New Product Launch Date Decisions: Promotion and Production

November 2008

Laura J. Kornish*
Leeds School of Business
University of Colorado at Boulder
UCB 419
Boulder, CO 80309-0419
Tel: 1-303-735-6385
kornish@colorado.edu

Moren Lévesque
Faculty of Engineering
University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada N2L 3G1
Tel: 1-519-888-4567 ext 35367
levesque@uwaterloo.ca

Lisa M. Maillart
Industrial Engineering Department
University of Pittsburgh
Pittsburgh, PA 15261
Tel: 1-412-624-9845
maillart@pitt.edu

* Corresponding author

The authors are grateful to Michael Armstrong, Richard Cassady and Paul Zipkin for their invaluable comments and suggestions. The authors also gratefully acknowledge financial support from the Weatherhead School of Management and from an NSF ADVANCE Institutional Transformation Grant, SBE-0245054, Academic Careers in Engineering and Science (ACES) at Case Western Reserve University. Moren Lévesque further acknowledges financial support from a Discovery Grant (#341679) from the Natural Sciences and Engineering Research Council (NSERC) of Canada.
New Product Launch Date Decisions: Promotion and Production

Abstract

We study two interconnected decisions that firms make in preparation for the launch of a new product: How much of the product to have available at the launch date and how much to promote the new product. With limited production capacity, the production quantity decision implies a launch date. Our focus is on settings in which demand is uncertain, waiting to launch negatively affects demand (i.e., demand is eroding), and spending on promotion positively affects demand. We compare the optimal production quantity in a scenario where production is time consuming to that of a generic newsvendor setting. We address two main research questions. First, under what conditions does time-consuming production have the strongest effect on probability of shortage at launch? Second, what is the effect of the magnitude and source of demand erosion on the optimal promotion and production decisions? For the first question, we find that accounting for time-consuming production is most important when the ratio of units lost during delay to production is high, or when the uncertainty of demand for the product is high. For the second question, we find that the greater the magnitude of demand erosion, the less a firm should promote (and produce) in order to launch sooner. Refining that conclusion, we find that the more that demand erosion springs from competitive product introductions rather than a limited selling season, the more the firm should promote.

Key words: new product launch timing, optimal production, optimal promotion, newsvendor
1. Introduction

Apple’s iPhone 3G went on sale in twenty-one countries on July 11, 2008. Three days later, the company announced that one million units had been sold (Apple.com, 2008). Other news reports from the launch painted a familiar picture of high-profile product launches: long lines and some customers going away empty-handed (Allison et al., 2008; Moran, 2008). Some of the people empty-handed might have asked, “Why didn’t they produce more?”

That question has a ready answer from inventory theory: demand is uncertain and supply should be chosen to balance the cost of having excess with the (opportunity) cost of a shortfall. However, that answer omits timing considerations. In particular, it does not consider that additional production time would delay the product launch. Firms may be in a rush to get to market to hit a selling season or to stay ahead of competitors, and a launch shortage may be part of the trade-off for these goals. The “Why didn’t they…” question also omits another consideration, the pre-launch promotion of the product. In other words, the question is not simply “Why didn’t they produce more?” but also “Why did they promote the product so much if they were not going to produce more?”

Apple is certainly not the only company to experience a shortage at product introduction. Amazon introduced its Kindle e-book reader in 2007 with insufficient supply (Meyers, 2007). Video game consoles have a bad track record on this front: Microsoft had shortages of the Xbox 360 in 2005 and Nintendo had shortages of the Wii in 2006. The Toyota Prius launch in the U.S. in 2000 resulted in long waiting lists. Popular holiday toys, like the Tickle Me Elmo doll, most recently with the tenth anniversary edition (Binkley, 2006), and Cabbage Patch dolls in the 1980s, are also notorious examples of shortages.
With these products, why is there urgency to launch? Competitive threats are an important reason. These threats can be specific races, such as Sony’s and Microsoft’s head-to-head competition for video game consoles or XM’s and Sirius’ competition for the initial customers for satellite radio. Or, the competitive threats can be less well defined. For example, Apple was not in a horse race with Nokia for the newest smart phone but still feels pressure. For companies like Apple, we can attribute urgency to the competitive and dynamic consumer electronics market. Gadgets are likely to have high switching costs, due to service contracts, a learning curve, and customization (e.g., phone lists and calendar information). Apple hopes to capture new customers that may not be available later if they are preempted by the latest BlackBerry or Treo. Another scenario that creates urgency is a selling season. In particular, many companies are mindful of running out the clock on the crucial holiday season. In referring to recent experiences with holiday electronics and toy introductions, one writer quipped, “These spectacular sellouts have become as much of a holiday staple as turkey and stuffing” (Harford, 2005).

In this paper, we study two key decisions in a setting with product launch urgency: how much to produce for launch (which also determines when the launch will happen), and how intensively to promote the product prior to launch. Our model and analysis are geared toward answering the following research questions:

1. How different is the production quantity decision a firm would make when considering time-consuming production versus ignoring it? And how different is the resulting probability of shortage at launch? What intensifies these differences?
2. How is the optimal level of promotion affected by the degree to which it is beneficial to launch sooner rather than later, i.e., by the rate at which demand erodes over time? How do
the optimal promotion and production decisions depend on whether this erosion is driven by a limited selling season or competition?

To answer these questions, we build a single-period mathematical model of the production and promotion decisions. Demand is uncertain and its distribution is positively influenced by promotion, but negatively influenced by delay. In Section 2, we present related work from the literature and explain our contribution. In Section 3, we present our model of time-consuming production and demand-building promotion. In Section 4, we answer the research questions posed in this introduction. In Section 5, we discuss the managerial implications and close the discussion with concluding remarks.

2. Theoretical background and contribution

This paper offers two main contributions to the literature at the operations-marketing interface. The first contribution is an analysis of a newsvendor inventory problem with time-consuming production. This time-consuming production affects, in turn, the launch date, the demand distribution, and the optimal production quantity, so the standard newsvendor solution does not apply. The second contribution is the combination of this time-consuming production problem with an investment-in-promotion decision. Whereas existing work on time-consuming product development examines a balance between a better product and a more timely product launch, we examine a balance between a better promoted product, a more abundant product, and a more timely product launch.

In the classic newsvendor problem (e.g., Arrow et al., 1951), a quantity is chosen to balance the costs of unsold inventory with the costs of having too little inventory at a single
selling opportunity. Fisher et al. (2001) explain that a long lead time relative to the duration of the selling season is an implicit part of the production dilemma in the newsvendor setting. Traditionally, this lead time is fixed and, therefore, the demand distribution does not depend on the production quantity. In our model, however, because production is time-consuming and demand deteriorates as time passes, the launch-date demand distribution does change with production quantity. Even newsvendor research that considers multiple production opportunities, such as Fisher and Raman (1996), Fisher et al. (2001), Milner and Kouvelis (2002, 2005), and Li et al. (2006), assume a fixed lead time for each lot regardless of its size.

Our treatment of lead time as proportional to quantity also distinguishes our work from many other classes of inventory problems. There are streams of work that consider stochastic lead times (e.g., Kaplan, 1970; Song, 1994; Song and Zipkin, 1996a, 1996b; Zipkin, 2000 chapter 7), however again, the lead time is associated only with the arrival of the order, not the size of the order. Two additional exceptions to the fixed-lead-time assumption are the queueing models of Zipkin (1986) and Karmarkar (1987). The former mainly focuses on the case in which processing time and order quantity are independent and argues that if these quantities are dependent then the convexity of the cost function is threatened. Processing time plays a more central role in Karmarkar (1987), where it is modeled as a linear function of batch size. Other works at the operations-marketing interface focus not on lead time, but on the interplay between production and word-of-mouth (diffusion) processes. Ho et al. (2002) and Kumar and Swaminathan (2003) address the question of production planning, including inventory build-up before launch. They both analyze the effect of constrained production capacity on the time-path of sales using a deterministic process for demand.
Of course, we are not the first to study launch timing for a new product. Launch timing has received considerable attention in both the operations and marketing literatures; see e.g. Armstrong and Lévesque, 2002; Bayus, 1997; Bayus et al., 1997; Cohen et al., 1996, 2000; Klastorin and Tsai, 2004; Lippman and Mamer, 1993; Morgan et al., 2001; Souza et al., 2004. In these papers, production is either ignored or considered to be instantaneous. For example, in Souza et al. (2004) an inventory decision is considered, but production is instantaneous. In that literature, the basic tension arises from the idea that by spending a longer time on product development, the development costs more and the competitive position is weakened, but the product is better and therefore more attractive to the market. Our paper treats the related problem in which improving the attractiveness of the product in the market place does not delay launch, but production does.

Furthermore, many authors consider means of manipulating the demand distribution in a newsvendor framework, most notably via price. Petruzzi and Dada (1999) review and extend that body of work (see also Porteus, 1990). We treat price as fixed, and focus on other ways the firm influences demand. Also related to our approach is the so-called inverse newsvendor problem, as in Carr and Lovejoy (2000), in which the firm has a given, but uncertain supply and selects a demand distribution. The major difference between Carr and Lovejoy’s work and ours is that in our model the launch date affects supply, whereas in their model the supply, although random, is not alterable. Rao et al. (2005) focus on lead time, but they look at an optimal lead time guarantee across customers. Like our work, Dana and Petruzzi (2001) capture a link between the quantity produced and the demand distribution. However, they model a positive relationship between inventory levels and demand, whereas in our framework a higher inventory level implies a later launch, with a worse demand distribution. To our knowledge, no one has studied
the newsvendor variant in which production is time-consuming and therefore influences the
selling date and the demand distribution.

### 3. Model

We consider a firm launching a new product and facing two decisions: a level of
promotional spending, \( k \), and an inventory level to be available at launch, \( q \). Production occurs at
a constant rate \( r \), and hence the inventory decision implies a launch date, which we denote
\( \ell \equiv \ell(q, r) = q/r \). This model is a variant of the newsvendor inventory model, adapted for time-
consuming production and demand-enhancing promotion.

The random demand for the product on the launch date is \( X \), with probability density
function \( f(x; k, \ell(q, r)) \), cumulative distribution function \( F(x; k, \ell(q, r)) \), and \( X > 0 \). The unit
price and unit production costs are given by \( p \) and \( c \), respectively, with \( p > c \). The promotional
spending \( k \) is a cost. Therefore, the firm’s profit on the launch date, with actual demand \( x \), is

\[
p \min\{x, q\} - cq - k
\]

and the expected profit, taking into account the demand uncertainty, is

\[
\pi(q, k) = -cq + \int_0^q px f(x; k, \ell(q, r))dx + \int_q^\infty p(q, r)dx - k.
\]

Defining the mean of the demand distribution as \( \mu(k, \ell(q, r)) \), (2) can be rewritten as

\[
\pi(q, k) = -cq + p\mu(k, \ell(q, r)) - p\int_q^\infty (x-q) f(x; k, \ell(q, r))dx - k.
\]

We further make the simplification that only the mean of the demand distribution, and
not the shape, depends on \( k \) and \( q \). This assumption allows us to work with any shape
distribution, and implies that changes in promotion and launch date shift the density function.
Therefore, we define $X \equiv \mu(k, \ell(q, r)) + Y$ and, without loss of generality, $E[Y] = 0$. We denote the probability density function for $Y$ as $h(Y)$, independent of $q$ and $k$, and the cumulative distribution function as $H(y)$. The variance of demand is thus constant in $q$ and 

$$F(\delta; k, \ell(q, r)) = H(\delta - \mu(k, \ell(q, r))),$$

or $f(\delta; k, \ell(q, r)) = h(\delta - \mu(k, \ell(q, r)))$, for all $\delta \geq 0$ and zero otherwise. Rewriting the objective function using this change of variable, we obtain

$$\pi(q, k) = -cq + p\mu(k, \ell(q, r)) - p\int_{q-\mu(k, r)}^{\infty} (y + \mu(k, \ell(q, r)) - q)h(y)dy - k.$$ 

The optimal levels of production quantity $q$ and promotional spending $k$ are determined by the two first order conditions

$$\frac{\partial \pi}{\partial q} = -c + p\left(1 - \frac{1}{r} \frac{\partial \mu(k, \ell)}{\partial \ell}\right)F(q; k, \ell) = 0 \quad (3)$$

and

$$\frac{\partial \pi}{\partial k} = p\frac{\partial \mu(k, \ell)}{\partial k}F(q; k, \ell) - 1 = 0 \quad (4)$$

The second order conditions for optimality are

$$\frac{\partial^2 \pi}{\partial q^2} < 0 \quad (5)$$

and

$$\frac{\partial^2 \pi}{\partial q^2} \frac{\partial^2 \pi}{\partial k^2} \left[\frac{\partial^2 \pi}{\partial q \partial k}\right]^2 > 0 \quad (6)$$

the conditions for the Hessian matrix to be negative semi-definite.\(^1\)

---

\(^1\) Following the approach of Whitin (1955) and Zabel (1970), which Petruzzi and Dada (1999) explain can be used to analyze newsvendor problems in which both price and quantity are decision variables by considering one variable (quantity) as a function of the other (price, or in our case, promotion) and subsequently solving for the other, the two second order conditions become
We analyze a setting in which promotion enhances the attractiveness of the product, and we capture that idea by assuming that the mean demand $\mu(k, \ell)$ increases in the level of promotional spending $k$. We also assume that the mean demand decays during production, i.e., during the time until launch. Formally, (assuming differentiability) these conditions can be expressed by $\partial \mu / \partial k \geq 0$ and $\partial \mu / \partial \ell \leq 0$, respectively. Finally, we define $\rho$ as the probability of incurring a shortage on the launch date, that is, $\rho \equiv 1 - F(q; k, \ell)$.

Before characterizing the optimal production quantity and promotion decisions, a few additional comments that extend the applicability of our decision framework are warranted. Some of the most prominent examples of shortages are based around the holiday cycle, which suggests situations in which the firm does not select a launch date, but rather has it dictated by that cycle (the hard deadline is similar to the structure in Kornish and Keeney, 2008). However, even with a seasonal deadline, the firm has a choice on when in a season to launch. Almost every deadline, holiday seasonal ones included, is in fact a “soft” deadline. By a soft deadline, we mean that there is no specific “drop dead” date. There may be target dates the firm is aiming for, but while there may be penalties for missing targets, the value of the new product does not become zero if a target is not met. The soft deadline idea is consistent with our model of mean demand decreasing as the launch date is delayed.

Further, although we have framed the question as when to launch the product based on production level and production rate, we could equivalently work backward from a hard

$$\frac{1}{r^2} \frac{\partial^2 \mu(k, \ell)}{\partial \ell^2} < \frac{f(q; k, \ell)}{F(q; k, \ell)} \left(1 - \frac{1}{r} \frac{\partial \mu(k, \ell)}{\partial \ell}\right) \quad \text{and} \quad \frac{1}{r} \frac{\partial \mu(k, \ell)}{\partial \ell} \left(1 - \frac{1}{r} \frac{\partial \mu(k, \ell)}{\partial \ell} \right) + \frac{1}{r} \frac{\partial^2 \mu(k, \ell)}{\partial \ell^2} + \frac{1}{r} \frac{\partial \mu(k, \ell)}{\partial k} \left(1 - \frac{1}{r} \frac{\partial \mu(k, \ell)}{\partial \ell} \right) \frac{\partial^2 \mu(k, \ell)}{\partial k \partial \ell} + \frac{1}{r} \frac{\partial^2 \mu(k, \ell)}{\partial k^2} \frac{\partial \mu(k, \ell)}{\partial \ell} \frac{\partial \mu(k, \ell)}{\partial k} \frac{\partial \mu(k, \ell)}{\partial \ell}$$

where

$$\left(1 - \frac{1}{r} \frac{\partial \mu(k, \ell)}{\partial \ell}\right) \frac{\partial^2 \mu(k, \ell)}{\partial k^2} \frac{\partial \mu(k, \ell)}{\partial \ell} \frac{\partial \mu(k, \ell)}{\partial k} \frac{\partial \mu(k, \ell)}{\partial \ell} > 0$$

$\ell^*$ and $q^*$ are the optimal solutions that come from solving (3).
deadline. That is, given a launch date, we could frame the question as one of when to start production and implement the promotion decision, after which demand will stochastically deteriorate leading up to the launch. Posed this way, the questions look more like those of Milner and Kouvelis (2002, 2005), who study order timing, as opposed to our launch timing. However, they consider a fixed time between order placement and arrival, while we consider lead time to be proportional to the order quantity.

4. Analysis

In this section we address the research questions posed in the introduction.

4.1. The importance of capturing time-consuming production

To isolate the effect of adding time-consuming production to the standard newsvendor model, we examine the production quantity decision \( q \), treating the level of promotional activity \( k \) as exogenous. Rewriting (3), we obtain

\[
F(q; k, \ell) = \frac{1 - \frac{c}{p}}{1 - \frac{1}{r} \frac{\partial \mu(k, \ell)}{\partial \ell}}. \tag{7}
\]

Ignoring production time is equivalent to assuming \( r = \infty \), in which case the denominator of the right hand side of (7) would be 1. That yields the standard newsvendor solution (the optimal quantile of the demand distribution is \( 1 - \frac{c}{p} = \frac{p - c}{p} \)). When we recognize that \( r < \infty \), \( \frac{\partial \mu}{\partial \ell} \) plays a crucial role. This term compares the rate at which expected demand erodes to the rate of production. (Note that for production to be attractive at all, we must have \( \frac{\partial \mu}{\partial \ell} < r \).)
For example, with the Apple iPhone, if expected demand is dropping by 200,000 units a month—because those people will not wait for the iPhone, but would rather select another phone—and production is 500,000 units per month, then the change in optimal quantity is significant when considering time-consuming production. The optimal quantile of the demand distribution changes from \(1 - \frac{c}{p}\) in the standard newsvendor analysis to

\[
\frac{1 - \frac{c}{p}}{1 + \frac{200,000}{500,000}} = \frac{5}{7} \left(1 - \frac{c}{p}\right).
\]

The probability of having sufficient supply therefore decreases by 2/7 or 28.6%. That is, if the probability of having sufficient supply \((1 - \rho)\) was 70% using the standard \(r = \infty\) assumption, then accounting for time-consuming production means that the probability of sufficient supply is now only 5/7 of 70% or 50%. The probability of shortage thus increases from 30% to 50%.

The optimal quantity produced can also be dramatically affected. In general, the greater the spread of the demand distribution (in the sense of a “mean-preserving spread” from Rothschild and Stiglitz, 1970), the less steep the cumulative distribution function (CDF) of demand, and the larger the absolute effect of a change in optimal quantile (i.e., what is measured on the vertical axis of the CDF) has on the change in optimal quantity (i.e., what is measured on the horizontal axis of the CDF).

In sum, considering time-consuming production is the most crucial when the ratio of units lost during delay to units produced during delay (i.e., the production rate) is high. The

---

These numbers are rough estimates based on Apple’s sales goals (as reported in Allison and Minto, 2008) and Apple’s competitive position in the market, with roughly 0.2% market share in smart phones (NetApplications 2008).
probability of having sufficient supply drops by a factor of \( r \frac{r}{r - \partial \mu / \partial \ell} \) relative to the standard newsvendor solution (that factor times the newsvendor solution \( 1 - \frac{c}{p} \) gives the right-hand side of (7)). These consequences are most dramatic when there is greater uncertainty about demand.

4.2. The effect of the magnitude and source of demand erosion

In the previous section, we examined the effect of time-consuming production on the optimal production level and the probability of shortage. In this section, we examine the effect of time-consuming production and the resultant demand erosion on the optimal level of promotion. A larger magnitude of demand erosion corresponds to a larger value of \( \left| \frac{\partial \mu(k, \ell)}{\partial \ell} \right| \) (recall that launch delay reduces demand because \( \partial \mu / \partial \ell \leq 0 \)), the rate at which the mean demand drops with delay in product launch. We first demonstrate that the higher the magnitude of demand erosion, the lower the optimal promotion. Second, we show that the source of demand erosion matters. If the erosion comes from a limited selling season, then the optimal promotion is lower and the launch date is sooner than if the erosion comes from a competitive environment.

We establish the first result as follows. We derive conditions for a positive relationship between the optimal levels of promotion and production. Then, because a higher magnitude of demand erosion implies a lower optimal production level, we conclude that the higher the magnitude of demand erosion, the lower the optimal promotion.

The first step, i.e., the positive relationship between promotion and production, comes from a comparative statics analysis (e.g., Varian 1992: 492). With a change in any parameter \( \theta \), the optimal promotion decision \( k \) changes according to
\[
\frac{dk^*}{d\theta} = \left[ \frac{\partial^2 \pi(k^*; \theta)}{\partial k \partial \theta} \right] \int \left[ -\frac{\partial^2 \pi(k^*; \theta)}{\partial k^2} \right].
\]

It follows that, under (6), the denominator is positive, therefore the sign of \( \frac{dk^*}{dq} \) equals that of

\[
\frac{\partial^2 \pi(q, k^*)}{\partial k \partial q} = p \left[ \frac{1}{r} \frac{\partial^2 \mu(k^*, \ell)}{\partial k \partial \ell} F(q; k^*, \ell) + \frac{\partial \mu(k^*, \ell)}{\partial k} f(q; k^*, \ell) \left( 1 - \frac{1}{r} \frac{\partial \mu(k^*, \ell)}{\partial \ell} \right) \right],
\]

which is positive given \( \partial \mu/\partial k \geq 0, \partial \mu/\partial \ell \leq 0 \) and \( \partial^2 \mu/\partial k \partial \ell \geq 0 \). The first two inequalities were given earlier: promotion increases mean demand, delay reduces it. The third inequality is a reasonable second-order effect: promotion mitigates reduction in mean demand from delay. With these conditions and (6), we obtain \( \frac{\partial^2 \pi(q, k^*)}{\partial k \partial q} \geq 0 \), which implies a higher optimal level of promotion for a higher level of production. Note that the reverse is also true—higher optimal production for higher promotion—under (5) as the numerator of \( \frac{dq^*}{dk} \) is the same as (8) and the denominator is positive by (5).

In section 4.1, we saw that the higher the magnitude of demand erosion \( \partial \mu/\partial \ell \), the lower the optimal production. From (7), we have

\[
F(q^*; k, \ell^*) = \frac{1 - \frac{c}{p}}{1 - \frac{1}{r} \frac{\partial \mu(k, \ell^*)}{\partial \ell}}.
\]

At any level of \( \ell = q/r \), a larger \( \frac{\partial \mu(k, \ell^*)}{\partial \ell} \) yields a lower \( F(q^*; k, \ell^*) \). Because \( F \) is increasing in its first argument, a larger \( \frac{\partial \mu(k, \ell^*)}{\partial \ell} \) yields a lower \( q^* \).

Our conclusion is that the faster the demand is eroding, the more quickly the firm should launch the product, and the firm should not waste too much money promoting the product. The
quick launch will imply a lower level of production, and it does not make sense to excessively build up demand with promotion if there will be insufficient production to satisfy it.

Further, we examine how the source of demand erosion matters. We compare the case of erosion due to a limited selling season to that due to a competitive environment. We capture the former case with $\frac{\partial^2 \mu}{\partial k \partial \ell} = 0$. If demand is related to the percentage of the season the product is available in the market, then promotion does not mitigate the effect of launch date delay. We capture the latter case with $\frac{\partial^2 \mu}{\partial k \partial \ell} > 0$, that is, promotion does mitigate the effect of delay. For example, Apple’s advertising may diminish some consumers’ interest in Treos or BlackBerries, giving Apple more leeway in launch timing.

We unify these two cases, $\frac{\partial^2 \mu}{\partial k \partial \ell} = 0$ and $\frac{\partial^2 \mu}{\partial k \partial \ell} > 0$, by looking at the effect of $\frac{\partial^2 \mu}{\partial k \partial \ell}$ on the optimal decisions. To simplify the analysis, we assume a constant $\frac{\partial^2 \mu}{\partial k \partial \ell}$ (i.e., $\mu(k, \ell)$ has a $k\ell$ term but not higher order terms including that product). This simplification is sufficient but not necessary for our result and allows us to examine the effect of $\frac{\partial^2 \mu}{\partial k \partial \ell}$ on the optimal decisions. In the appendix, we prove that with constant and non-negative $\frac{\partial^2 \mu}{\partial k \partial \ell}$ and $\frac{\partial^2 \mu}{\partial k^2} < 0$ (i.e., promotional dollars have a decreasing marginal effect on mean demand), then the optimal level of promotion $k^*$ is increasing in $\frac{\partial^2 \mu}{\partial k \partial \ell}$.

Our conclusion is an extension of our result above about the negative relationship between the magnitude of demand erosion and optimal production. We have now shown how this negative relationship is affected by a change in the source of demand erosion. There is a
lower optimal level of promotion when demand erosion is due to a limited selling season rather than a competitive environment. In fact, the stronger the competitive environment, i.e., the higher $\frac{\partial^2 \mu}{\partial k \partial \ell}$, the higher the optimal level of promotion, the larger the optimal production, and the later the launch date.

5. Conclusion

5.1. Managerial takeaways

What did we learn from analyzing the effects of promotion and time-consuming production on key decisions for product launch? First, the importance of the launch timing depends crucially on the ratio of units lost during delay to production rate. The higher this ratio, the more important it is to account for time-consuming production.

Second, the effect of time-consuming production on the optimal quantity is greater when there is more uncertainty in the demand distribution. With great uncertainty (in the sense of a mean-preserving spread), small changes in the optimal quantile translate to large changes in the optimal quantity. Therefore, it is most important to consider time-consuming production for a truly novel product where demand can vary widely.

Third, we find that greater promotion is the not best response to greater magnitude of demand erosion. Instead, the greater that magnitude, the less a firm should promote (and, the less it should produce) in order to launch sooner. However, the more that the demand erosion springs from competitive product introductions rather than a limited selling season, the higher the optimal level of promotion. In those competitive situations, firms should buy production time with promotion, allowing them to produce more for a later launch.
Finally, we see a correspondence between highly promoted launches and higher quantity available at launch. Thus, even though there are notable examples of high-profile launch-date shortages, it is not a foregone conclusion that more promotion will lead to a higher chance of stockout. In some cases, the increased quantity and the increased promotion cancel each other out with regards to shortage probability. Overall, these conclusions give a fuller picture of how to think about promotion and production for a launch date when production is time-consuming due to limited capacity.

5.2. Limitations and future work

In the introduction, we suppose that empty-handed customers are asking, “Why didn’t they produce more?” We also propose an extended version of the question, “Why did they promote the product so much if they were not going to produce more?” The model we presented in this paper highlights the balance between the two implied decisions: production quantity and promotion. There are other considerations that we did not address, such as the time value of money, replenishment, and pricing. We comment on each of those issues next.

The time value of money, or discounting, is also an answer to the question of why firms will launch products to a shortage. Dollars earned earlier are worth more than dollars earned later, encouraging firms to hurry their launches. Including this element in the analysis would strengthen the idea that firms often have an urgency to launch their product. The discount factor is likely quite small (say, less than 1% a month), while the lost demand and increased production can be much more significant factors. Our work shows that discounting is not necessarily the major driver for firms’ concern with launch timing.
Another issue we did not model is replenishment. We have analyzed the firm’s decisions for a launch event, but not subsequent restocking of inventory. While this is a simplification, we maintain that there is an importance around the launch event that makes it an acceptable simplification. In many cases, such as the consumer electronics and holiday toy examples in the introduction, firms invest a lot in making customer aware of the product launch and in working with channel partners to make sure they are ready for it. A model with replenishment would be like a two-period (or multi-period) version of our model, presumably with learning about the residual demand distribution after the first period. We would expect many of the same conclusions to hold, such as the centrality of the ratio of lost demand during delay to the production rate and the prescription that stronger timing urgency leads to less promotion and sooner introduction and replenishment.

A third issue is pricing. In addition to the questions in the minds of customers that we pose above, another way to frame the question would be to ask, “Why didn’t they charge more?” A shortage occurs when the quantity demanded at a given price is less than the quantity supplied. A higher price would bring those two quantities in line. However, when demand is uncertain, there is almost never an exact match between quantity supplied and demanded. At any price, shortage is a possibility (as is excess). Firms do have the option of dynamically adjusting their prices in response to demand: raise it with evidence of strong demand and lower it with evidence of weak demand. This type of revenue management seems to work in some categories, such as hotel rooms and airlines seats, where the pricing structure is far from transparent. However, such a dynamic strategy is risky in a launch of a physical product that has a clearly defined price. Dynamic price adjustments could be seen as firms “taking advantage” of customers, thereby taking an event that can be construed as positive news—“there is so much demand for our
product we sold out”—and turning it into negative news—“the heartless firm gouges their loyal customers.”

These limitations are possible directions for future work, further demonstrating that time-consuming production is an important consideration in launch date decisions about promotion and production.
Appendix for Section 4.2

Proposition: For constant $\frac{\partial^2 \mu}{\partial k \partial \ell}$, if $\frac{\partial^2 \mu}{\partial k \partial \ell} \geq 0$ and $\frac{\partial^2 \mu}{\partial k^2} < 0$, then $k^*$ is increasing in $\frac{\partial^2 \mu}{\partial k \partial \ell}$.

Proof: Combining the first-order conditions (3) and (4) yields

$$\frac{1}{p \frac{\partial \mu(k^*, \ell^*)}{\partial k}} = \frac{1 - \frac{c}{p}}{1 - \frac{1}{r} \frac{\partial \mu(k^*, \ell^*)}{\partial \ell}}. \tag{9}$$

The slope of the left-hand side of (9) with respect to $k$ is $\frac{1}{p \left(\frac{\partial \mu}{\partial k}\right)^2}$, which is positive when $\frac{\partial^2 \mu}{\partial k^2} < 0$. The slope of the right-hand side of (9) is $\frac{\left(1 - \frac{c}{p}\right) \left(\frac{\partial^2 \mu}{\partial \ell \partial k}\right)}{\left(1 - \frac{1}{r} \frac{\partial \mu}{\partial \ell}\right)^2}$, which is zero when $\frac{\partial^2 \mu}{\partial k \partial \ell} = 0$, positive when $\frac{\partial^2 \mu}{\partial k \partial \ell} > 0$, and increasing in $\frac{\partial^2 \mu}{\partial k \partial \ell}$. Further, the value of the right-hand side at $k = 0$ is independent of $\frac{\partial^2 \mu}{\partial k \partial \ell}$, when $\frac{\partial^2 \mu}{\partial k \partial \ell}$ is a constant. Therefore, increases in $\frac{\partial^2 \mu}{\partial k \partial \ell}$ pivot the right-hand side counterclockwise around the y-intercept. Because the left-hand side is increasing in $k$, the result is a higher point of intersection $k^*$ for a higher $\frac{\partial^2 \mu}{\partial k \partial \ell}$. Figure 1 illustrates these relationships.
**Figure 1:** This figure shows a schematic of the relationships described in the proof above. The conclusion is that \( k^* \) is increasing in \( \frac{\partial^2 \mu}{\partial k \partial \ell} \). The expressions for the left hand side and right hand side of equation (9) are not restricted to be linear in \( k \), as shown in the figure, but they are both increasing in \( k \) as described in the proof.
References


Bayus BL, Jain S, Rao AG. Too little, too early: Introduction timing and new product performance in the personal digital assistant industry. Journal of Marketing Research 1997; 34(1); 50-63

Binkley C. What’s the newest Elmo’s gimmick? If we told you, we’d have to kill you. The Wall Street Journal September 15, 2006; B1.

Carr S, Lovejoy W. The inverse newsvendor problem: Choosing an optimal demand portfolio for capacitated resources. Management Science 2000; 46(7); 912-927.


Dana JD, Petruzzi NC. The newsvendor model with endogenous demand. Management Science 2001; 47(11); 1488-1497.


Kaplan RS. A dynamic inventory model with stochastic lead times. Management Science 1970; 16(7); 491-507.

Karmarkar U. Lot sizes, lead times and in-process inventories. Management Science 1987; 33(3); 409-418.


Kornish LJ, Keeney RL. Repeated commit-or-defer decisions with a deadline: The influenza vaccine composition.” Operations Research 2008; 56(3); 527-541.


Milner JM, Kouvelis P. Order quantity and timing flexibility in supply chains: The role of demand characteristics. Management Science 2005; 51(6); 970-985.

Moran R. New iPhone may be faster, but lines for it still slow. Philadelphia Inquirer; July 12, 2008.

Morgan LO, Morgan RM, Moore WL. Quality and time-to-market trade-offs when there are multiple product generations. Manufacturing & Service Operations Management 2001; 3(2); 89-102.


Song J-S, Zipkin PH. Managing inventory with the prospect of obsolescence. Operations Research 1996a; 44(1); 215-222.

Song J-S, Zipkin PH. Inventory control with information about supply conditions. Management Science 1996b; 42(10); 1409-1419.


Whitin TM. Inventory control and price theory. Management Science 1955; 2(1); 61-68.


Zipkin PH. Models for design and control of stochastic, multi-item batch production systems. Operations Research 1986; 34(1); 91-104.