Optimizing a Tradeoff:

Production Efficiency

versus

Inventory Obsolescence

Combining Make-to-Order and

Make-to-Stock Production

Ely Dahan and Glen Schmidt

January 8, 1999

Table of Contents

Abstract:
We consider a single-period, single product decision by a profit-maximizing firm facing uncertain demand. The firm must make two decisions; (1) select a capacity level, and (2) allocate it between low-cost mass production and higher-cost make-to-order production. We show that both solutions take the form of a critical fractile solution to different Newsvendor problems. A boundary on marginal capacity cost exists beyond which make-to-stock production is strictly favored. If product obsolescence is relatively costly and technology lowers the cost of make-to-order production, then make-to-order production is favored. As improved information reduces the degree of uncertainty about demand, make-to-stock production becomes more attractive once again.

Introduction:

Recently, several PC manufacturers have announced their intention to convert at least part of their manufacturing capacity from mass production which for simplicity we characterize as make-to-stock, to a form of mass customization which we characterize as make-to-order. Specifically, they have decided to produce PC's *after* having received orders from customers. In the past, these firms built products to stock and filled orders from inventory.

This shift in manufacturing strategy raises some interesting questions:

- Why should a firm choose to devote some of its capacity to a make-to-order strategy?
- Why did PC firms wait until now to adopt a make-to-order strategy?
- Will future trends in production technology and information gathering *necessarily* favor a shift towards mass-customization?
- Could the make-to-order strategy be justified even if the product being delivered were *identical* to the mass-produced product?
- How much capacity of each type (make-to-stock and make-to-order) is optimal in the face of uncertain demands? How much should be produced in the period?
- How does the optimal allocation of capacity and the production decision change as product obsolescence accelerates? As technology improves (so that make-to-stock doesn't cost much more than mass production? As a function of the cost of capacity?

To answer these questions, we consider a simplified model of a firm making a single period, single product decision on quantity. The firm seeks to maximize its total profit by optimizing two tradeoffs. First, it selects a total capacity level, K, by comparing the cost of that capacity with the benefit of being able to fill uncertain demand.

Second, the firm decides the fraction of the capacity to allocate between the low direct cost mass production (α) and make-to-order production ($1 - a$). The cost of producing one unit to stock is c_s and the cost of producing one unit to order is c_o, with $c_o > c_s$. (Clearly, if make-to-order production costs the same or less as making-to-stock, it would be the preferred method since it would eliminate the cost of overstocks without adding any other costs.)

For each unit sold, the firm realizes revenue of p, which can be thought of as the revenue in the current period plus the net present value of the gross margin from future sales. For each unit left in stock after demand is realized, the firm incurs a holding cost h that represents the expected cost of obsolescence and inventory holding. The cost of K units of capacity (of either type of production) is $q(K)$, which is convex in K. Demand is characterized as a non-negative, continuous random variable with probability density $f(K)$ and cumulative density $\Phi(K)$.

The analysis proceeds along the line of a traditional Newsvendor problem, but with the added feature of two sets of overage and underage costs. An objective function that reflects the two stage decision is optimized using first order conditions. The objective function, $F(\alpha, K)$, incorporates decision variables for total capacity and the allocation of that capacity to stock- and order-production.

Model Parameters:

$K = \text{capacity}$

$q(K) = \text{total cost of capcity K (in \$ per period)}$

$a = \text{\% of capacity that is make} - \text{to} - \text{stock}$

$p = \text{price / unit}$

$c_s = \text{unit cost for stock production}$

$c_o = \text{unit cost of make} - \text{to} - \text{order production}$

$h = \text{holding cost / unit (obsolescence cost)}$

$f(x) = \text{demand distribution probability density}$

$\Phi(x) = \text{demand distribution cumulative density}$

The objective function maximizes the sum of five terms:

- Low Demand Gross Profit: The expected amount realized given that demand was satisfied fully with stock production (with some inventory left over).
- +Medium Demand Gross Profit: The expected gross profit given that demand was met out of stock production and some make-to-order.
- + High Demand Gross Profit: The expected profit made from both kinds of production given that a complete stock-out occurred.
- - Low Demand Holding Costs: The expected holding and obsolescence costs given that stock production exceeded the demand for the period.
- - Capacity Costs. The total cost of investing in capacity level K.

It is expressed mathematically below:

Objective Function:

$$F(a,K) = \max_{a,K} \int_0^{aK} (p-c_s)xf(x)dx + \int_{aK}^{K} [(p-c_o)(x-aK)+(p-c_s)aK]f(x)dx$$
$$+ \int_K^{\infty} K[a(p-c_s)+(1-a)(p-c_o)]f(x)dx - \int_0^{aK} h(aK-x)f(x)dx - q(K)$$

Now we seek the first order conditions for a maximum, so we differentiate the objective with respect to the two decision variables (see Appendix). The first order condition with respect to α yields the following result.

Critical Fractile for Make-to-Stock Percentage:

$$\Phi(aK) = \frac{c_o - c_s}{h + c_o - c_s} \qquad (1)$$

This is an interesting result. It says that the proportion of capacity that should be reserved for stock production is the critical fractile solution to a particular Newsvendor problem. The cost of underage is the difference between the unit costs of order- versus stock-production. The overage cost is the holding cost per unit which occurs when stock production exceeds demand.

As production technology improves, and make-to-order production is very cost competitive with stock production, the critical fractile decreases and so does the proportion of make-to-stock production for a given capacity. This provides one reason that the PC firms moved to last minute production: new technology enabled the move.

Also, as holding and obsolescence costs increase, the critical fractile decreases, again favoring last minute production. When Compaq and IBM got stuck with excess 80486 PC inventory, they may have realized that inventory holding costs were simply too high. This provides a second motivation for the move towards make-to-order production.

The first order condition with respect to K also yields interesting results.

Critical Fractile for Total Capacity

$$q'(K) = (p - c_0) \times \left[1 - \Phi(K)\right] \tag{2}$$

This second result tells us that total capacity will be chosen so that the marginal cost of capacity, $q'(K)$, will equal $(p - c_0) \times \left[1 - \Phi(K)\right]$, which is the expected revenue from one more unit of make-to-order sales. Equation (2) can also be rewritten as a critical fractile:

$$\Phi(K) = \frac{p - c_o - q'(K)}{\left(p - c_o - q'(K)\right) + q'(K)}$$

Here we see that the cost of underage of capacity is the lost profit of one unit of make-to-order production minus the savings on the capacity itself. The reason that profits are determined by c_o is that whenever α is less than 1, as assumed here, the next unit of capacity must be make-to-order. If α is equal to 1, that is if all capacity is make-to-stock, then the entire problem reduces to another Newsvendor solution as shown in the next section.

Numerical Example:

$q(K)$ = total cost of capcity K (in \$ per period) = \$3K

p = price / unit = \$50

c_s = unit cost for stock production = \$25

c_o = unit cost of make – to – order production = \$40

h = holding cost / unit (obsolescence cost) = \$10

$\Phi(x)$ = demand distribution cumulative density ~ $Uniform(0,100)$

$f(x)$ = demand distribution probability density = 0.01

$$\Phi(K^*) = \frac{p - c_o - q'(K)}{\left(p - c_o - q'(K)\right) + q'(K)} = \frac{\$50 - \$40 - \$3}{\left(\$50 - \$40 - \$3\right) + \$3} = \frac{\$7}{\$10} = 0.7, \text{so } \mathbf{K^* = 70}$$

$$\Phi(a^*K^*) = \frac{c_o - c_s}{h + c_o - c_s} = \frac{\$40 - \$25}{\$10 + \$40 - \$25} = \frac{\$15}{\$25} = 0.6, \text{so } \boldsymbol{a^* = 6/7 = 86\%} \text{ make-to-stock.}$$

These results are illustrated in the graph on the next page. The objective function is maximized when $K = 70$ units of capacity and $a = 86\%$ of capacity dedicated to stock production. Note that the objective function does not appear to be concave for low values of α and K, but rather has a slight "bell" shape.

Global Optimality Requirements

It is not obvious from simple observation that the second order conditions for concavity of the objective function are met by the Hessian in the Appendix. In fact, there are additional conditions on the parameters of the problem to ensure that our optimality results hold.

One condition that must logically be met is that $\Phi(aK) \le \Phi(K)$ since α is the proportion of capacity allocated to make-to-stock production. This condition can be restated in the terms of equations (1) and (2) as follows:

$\Phi(aK) \le \Phi(K)$, but since $\Phi(aK) = \dfrac{c_o - c_s}{h + c_o - c_s}$ and $\Phi(K) = 1 - \dfrac{q'(K)}{(p - c_0)}$, we have that

$\dfrac{c_o - c_s}{h + c_o - c_s} \le 1 - \dfrac{q'(K)}{(p - c_0)}$, which can be rewritten in the following useful form:

Boundary Condition for Marginal Capacity Cost:

$$q'(K) \le \frac{h(p-c_o)}{h+c_o-c_s} \tag{3}$$

Equation (3) provides a convenient method of checking whether it is worthwhile to invest in make-to-order capacity. If the marginal cost of capacity is too high, that is if θ'(*K*) exceeds the ratio on the right hand side of equation (3), then α will equal 1 and the firm will make everything to stock.

In fact, assuming linear capacity costs, the problem reduces to the classic Newsvendor problem where the cost of overage is $h+q'(K)$, the sum of one unit of holding and wasted capacity costs. The cost of underage is $p-c_s-q'(K)$ since by under-stocking by one unit we lose one unit of profit, $p-c_s$, but we also save the cost of one unit of capacity, $q'(K)$.

Classic Newsvendor Problem When α = 100%

Another way of deriving this result is to add a constraint that α be less than or equal to 1 to our original optimization. This yields the following changes to equation (1):

$$\Phi(aK) = \frac{c_o - c_s - \frac{l}{K}}{h+c_o-c_s} \qquad (1)'$$

where λ is the LaGrange multiplier for the constraint on α. This equation reduces to (1) when the constraint is non-binding, that is when α is less than or equal to 1.

Given that the *constraint is binding*, i.e. that α=1, equation (2) becomes:

$$\Phi(K) = \frac{p-c_s-q'(K)}{[p-c_s-q'(K)]+[h+q'(K)]} = \frac{p-c_s-q'(K)}{p-c_s+h} \qquad (2)'$$

which is exactly the critical fractile solution for the Newsvendor problem.

Summary and Conclusions:

We have derived three simple results in the form of equations (1), (2) and (3). Equation (1) shows that the optimal allocation of make-to-stock versus make-to-order capacity takes the form of a critical fractile solution where overage costs come from excess capacity and inventory obsolescence and underage comes from lost profits.

Likewise, equation (2) reveals that the optimal total capacity is the critical fractile solution to a different Newsvendor problem. The optimal capacity can be shown to occur when marginal capacity cost is equated with the expected marginal revenue.

Finally, equation (3) sets a boundary on the parameters of the problem by recognizing that if marginal capacity costs are high enough, the firm will make everything to stock.

One troubling question remains: Doesn't the Hessian need to be negative semi-definite for the objective function to be concave? If the objective function is not concave, how are we sure that our solution is optimal?

Consider two possibilities: (a) the constraint on $\alpha \leq 1$ is binding so all production is make-to-stock, or (b) the constraint on α does not bind, i.e. $0 \leq \alpha < 1$. If condition (a) holds, then we have shown that the problem reduces to a classic Newsvendor problem which is concave in its objective, so the solution is a global maximum. In other words, the firm can choose any capacity level (assuming linear capacity costs). If (b) holds, then we know that aK can be determined independently of K since for any value of K, there exists aK such that equation (1) may be solved.

Thus, (1) and (2) may be solved as functions of single variables and we have shown that each is concave in its respective decision variable. So the solutions expressed in equations (1) and (2) are optimal if $\alpha < 1$, that is if at least some make-to-order production is suggested.

Our analysis provides insight into the PC firms' decision to convert some capacity to make-to-order production. Several reasons for the move to mass customization are suggested:
(a) obsolescence costs may have become more significant as product innovation accelerated,
(b) production technology improved to the point that the cost of last-minute production became competitive with stock production, and (c) the cost of marginal capacity declined so as to make additional make-to-order capacity economical.

Interestingly, these three justifications for make-to-order production do not depend on another natural advantage, namely the ability to customize products to meet specific customer needs. The advantages of product differentiation are in addition to the probabilistic advantages of responding to demand after it is known with certainty.

Appendix:

First Order Condition With Respect to Make-to-Stock Percentage (α):

We employ Liebnitz's rule (stated below) to differentiate each term of $F(\alpha,K)$.

If $F(\alpha)=\int_{a(\alpha)}^{b(\alpha)} f(x,\alpha)dx$, then $\frac{\partial F}{\partial \alpha}=\int_{a(\alpha)}^{b(\alpha)}\frac{\partial f}{\partial \alpha}dx + f(b(\alpha),\alpha)\frac{\partial b}{\partial \alpha} - f(a(\alpha),\alpha)\frac{\partial a}{\partial \alpha}$

For the first term, $a(\alpha)=0, \partial a/\partial \alpha = 0, b(\alpha)=\alpha K$, and $\partial b/\partial \alpha = K$, so the derivative is :

$0+(p-c_s)\alpha K^2 f(\alpha K)+0$

For the second term, $a(\alpha)=\alpha K, \partial a/\partial \alpha = K, b(\alpha)=K$, and $\partial b/\partial \alpha = 0$, so the derivative is :

$K(c_o-c_s)(\Phi(K)-\Phi(\alpha K))-(p-c_s)\alpha K^2 f(\alpha K)$

For the third term, $a(\alpha)=K, \partial a/\partial \alpha = 0, b(\alpha)=\infty$, and $\partial b/\partial \alpha = 0$, so the derivative is :

$\int_K^\infty K(p-c_s-p+c_o)f(x)dx = K(c_o-c_s)\times(1-\Phi(K))$

Finally, for the fourth term, $a(\alpha)=0, \partial a/\partial \alpha = 0, b(\alpha)=\alpha K$, and $\partial b/\partial \alpha = K$, so the derivative is : $-hK\Phi(\alpha K)$.

Assembling all four results gives the first order condition :

$$(p-c_s)\alpha K^2 f(\alpha K)+K(c_o-c_s)(\Phi(K)-\Phi(\alpha K))-(p-c_s)\alpha K^2 f(\alpha K)+ K(c_o-c_s)\times(1-\Phi(K))-hK\Phi(\alpha K)=0,$$

which gives the following as $\frac{\partial F(\alpha,K)}{\partial \alpha}$:

$K(c_o-c_s)(\Phi(K)-\Phi(\alpha K))+K(c_o-c_s)\times(1-\Phi(K))-hK\Phi(\alpha K)=0$, which simplifies to:

$$\Phi(\alpha K)=\frac{c_o-c_s}{h+c_o-c_s}$$

Second Order Condition with respect to α:

We calculate $\frac{\partial^2 F(\alpha,K)}{\partial \alpha^2}$ and $\frac{\partial^2 F(\alpha,K)}{\partial \alpha \partial K}$ for use in the Hessian to check concavity:

$$\frac{\partial F}{\partial \alpha}=K(c_o-c_s)(\Phi(K)-\Phi(\alpha K))+K(c_o-c_s)\times(1-\Phi(K))-hK\Phi(\alpha K)=0, \text{so}$$

$$\frac{\partial^2 F}{\partial \alpha^2}=-K^2(h+c_o-c_s)f(\alpha K)$$

$$\frac{\P^2 F}{\P a \P K} = (c_o - c_s)[(\Phi(K) - \Phi(aK)] + K(c_o - c_s)[(f(K) - a f(aK)] + (c_o - c_s) \times (1 - \Phi(K))$$
$$- K(c_o - c_s) \times f(K) - h\Phi(aK) - ahKf(aK)$$
$$= -(h + c_o - c_s)\Phi(aK) - aK(h + c_o - c_s) f(aK) + (c_o - c_s)$$

We note that if K were given, that is if total capacity were fixed, $\frac{\P^2 F(a,K)}{\P a^2}$ <0 shows that the objective is concave in α. More on this later.

<u>First Order Condition With Respect to Capacity *(K)*:</u>

Use Liebnitz's Rule to differentiate the objective function for first order conditions :

If $F(K) = \int_{a(K)}^{b(K)} f(x,K)dx$, then $\frac{\P F}{\P K} = \int_{a(K)}^{b(K)} \frac{\P f}{\P K} dx + f(b(K),K)\frac{\P b}{\P K} - f(a(K),K)\frac{\P a}{\P K}$

For the first term, $a(K) = 0, \P a / \P K = 0, b(K) = aK$, and $\P b / \P a = a$, so the derivative is : $(p - c_s)a^2 K f(aK)$

For the second term, $a(K) = aK, \P a / \P K = a, b(K) = K$, and $\P b / \P K = 1$, so the derivative is : $a(c_o - c_s)[\Phi(K) - \Phi(aK)] + Kf(K)[a(p - c_s) + (1 - a)(p - c_o)]$
$- (p - c_s)a^2 K f(aK)$

For the third term, $a(K) = K, \P a / \P K = 1, b(K) = \infty$, and $\P b / \P K = 0$, so the derivative is : $[a(p - c_s) + (1 - a)(p - c_o)] \times (1 - \Phi(K)) + 0$
$- Kf(K)[a(p - c_s) + (1 - a)(p - c_o)]$

For the fourth term, $a(K) = 0, \P a / \P K = 0, b(K) = aK$, and $\P b / \P K = a$, so the derivative is : $-ha\Phi(aK) - 0 + 0$

Finally, the derivative of the fifth term is : $-q'(K)$

Assembling all five results gives the first order condition :

$$(p - c_s)a^2 K f(aK) + a(c_o - c_s)[\Phi(K) - \Phi(aK)] +$$
$$Kf(K)[a(p - c_s) + (1 - a)(p - c_o)] - (p - c_s)a^2 K f(aK)$$
$$+ [a(p\text{-}c_s) + (1 - a)(p - c_o)] \times [1 - \Phi(K)]$$
$$- Kf(K)[a(p\text{-}c_s) + (1 - a)(p - c_o)] - ha\Phi(aK) - q'(K) = 0$$

which simplifies to :

$$\Phi(aK) = \frac{(p - c_o) \times [1 - \Phi(K)] + a(c_o - c_s) - q'(K)}{a[h + (c_o - c_s)]}$$

Recalling the previous result that $\Phi(aK) = \dfrac{c_o - c_s}{h + c_o - c_s}$, we have

$$\frac{c_o - c_s}{h + c_o - c_s} = \frac{(p - c_o)\times[1 - \Phi(K)] + a(c_o - c_s) - q'(K)}{a[h + (c_o - c_s)]} \Rightarrow$$

$$\boxed{q'(K) = (p - c_0)\times[1 - \Phi(K)]}$$

Second Order Condition with respect to K:

We again calculate $\dfrac{\partial^2 F(a,K)}{\partial K^2}$ and $\dfrac{\partial^2 F(a,K)}{\partial K \partial a}$ for use in the Hessian to check concavity:

$$\frac{\partial F}{\partial K} = -a[h + (c_o - c_s)]\Phi(aK) + (p - c_o)\times[1 - \Phi(K)] + a(c_o - c_s) - q'(K), \text{ so}$$

$$\frac{\partial^2 F}{\partial K^2} = -a^2[h + (c_o - c_s)]f(aK) - (p - c_o)f(K) - q''(K)$$

$$\frac{\partial^2 F}{\partial K \partial a} = -[h + (c_o - c_s)]\Phi(aK) - aK[h + (c_o - c_s)]f(aK) + (c_o - c_s),$$

which is exactly the same as the cross partial we had before $\left[\dfrac{\partial \mathrm{F}}{\partial a \, \partial \mathrm{K}}\right]$, as expected.

Hessian Matrix:

The second order conditions can be combined to form the following Hessian:

$$\begin{bmatrix} -K^2(h + c_o - c_s)f(aK) & \begin{matrix}(c_o - c_s) - [h + (c_o - c_s)]\Phi(aK) \\ -aK[h + (c_o - c_s)]f(aK)\end{matrix} \\ \begin{matrix}(c_o - c_s) - [h + (c_o - c_s)]\Phi(aK) \\ -aK[h + (c_o - c_s)]f(aK)\end{matrix} & \begin{matrix}-a^2[h + (c_o - c_s)]f(aK) \\ -(p - c_o)f(K) - q''(K)\end{matrix} \end{bmatrix}$$