if consumers are strategic and learn about product value over time, this is not to say that quick response is always harmful to the firm in this setting. As Theorem 2 and Figure 2 demonstrate, quick response can increase the profitability of the firm if c_2 is small enough. Nevertheless, a result of Theorems 1 and 2 is that it may be in the best interests of the firm to forgo quick response tactics and the option to procure additional inventory, and further to ensure that consumers are aware of this operating regime. Particularly in light of additional fixed costs that inevitably accompany the adoption of any quick response system (e.g., shipping and fulfillment infrastructure, IT systems, and production capacity or reservation costs), it is clear that the firm is less likely to benefit from a quick response system when customers are strategic and learn about product value over time.

This relates, in part, to the rationing risk results in the literature on strategic consumer purchasing–see, e.g., Su and Zhang (2008) and Debo and van Ryzin (2007). In contrast to the mere reduction of inventory described in this literature, Theorem 2 implies that the firm may be better off with an entirely different operating policy (Single Procurement vs. Quick Response) when consumers behave strategically–the inability to react to updated demand information in a timely and responsive way can *benefit* the firm by generating a credible mismatch between supply

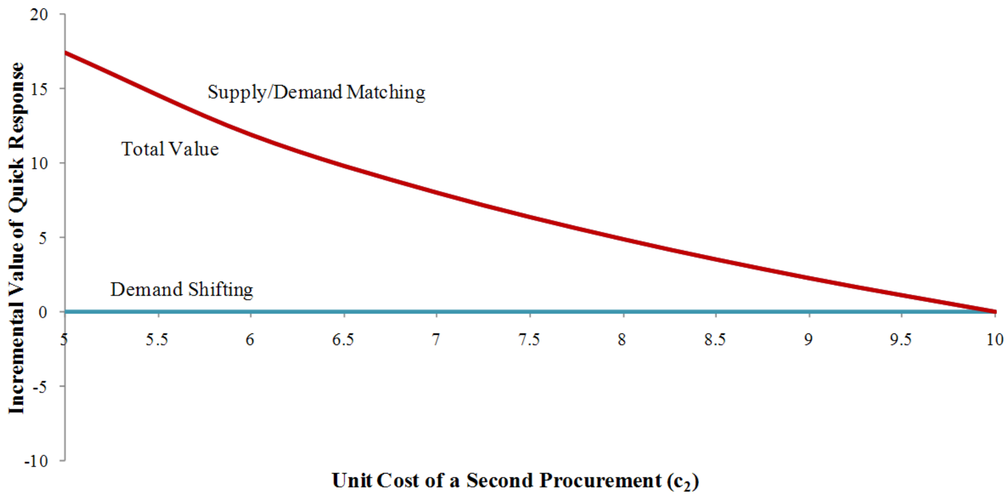

Figure 3. The incremental value of quick response ($\pi^{qr} - \pi^{sp}$) as a function of the cost of an expedited procurement (c_2) when $\delta = 0$. Compared to Figure 2, in which $\delta = 1$, all the curves are shifted upwards.

and demand and inducing more consumers to purchase prior to learning their value.

The fact that it may be optimal for the firm to operate without quick response lends justification to publicized "limited edition" runs of certain products: such tactics induce consumers who are otherwise "on the fence" to purchase the product prior to learning if they truly value it, lest no inventory remain once valuations are revealed. For example, Disney is famous for releasing its classic films on video for very limited periods of time, after which the films are "placed in the Disney vault," not to be released again for a period of several years. However, as we shall see in the following sections, this result is sensitive to at least two key assumptions regarding the nature of the product.

6 Consumer Returns

The preceding analysis assumed that any consumer who purchased an item early had no recourse if their value for that item turned out to be low–that is, the possibility that a consumer could return a product if she is dissatisfied was excluded. In some industries, this assumption is appropriate. For example, with most types of media (e.g., movies, music, video games, or computer software) returns are forbidden once an item has been opened (often due to fears of piracy), and Amazon.com does not allow returns on large televisions due to the logistical challenges of return shipping.

In some cases, however, product returns are a common and important component of firm strategy. Satisfaction guarantees abound in many settings (clothing, electronics, etc.), with firms encouraging customers to try new products "risk free" while promoting generous return policies. At both Amazon.com and the electronics retailer Best Buy, for example, returns are allowed for full refunds on most items within a 30 day period; during the holidays this return window is extended up to a maximum of 90 days. Such policies increase the consumer incentive to purchase early by reducing the consequences of buying a product which is not valued.

In this section, we consider the effect of returns policies on our model of consumer and firm learning. Such policies have been addressed–see the discussion in §2–though unlike some previous papers, we do not address the issue of designing the optimal return policy, but rather we assume that the firm offers a fixed, exogenous return policy to any dissatisfied customer (possibly for marketing or competitive reasons).

Returns are allowed throughout the selling season, and each return is for a full refund minus a consumer restocking fee, $r_c \geq 0$ (i.e., the net refund is $p - r_c$). Returns occur immediately after a consumer who purchased early learns her valuation (e.g., uniformly throughout the selling season). We assume that returned products are resalable–that is, the firm may immediately repackage and resell any returns that it receives. Furthermore, we assume that any consumer who wishes to purchase and finds the firm out-of-stock costlessly waits to see if any returned products become available to purchase during the selling season.

Consumers who make a return incur a hassle cost $h \geq 0$ deriving from, for instance, the travel cost of returning to a store. Returns are also costly to the firm, incurring an internal firm restocking fee of $r_f \geq 0$ on each returned item (due to, for example, repackaging costs or the cost of employee time). We assume that $p - h - r_c \geq 0$, i.e., a dissatisfied consumer benefits from a return. This implies that if $\gamma_h(\alpha) v - p \geq 0$, then

$$\gamma_h(\alpha) v - p + (1 - \gamma_h(\alpha))(p - h - r_c) \geq \gamma_h(\alpha) v - p \geq 0, \tag{6}$$

i.e., with returns, high signal consumers have greater incentive to purchase early than without returns. We assume also that returns are enough of a hassle ($h + r_c$ is large enough) that low signal consumers still do not purchase before learning their valuations.[10]

We are interested in how the addition of the described return policy changes the results of §5, specifically the results provided in Theorems 1 and 2. As we might expect from (6), by increasing expected surplus from an early purchase, returns encourage more consumers to purchase before learning their values. While this would seem to benefit the firm, the increase in advance purchasing comes at a price: consumers who purchase early and are dissatisfied can be costly to the firm, due to the fact that each returned unit costs the firm the price of the refund minus the charged consumer restocking fee, $p - r_c$, and the internal firm restocking fee, r_f. Thus, the value of quick response practices–which as we have already mentioned shift demand by lessening the availability risk associated with delaying a purchase–will depend upon the magnitude of these restocking fees, as the following theorem shows. In what follows, we use the subscript r to denote equilibrium values (profits, quantities, signal strengths) in a model with returns.

Theorem 3 *If consumer returns are allowed:*

1. *If $r_f \geq r_c$, equilibrium firm profit (in either regime) is greater if $\delta = 1$ than if $\delta = 0$. Otherwise, profit is greater if $\delta = 0$.*

2. *If $r_f \geq r_c$, the incremental value of quick response ($\pi_r^{qr} - \pi_r^{sp}$) is always positive. Otherwise, the value may be positive or negative.*

3. *If $r_f \geq r_c$, quick response is more valuable if $\delta = 1$ than if $\delta = 0$. Otherwise, quick response is less valuable if $\delta = 1$.*

Proof. 1. The proofs of equilibrium existence and uniqueness are similar to Lemmas 2 and 3, and are hence omitted. First, we note that with consumer returns, any consumers who purchase early and are dissatisfied with the product will return the item. Because we assume that these products are resalable, the total demand to the firm is simply θN. Thus, the expected profit (without quick response) is

$$\pi_r^{sp}(q) = \mathbb{E}\left[p\theta N - p(\theta N - q)^+ - c_1 q - (r_f - r_c)(1-\theta) N \int_{\alpha_r^{sp}}^{1} (1-x)\, g(x)\, dx\right],$$

[10] Specifically, this implies $\gamma_l(1/2)\, v - p + (1 - \gamma_l(1/2))(p - h - r_c) < 0$.

where α_r^{sp} refers to the equilibrium critical consumer signal strength with returns, determined by equating early purchase and late purchase surplus, yielding

$$\alpha_r^{sp} = \frac{(h + r_c)(1-\theta)}{(h + r_c)(1-\theta) + \theta(v-p)\left(1 - \delta\widehat{\phi}\right)}.$$

Differentiating $\pi_r^{sp}(q)$, we see

$$\frac{d\pi_r^{sp}(q)}{dq} = p(1 - F(q/\theta)) - c_1 \text{ and } \frac{d\pi_r^{sp}(q)}{dq} = -pf(q/\theta).$$

Hence, $\pi_r^{sp}(q)$ is concave in q and yields an optimal inventory level equal to $q_r^{sp} = \theta F^{-1}\left(\frac{p-c_1}{p}\right)$. Note that the optimal inventory level is independent of the critical signal strength, α_r^{sp}, and as a result so is the fill rate, which we denote ϕ_r^{sp}. Thus,

$$\frac{d\pi_r^{sp}}{d\delta} = (r_f - r_c)(1-\theta)\mu(1-\alpha_r^{sp})g(\alpha_r^{sp})\frac{d\alpha_r^{sp}}{d\delta}.$$

Because

$$\frac{d\alpha_r^{sp}}{d\delta} = \alpha_r^{sp}\frac{\theta(v-p)\phi_r^{sp}}{(h + r_c)(1-\theta) + \theta(v-p)(1-\delta\phi_r^{sp})} > 0,$$

it follows that $\frac{d\pi_r^{sp}}{d\delta} \geq 0$ if $r_f \geq r_c$ (and $\frac{d\pi_r^{sp}}{d\delta} \leq 0$ if $r_f \leq r_c$). Similarly, in the quick response regime, as in the case without returns, quick response induces $\widehat{\phi} = 1$, hence

$$\alpha_r^{qr} = \frac{(h + r_c)(1-\theta)}{(h + r_c)(1-\theta) + \theta(v-p)(1-\delta)}$$

and $\alpha_r^{sp} \leq \alpha_r^{qr}$ for any equilibrium belief concerning the fill rate in the SP regime. Thus, the expected profit with quick response is

$$\pi_r^{qr}(q) = \mathbb{E}\left[p\theta N - c_2(\theta N - q)^+ - c_1 q - (r_f - r_c)(1-\theta)N\int_{\alpha_r^{qr}}^1 (1-x)g(x)\,dx\right].$$

Differentiating $\pi_r^{qr}(q)$, we see

$$\frac{d\pi_r^{qr}(q)}{dq} = c_2(1 - F(q/\theta)) - c_1 \text{ and } \frac{d\pi_r^{sp}(q)}{dq} = -c_2 f(q/\theta).$$

$\pi_r^{qr}(q)$ is thus concave in q and yields an optimal inventory level equal to $q_r^{qr} = \theta F^{-1}\left(\frac{c_2-c_1}{c_2}\right)$. Again, the optimal inventory level is independent of the critical signal strength, α_r^{qr}, and as a result so is the fill rate, ϕ_r^{sp}. As before,

$$\frac{d\pi_r^{qr}}{d\delta} = (r_f - r_c)(1-\theta)\mu(1-\alpha_r^{qr})\, g(\alpha_r^{qr})\frac{d\alpha_r^{qr}}{d\delta},$$

where

$$\frac{d\alpha_r^{qr}}{d\delta} = \alpha_r^{qr}\frac{\theta(v-p)}{(h+r_c)(1-\theta)+\theta(v-p)(1-\delta)} > 0,$$

hence $\frac{d\pi_r^{qr}}{d\delta} \geq 0$ if $r_f \geq r_c$ and $\frac{d\pi_r^{qr}}{d\delta} \leq 0$ if $r_f \leq r_c$, proving the result.

2. Let $\Pi_r^{sp} = \mathbb{E}\left[p\theta N - p(\theta N - q_r^{sp})^+ - c_1 q_r^{sp}\right]$ such that

$$\pi_r^{sp} = \Pi_r^{sp} - (r_f - r_c)(1-\theta)\mu\int_{\alpha_r^{sp}}^{1}(1-x)\, g(x)\, dx,$$

and let Π_r^{qr} be defined analogously such that

$$\pi_r^{qr} = \Pi_r^{qr} - (r_f - r_c)(1-\theta)\mu\int_{\alpha_r^{qr}}^{1}(1-x)\, g(x)\, dx.$$

Note that Π_r^{sp} and Π_r^{qr} are the optimal profits (without and with quick response, respectively) of a newsvendor facing demand θN, hence $\Pi_r^{qr} \geq \Pi_r^{sp}$ and both are independent of δ. Thus,

$$\pi_r^{qr} - \pi_r^{sp} = \Pi_r^{qr} - \Pi_r^{sp} + (r_f - r_c)(1-\theta)\mu\int_{\alpha_r^{sp}}^{\alpha_r^{qr}}(1-x)\, g(x)\, dx.$$

If $r_f \geq r_c$, then clearly $\pi_r^{qr} \geq \pi_r^{sp}$.

3. If $\delta = 0$, then $\alpha_r^{sp} = \alpha_r^{qr}$ and thus $\pi_r^{qr} - \pi_r^{sp} = \Pi_r^{qr} - \Pi_r^{sp}$. If $\delta > 0$ and $r_f \geq r_c$, then because $\alpha_r^{sp} \leq \alpha_r^{qr}$, $\pi_r^{qr} - \pi_r^{sp} \geq \Pi_r^{qr} - \Pi_r^{sp}$. If $\delta > 0$ and $r_f \leq r_c$, then because $\alpha_r^{sp} \leq \alpha_r^{qr}$, $\pi_r^{qr} - \pi_r^{sp} \leq \Pi_r^{qr} - \Pi_r^{sp}$. ■

The preceding theorem yields several intriguing results. Part (1) shows that under consumer returns, if $r_f \geq r_c$ firm profit in either regime is *greater* if customers exhibit strategic behavior than if they are non-strategic. The key to this result lies in the fact that, if $r_f \geq r_c$, returns (a) are costly to the firm on a marginal basis and (b) ensure that no consumer who doesn't value the product receives the product, thereby eliminating the advance selling effect and guaranteeing that firm demand (in the sense of non-returned purchases) is always θN regardless of the value of δ. Thus, there is no benefit to selling a unit to a consumer who ultimately possesses low value for the product; on the contrary, this is costly to the firm because of the restocking fees and refund amount. The firm seeks to minimize the number returns, and the number of returns is lower when consumers are strategic (and hence wait to learn about product value before purchasing) than when they are non-strategic (blindly purchasing before knowing their real valuation, only to return the item later). If, on the other hand, $r_f < r_c$, then the firm charges customers more for a return than its own internal costs associated with a return; in this case, the firm profits from each individual return and so, just as in the model without consumer returns, prefers if customers purchase before learning their valuations. Consequently, the firm prefers a non-strategic customer population.

Part (2) of Theorem 3 shows that if $r_f \geq r_c$, quick response always increases firm profit ($\pi_r^{qr} - \pi_r^{sp} \geq 0$). Just as we saw in part (1) of the theorem, the firm benefits from minimizing the number of costly returns–hence, the tendency of quick response to shift demand also *increases* firm profit.[11] When $r_f < r_c$, however, this may or may not be the case; just as in the model without returns, the firm is hurt by demand shifting as it reduces advance selling and profitable returns. Finally, part (3) of Theorem 3 shows that if $r_f \geq r_c$ the result of Theorem 1 is reversed: the value of quick response is *greater* if customers exhibit strategic behavior than if they are non-strategic. Intuitively, the ability of a quick response system to induce demand shifting (which is profitable if $r_f \geq r_c$) is most effective when consumers are strategic (indeed, when consumers are completely non-strategic, quick response induces no demand shifting at all). Hence, the value of a quick response is greatest under forward-looking customer behavior. Alternatively, when $r_f < r_c$, we again have a result similar to Theorem 1: quick response is less valuable when customers are strategic because it generates demand shifting and causes the firm to lose profitable returns.

The results of Theorem 3 are due to the inclination of consumers to hoard inventory: given that returns are possible, a consumer would rather purchase an item early and run the risk of having to return the product, as opposed to delaying the purchase and risking a stock-out. Two ways to reduce hoarding are to increase availability (e.g., adopt quick response) and make consumers

[11] Indeed, this result extends to the case when some or all of the returned goods are not resalable–in that case, returns are even more costly to the firm due to the lost opportunity of reselling a returned product, further strengthening the firm's desire to minimize returns.

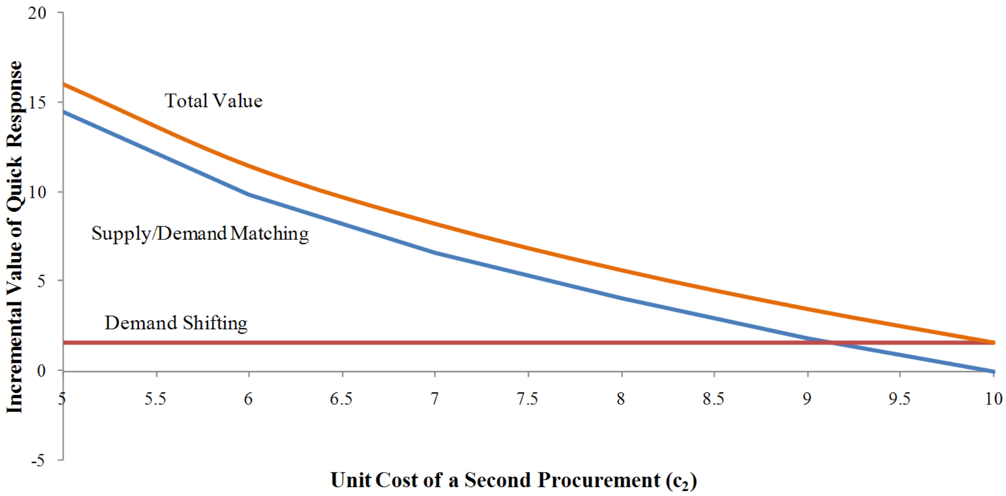

Figure 4. The incremental value of quick response ($\pi_r^{qr} - \pi_r^{sp}$) as a function of the cost of an expedited procurement (c_2) when $\delta = 1$, $h = 2$, $r_f = 1$ and $r_c = 0$, separated into component factors. With costly returns, both matching supply and demand and shifting demand provide positive value.

strategic (increase δ from 0 to 1). If $r_f \geq r_c$, then hoarding is costly to the firm and so both strategic behavior and quick response help to minimize this negative behavior. See Figure 4 for a graphical depiction of this effect on the value of quick response under costly returns ($r_f \geq r_c$).

The fact that in some cases strategic customer behavior can be *good* for the firm (and for the value of quick response) runs contrary to the vast majority of the strategic consumer literature. This is because, in our model, forward-looking behavior results in actions that benefit customers (due to the avoidance of hassle costs and consumer return fees) and the firm (due to the avoidance of internal firm restocking costs). Of course, this result depends critically on the whether individual returns are costly or profitable to the firm; however, in many industries (e.g., retailing) the vast majority of returns are for full (or nearly full) refunds due to competitive pressure, and are subsequently costly to firms–see Stock et al. (2006) for a discussion of how firms actively attempt to minimize returns, and Moorthy and Srinivasan (1995) for a discussion of costly returns. Thus, our model demonstrates how in many cases of practical relevance, the interaction of two effects–consumer learning and costly product returns–implies that the firm benefits from both quick response practices and a very strategic customer population.

7 Pricing

In this section, we endogenize pricing in our original model and address how the value of quick response is affected. We consider two types of pricing: fixed pricing (in which the retailer sets a single price for the entire selling season) and introductory pricing (in which the retailer may set a different price during the initial–or introductory–release of the product, e.g., when consumer valuations are still unknown).

7.1 Fixed Pricing

Unlike the inventory level, price is directly observed by consumers, and hence the firm acts as a Stackelberg leader in the price game. Thus, the model with fixed pricing entails a first stage in which the firm sets the (constant) selling price, and a second stage which behaves identically to the games analyzed in §§3–5. As a result, given a particular price, the previous results continue to hold (notably the equilibrium existence results) in the second stage of the game, and we need only analyze the firm's choice of the selling price by comparing expected profits in the inventory/purchasing subgames using various price levels. The following theorem confirms that the result of Theorem 2–quick response may decrease firm profit–continues to hold even when the firm may set a (constant) price level. In what follows, we use the subscript fp to denote equilibrium values (profits, quantities, signal strengths) in a model with fixed endogenous pricing.

Theorem 4 *The incremental equilibrium value of quick response with fixed pricing ($\pi_{fp}^{qr} - \pi_{fp}^{sp}$) is strictly decreasing in the cost of quick response (c_2), and if $c_2 = v$, $\pi_{fp}^{qr} \leq \pi_{fp}^{sp}$.*

Proof. Note that the existence of an equilibrium is immediate, due to the fact that we have already shown an equilibrium exists to the inventory/purchasing subgames and the firm's expected payoffs are bounded (by 0 and $\mathbb{E}N(v - c_1)$) and its strategy space is a compact interval $[c_1, v]$ in the pricing supergame ($[c_2, v]$ when using quick response–if price is less than c_2 but greater than c_1, the firm will never use QR and reverts to the SP regime). Let π_{fp}^{qr}, p_{fp}^{qr}, and q_{fp}^{qr} be the equilibrium profit, price, and inventory of the firm with quick response and fixed pricing, and let π_{fp}^{sp} be the equilibrium profit without QR. Differentiating π_{fp}^{qr} with respect to c_2, we have, from the Envelope

Theorem,

$$\frac{d\pi_{fp}^{qr}}{dc_2} = \frac{\partial\pi_{fp}^{qr}}{\partial c_2} + \frac{\partial\pi_{fp}^{qr}}{\partial p}\frac{dp_{fp}^{qr}}{dc_2} + \frac{\partial\pi_{fp}^{qr}}{\partial\alpha}\frac{d\alpha_{fp}^{qr}}{dc_2} \leq \frac{\partial\pi_{fp}^{qr}}{\partial c_2}.$$

Observe that either $\frac{\partial\pi_{fp}^{qr}}{\partial p} = 0$ (the firm prices at an interior optimum) or $\frac{dp_{fp}^{qr}}{dc_2} = 0$ (the firm prices on the boundary, i.e., c_2 or v). Unlike the case without pricing, $\frac{d\alpha_{fp}^{qr}}{dc_2}$ in general does not equal zero. This is due to the fact that $\frac{dp_{fp}^{qr}}{dc_2} \geq 0$ and $\frac{d\alpha_{fp}^{qr}}{dp} \geq 0$–in other words, higher costs of quick response lead to higher prices (a natural result) and higher prices lead to more consumers waiting, see equation (4). Because $\frac{\partial\pi_{fp}^{qr}}{\partial\alpha} \leq 0$ (the more consumers that wait, the lower the firm's profits), it follows that the $\frac{\partial\pi_{fp}^{qr}}{\partial\alpha}\frac{d\alpha_{fp}^{qr}}{dc_2} \leq 0$. Finally, since

$$\frac{d\pi_{fp}^{qr}}{dc_2} \leq \frac{\partial\pi_{fp}^{qr}}{\partial c_2} = -\Pr\left(D > q_{fp}^{qr}\right) = -1 + \frac{c_2 - c_1}{c_2} < 0,$$

we find that profit is decreasing in c_2, precisely as in the case without pricing, and $\pi_{fp}^{qr} - \pi_{fp}^{sp}$ is similarly decreasing in c_2. In the limit as $c_2 \to v$, the firm's optimal price with QR goes to v, and margin on each unit sold that is procured via QR goes to zero. Hence, the firm's profit effectively becomes the same as if it did not have QR capabilities, with two caveats: it is constrained to price at v (in the SP regime, the firm can price anywhere in the interval $[c_1, v]$), and in equilibrium, more consumers will wait than if the firm did not have QR due to the fact that QR naturally shifts demand. In other words, if $c_2 = v$,

$$\pi_{fp}^{qr} - \pi_{fp}^{sp} = \pi^{qr}|_{p=v} - \max_{p\in[c_1,v]} \pi^{sp} \leq \pi^{qr}|_{p=v} - \pi^{sp}|_{p=v} \leq 0$$

where the last inequality follows from Theorem 2. ■

The key to this result is the following: when prices are fixed across time, regardless of the optimal price level, adopting quick response increases the consumer incentive to wait and hence decreases advance selling and firm profit. The freedom to set the price is of little value in the quick response regime when c_2 is large, as the firm's optimal price lies in the interval $[c_2, v]$–if the the price is lower than c_2, then quick response is never used, hence the firm essentially moves to the single procurement regime. In the single procurement regime, the firm remains free to price anywhere in the interval $[c_1, v]$. When the cost of quick response is large, the quick response regime has two detrimental effects to the firm: pricing is constrained and more consumers delay purchasing due to higher availability. As a result, the single procurement regime becomes even more attractive than in the exogenous price case. Thus, Theorem 4 mirrors the result of Theorem 2: it is possible for quick response to decrease profit even when the margin is positive ($c_2 \leq p \leq v$).

7.2 Introductory Pricing

In the introductory pricing case, we assume that the firm charges two different prices: an introductory price and a regular price. The introductory price is valid only at the start of the selling season (e.g., the first week) when consumers make their initial decision on when to purchase, while the regular price is valid thereafter.[12]

Consumers develop rational expectations of future prices–that is, they correctly anticipate the regular price (or, equivalently, the firm credibly announces the regular price along with the introductory price). We first note that if the firm is free to set different prices but is constrained only to mark prices *down* over time, Theorem 4 continues to hold. The reason is that it is never optimal in the current model to set an introductory price that is higher than the regular price–the lower regular price would only encourage more consumers to delay purchasing and hence decrease the amount of advance selling. Thus, a firm constrained to mark down over time chooses to set a constant price, and the model reduces to the fixed pricing case analyzed above.

If the firm can *raise* prices over time, however, a different picture emerges. Let p_1 and p_2 be the introductory price and the regular price, respectively. Note that the optimal regular price is $p_2 = v$; all consumers know their values when purchasing at the regular price, and possess values equal to v or 0 for the product. Hence, the firm extracts all surplus from consumers purchasing after learning the product's value by charging the valuation of the high type consumers. Consequently, all consumers have zero surplus from delaying a purchase (both high and low types, regardless of whether they successfully procure a unit), and all consumers with positive expected surplus from an early purchase will choose to buy before learning their valuations. In general, the optimal introductory price satisfies $p_1 \leq v$, i.e., the firm charges a lower introductory price to induce some advance selling among consumers.

Because all consumers have identically zero surplus from a delayed purchase, if the firm adopts

[12] For concreteness, we may thinking of dividing the selling season into two periods, a short introductory period during which no consumer learning occurs and a longer regular period during which consumers learn about product value.

quick response and raises the consumer expectation of product availability $(\widehat{\phi})$, the firm does *not* raise the expected surplus to any consumers from a delayed purchase. Thus, quick response no longer shifts demand–the only effect remaining is matching supply and demand, hence quick response always has positive value. The following theorem summarizes this result. In the following theorem, we use the subscript *ip* to denote equilibrium values (profits, quantities, signal strengths) in a model with endogenous introductory pricing.

Theorem 5 *The incremental equilibrium value of quick response with introductory pricing* $(\pi_{ip}^{qr} - \pi_{ip}^{sp})$ *is always positive if* $c_2 \leq v$.

Proof. From the preceding discussion, $p_2 = v$, and the existence of an equilibrium to the introductory pricing supergame follows from a similar argument to Theorem 4. Combined with the rational expectations assumption of consumer beliefs concerning future pricing, this implies the critical α is determined by the solution to $\frac{\alpha\theta}{\alpha\theta+(1-\alpha)(1-\theta)}v - p_1 = 0$, yielding

$$\alpha^* = \frac{p_1 (1-\theta)}{\theta (v - p_1) + p_1 (1-\theta)}$$

regardless of the firm's operating regime. Note that because the regular price is equal to v the firm always makes a profit on a unit procured using quick response, and furthermore introductory period price may lie anywhere in the interval $[c_1, v]$. It is straightforward to see that $\pi_{ip}^{qr} - \pi_{ip}^{sp}$ is decreasing in c_2. Next observe that, as a function of q and p_1, profit in the SP regime is

$$\pi_{ip}^{sp} (q, p_1) = \mathbb{E} \left[p_1 \min (\xi_1 N, q) + v \min \left(\xi_2 N, (q - \xi_1 N)^+ \right) - c_1 q \right],$$

where ξ_1 and ξ_2 are functions of p_1 (implicitly via α^*), defined similarly to the expressions in Lemma 2. Similarly, profit in the QR regime is

$$\pi_{ip}^{qr} (q, p_1) = \mathbb{E} \left[p_1 \min (\xi_1 N, q) + v \min \left(\xi_2 N, (q - \xi_1 N)^+ \right) + (v - c_2) \left((\xi_1 + \xi_2) N - q \right)^+ - c_1 q \right],$$

where ξ_1 and ξ_2 are the precise same functions as in the expression for $\pi_{ip}^{sp} (q, p_1)$. Thus, for any given combination of q and p_1,

$$\pi_{ip}^{qr}(q,p_1) - \pi_{ip}^{sp}(q,p_1) = \mathbb{E}\left[(v-c_2)\left((\xi_1+\xi_2)N-q\right)^+\right] \geq 0,$$

hence it must be true that the optimal prices and quantities, $\pi_{ip}^{qr} - \pi_{ip}^{sp} \geq 0$ as well. ■

The key to Theorem 5 is that increasing prices over time provides consumers with greater incentive to purchase early, shifting demand from *later* purchases to the *earlier* purchases. This effect counteracts the tendency of quick response to shift demand in the opposite direction. Thus, introductory pricing and quick response are complimentary in the sense that they enhance one another's value: increasing prices reduces costly demand shifting due to quick response, and quick response eliminates costly supply/demand mismatches (mismatches which are particularly costly under introductory pricing due to the higher regular price).

We note that due to the assumption that consumer values follow a two point distribution, introductory pricing in the present model completely eliminates strategic waiting in the sense that all consumers receive zero surplus from a delayed purchase and hence consumers purchase early if and only if they have positive expected surplus (e.g., as if they were non-strategic). Should consumers have more than one positive valuation, in general introductory pricing will not eliminate all strategic waiting, i.e., it will provide positive surplus to some consumers who delay purchasing. In that case, the adoption of quick response once again shifts demand to later times and decreases advance selling; nevertheless, increasing prices over time continues to reduce the amount of strategic waiting that occurs and hence minimizes the negative aspects of demand shifting due to quick response. Thus, while strategic waiting will not in general be eliminated by adopting introductory pricing if consumers have a more complicated valuation distribution, it will be *reduced* by introductory pricing, a fact which increases the value of quick response relative to the fixed pricing case.

8 Discussion

Quick response systems–or, more generally, leadtime reduction and rapid inventory replenishment–are often suggested as potential panaceas to the ill effects of supply and demand mismatches. In this paper, we show that such strategies are less valuable to the firm when consumers are forward-looking and have uncertain value for a product about which they learn over time. Furthermore, even if the fixed cost of implementing a quick response system is zero, it is possible that the option to receive additional inventory after a forecast update *decreases* the firm's profit once the consumer response to increased availability is taken into account. Though it is a commonly held belief that a faster, more responsive supply chain is a more profitable supply chain, we show that such responsiveness is not necessarily beneficial to a firm: when returns are forbidden or when prices are constant, the firm can exploit valuation uncertainty by advance selling, and quick response decreases the extent to which the firm can advance sell. By operating with rapid fulfillment capabilities, the firm loses its ability to credibly restrict inventory to create a stock-out risk, and thus may reduce its overall profitability in certain situations.

Given this result, a natural question to ask is: what happens to the value of quick response if the firm *can* credibly commit to future availability risk (as might be the case if, for example, the firm and consumers interact repeatedly over many periods)? It is possible to show that if the firm can commit to any arbitrary fill rate (a model analyzed in the technical appendix), it is optimal for the firm to commit to *not* raise fill rates by adopting quick response–in other words, the firm commits to the same fill rate in both operating regimes. This eliminates the demand shifting effect, and a result, the value of quick response can never be negative with fill rate commitment. However, even in this setting, it remains true that strategic customer behavior reduces the value of quick response relative to the non-strategic customer case. Thus, while the value of quick response cannot be negative with fill rate commitment, it is still lower with forward-looking customers than with myopic customers; Theorem 1 continues to hold in this case, though Theorem 2 does not.[13]

In that regard, the first model that we analyzed represents a worst case scenario for the firm. In this scenario, the firm cannot credibly commit to future availability risk and implicitly cannot hide its quick response capabilities from consumers (a consequence of the rational expectations hypothesis), and as a result, the value of quick response can be negative. In the absence of costly consumer returns, even in the best case scenario–in which the firm can credibly commit to any arbitrary fill rate and eliminate demand shifting–the value of quick response is reduced by the mere

[13] The fact that the fill rate is the same in both operating regimes implies customers need not be aware–even implicitly–of the operating regime of the firm, i.e., customers may be oblivious to the change in fill rates resulting from quick response. Thus, the results described here also apply to a model in which customers are completely ignorant of the firm's operating regime and its consequences with regard to the fill rate.

presence of forward-looking customers, a fact which may lead the adoption of such a strategy to be unprofitable once fixed costs are taken into account.

This result persists even though our model of a quick response system overly optimistic as to its value: in particular, we have assumed that the firm perfectly and exogenously learns demand prior to the start of the selling season and can procure any amount of inventory in the second production run. Nevertheless, even in this idealized setting, the value of quick response can become negative. In a more realistic quick response system, one in which forecast updates are imperfect or derive from endogenous updating based on early sales, the value of the second procurement opportunity will be even lower. Indeed, if forecast updates are based on monitoring early sales data and updating demand estimates, it is likely that demand shifting–which causes a greater fraction of demand to occur later in the selling season–is even more harmful to the firm, as it reduces the early sales data and hence the accuracy of the forecast update.

It is worth noting that quick response is an operational proxy for (more generically) information. Taken in that context, our results on the value of *quick response* are essentially results on the value of *accurate information* concerning product demand. The fact that quick response sometimes yields negative value supports the maxim that ignorance can be bliss; the lack of accurate information on demand can serve as a commitment mechanism to keep inventory scarce and increase advance selling. Taken together, our results provide insight into when a firm should adopt a fast supply chain that allows action on improved demand information. Whether selling to consumers or retailers, the value of matching supply and demand depends not only on the reduction of lost sales and excess inventory, but also on the strategic response of the firm's customers to increased product availability. This response can be harmful (if advance selling decreases as a result), beneficial (if costly returns are allowed and hoarding is an issue), and even diminished or eliminated by the appropriate pricing strategy (increasing prices over time in the optimal manner). Care must thus be taken when assessing the value of supply chain responsiveness in order to assess all consequences of such a strategy–both operational and behavioral.

Acknowledgements. Many thanks to the associate editor and two anonymous referees, as well as Gérard Cachon, Marshall Fisher, Serguei Netessine, Senthil Veeraraghavan, Arvind Tripathi, and seminar participants at the University of Pennsylvania, the University of Texas at Dallas, the University of Rochester, the University of Washington, New York University, London Business

School, Washington University in St. Louis, Northwestern University, the University of Chicago, Columbia University, Duke University, Stanford University, and the INFORMS Annual Meeting in Seattle for numerous comments and suggestions.

References

Akan, M., B. Ata, J. Dana. 2007. Revenue management by sequential screening. Working paper, Northwestern University.

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.

Bikhchandani, S., D. Hirshleifer, I. Welch. 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *The Journal of the Political Economy* **100**(5) 992–1026.

Cachon, G. P., R. Swinney. 2009. Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Sci.* Forthcoming.

Coughlan, A., C. Savaskan, J. Schulman. 2007. Effect of return policies on retail competition. Working paper, Northwestern University.

Dana, J. 1998. Advance-purchase discounts and price discrimination in competitive markets. *Journal of the Political Economy* **106**(2) 395–422.

Davis, S., E. Gerstner, M. Hagerty. 1995. Money back guarantees in retailing: Matching products to consumer tastes. *Journal of Retailing* **71**(1) 7–22.

Debo, L. 2007. In the shadow of the past: Optimal pricing strategies for products with unknown quality and observable price and sales history. Working paper, Carnegie Mellon University.

Debo, L., G. van Ryzin. 2007. Stock-outs and customer purchasing behavior when product quality is uncertain. Working paper, University of Chicago.

DeGraba, P. 1995. Buying frenzies and seller-induced excess demand. *The RAND Journal of Economics* **26**(2) 331–342.

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.

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

Jerath, K., S. Netessine, S. K. Veeraraghavan. 2007. Revenue management with strategic customers: Last minute selling and opaque selling. Working paper, University of Pennsylvania.

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

Liu, Q., W. Xiao. 2008. Selling to heterogeneous customers with uncertain valuations under returns policies. Working paper, Hong Kong University of Science and Technology.

Moorthy, S., K. Srinivasan. 1995. Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Sci.* **14**(4) 442–466.

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

Richtel, M. 2007. A year later, the same scene: Long lines for the elusive Wii. The New York Times.

Slatalla, M. 2002. Hot holiday toys: A parent's strategy. The New York Times.

Stock, J., T. Speh, H. Shear. 2006. Managing product returns for competitive advantage. *Sloan Management Review* **48**(1) 57–62.

Stoll, J. 2007. How GM handles a hit: Build fewer. The Wall Street Journal.

Su, X. 2009. Consumer returns policies and supply chain performance. *Manufacturing Service Oper. Management* Forthcoming.

Su, X., F. Zhang. 2008. Strategic consumer behavior, commitment, and supply chain performance. *Management Sci.* **54**(10) 1759–1773.

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.

Technical Appendix to "Selling to Strategic Consumers When Product Value is Uncertain: The Value of Matching Supply and Demand"

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

May, 2008. Revised January, 2009.

1 Lemma 2

The following derivations show that the equilibrium in the SP regime is unique, a fact which is used in the proof of Lemma 2. First, recall that the firm's optimal inventory level (as a function of the consumer critical signal strength, α) is

$$q(\alpha) = (\xi(\alpha)) F^{-1}\left(\frac{p-c_1}{p}\right),$$

With a random allocation rule, the actual second fill rate for any (q, α) is given by

$$\phi(q,\alpha) = \mathbb{E}\left[\min\left((q-\xi(\alpha)N)^+, \xi(\alpha)N\right)/\xi(\alpha)N\right],$$

or equivalently,

$$\phi(q,\alpha) = \int_0^{q/\xi(\alpha)} f(x)\,dx + \int_{q/\xi(\alpha)}^{\infty} \left(\frac{q-\xi(\alpha)x}{\xi(\alpha)x}\right) f(x)\,dx.$$

Substituting the firm's optimal inventory level, we see that in any equilibrium,

$$\phi(q,\alpha) = \int_0^{F^{-1}\left(\frac{p-c_1}{p}\right)} f(x)\,dx + \int_{F^{-1}\left(\frac{p-c_1}{p}\right)}^{\infty} \left(\frac{F^{-1}\left(\frac{p-c_1}{p}\right)-x}{x}\right) f(x)\,dx. \tag{1}$$

Note that this expression is independent of α. This is because, given some α, it is optimal for the firm to adjust the inventory level such that the fill rate is constant–hence, in an equilibrium, the fill rate is equal to (1). Next, we need to show that a unique solution in α exists to the expression

$$\alpha = \frac{(1-\theta)p}{(1-\theta)p + \theta(v-p)(1-\delta\phi(q,\alpha))}, \tag{2}$$

which will imply that a unique consumer action occurs in any equilibrium. Note that the left hand side of (2) is increasing in α with a slope of 1, and the right hand side is increasing independent of α in any equilibrium because $\phi(q,\alpha)$ is independent of α in equilibrium. Hence, there clearly exists a unique α which satisfies (2), and thus the equilibrium to the game is unique.

2 Fill Rate Commitment

In this extension, we consider the effect of allowing the firm to credibly commit to a maximum fill rate. Specifically, we allow the firm to announce, at the start of the game, that at most a fraction $\phi \leq 1$ of demand will be fulfilled. In other words, if the total demand is D, the firm only fulfills ϕD customers, and should the firm have inventory in excess of this demand, it costlessly disposes of the excess units. Note that this is inherently not a subgame perfect strategy on the part of the firm; however, it may be a desirable (profitable) strategy if the firm has some mechanism to enforce its actions, such as reputational effects or repeated interaction with consumers.

After announcing ϕ to consumers, the firm chooses an inventory level which maximizes expected profit, given the choice of ϕ. When the firm announces ϕ, consumers believe this announcement and interpret it to mean that every consumers has a probability ϕ of obtaining a unit (i.e., no consumer believes he or she is "special" or has a higher chance of getting a unit). The firm is thus a leader in a sequential game, choosing ϕ to induce a particular critical $\alpha(\phi)$:

$$\alpha(\phi) = \frac{(1-\theta)p}{(1-\theta)p + \theta(v-p)(1-\delta\phi)}.$$

Note the absence of $\hat{\cdot}$ symbols; consumers need no longer form *beliefs* about fill rate since the firm credibly conveys this information at the start of the game. Let

$$\xi(\phi) = \theta + (1-\theta)\int_{\alpha(\phi)}^{1}(1-x)\,g(x)\,dx$$

be the total equilibrium fraction of demand that attempts to purchase the product (analogous to the result derived in Lemma 1 of the main text). Note that $\alpha(\phi)$ is increasing in δ, and hence $\xi(\phi)$ is *decreasing* in δ. Thus, total firm demand is $\phi\xi(\phi)N$–the total number of customers willing to buy from the firm, $\xi(\phi)N$, times the maximum fill rate ϕ. Expected profit as a function of ϕ is

$$\pi^{sp}(q,\phi) = p\min(\phi\xi(\phi)N, q) - c_1 q.$$

Then, given a committed ϕ, the firm essentially solves a newsvendor problem with demand $\phi\xi(\phi)N$. The optimal inventory level given ϕ is thus

$$q^{sp}(\phi) = \phi\xi(\phi)F^{-1}\left(\frac{p-c_1}{p}\right).$$

Similarly, in the quick response regime, firm profit is

$$\pi^{qr}(q,\phi) = p\min(\phi\xi(\phi)N, q) - c_1 q + (p-c_2)\mathbb{E}(\phi\xi(\phi)N - q)^{+}.$$

The optimal q given ϕ is the QR regime is

$$q^{qr}(\phi) = \phi\xi(\phi)F^{-1}\left(\frac{c_2-c_1}{c_2}\right).$$

There are two items to note concerning the expressions for expected profit:

1. For any q and ϕ, $\pi^{qr}(q,\phi) - \pi^{sp}(q,\phi) \geq 0$. Hence, the QR regime always yields greater profit than the SP regime.

2. For any given ϕ, $q^{qr}(\phi) \leq q^{sp}(\phi)$. Thus, it takes a lower inventory level to satisfy a given

second period fill rate in the QR regime.

Consequently, the optimal profits as a function of the second period fill rate are, in the SP regime,

$$\pi^{sp}(\phi) = \phi\xi(\phi)\left[p\min\left(N, F^{-1}\left(\frac{p-c_1}{p}\right)\right) - c_1F^{-1}\left(\frac{p-c_1}{p}\right)\right],$$

and in the QR regime,

$$\begin{aligned}\pi^{qr}(\phi) &= \phi\xi(\phi)\left[p\min\left(N, F^{-1}\left(\frac{c_2-c_1}{c_2}\right)\right) - c_1F^{-1}\left(\frac{c_2-c_1}{c_2}\right)\right.\\ &\quad \left.+(p-c_2)\,\mathbb{E}\left(N - F^{-1}\left(\frac{c_2-c_1}{c_2}\right)\right)^+\right].\end{aligned}$$

In each expression, the term inside the bracket is the maximum expected profit without strategic customers (i.e., a traditional newsvendor and a newsvendor with quick response, respectively). Denote the bracket term in regime R by B^R, $R = sp, qr$. Note that $B^{qr} \geq B^{sp}$.

Here we note a crucial result. Because both profit expressions are of the form $\pi^R(\phi) = \phi\xi(\phi)B^R$, it follows that maximizing $\pi^R(\phi)$–*in either operating regime*–is equivalent to maximizing $\phi\xi(\phi)$. Thus, both operating regimes possess the same optimal second period fill rate, which we denote ϕ^*. Immediately, this implies that in the presence of perfectly credible fill rate commitment, demand shifting as a result of QR is eliminated.

Furthermore, the incremental value of quick response is given by

$$\pi^{qr} - \pi^{sp} = \phi^*\xi(\phi^*)\left(B^{qr} - B^{sp}\right).$$

As noted above, $\partial\xi(\phi)/\partial\delta < 0$, hence from the Envelope Theorem

$$\frac{d(\phi^*\xi(\phi^*))}{d\delta} = \left.\frac{\partial(\phi\xi(\phi))}{\partial\delta}\right|_{\phi=\phi^*} + \left.\frac{\partial(\phi\xi(\phi))}{\partial\phi}\right|_{\phi=\phi^*}\frac{d\phi^*}{d\delta}.$$

Because $\left.\frac{\partial(\phi\xi(\phi))}{\partial\phi}\right|_{\phi=\phi^*} = 0$ by virtue of the firm optimizing the fill rate, it follows that

$$\frac{d(\phi^*\xi(\phi^*))}{d\delta} = \phi\left.\frac{\partial\xi(\phi)}{\partial\delta}\right|_{\phi=\phi^*} < 0.$$

In other words, the more "strategic" consumers are (the higher δ), the lower the total demand. Thus, because $B^{qr} \geq B^{sp}$,

$$\frac{d}{d\delta}\left(\pi^{qr} - \pi^{sp}\right) = \phi^*\frac{d\xi(\phi^*)}{d\delta}\left(B^{qr} - B^{sp}\right) < 0.$$

In other words, the incremental value of quick response is decreasing in the δ. In particular, this implies that quick response possesses *lower* value if customers are strategic ($\delta = 1$) than if they are non-strategic ($\delta = 0$).