THE VALUE OF SKU RATIONALIZATION IN PRACTICE (THE POOLING EFFECT UNDER SUBOPTIMAL INVENTORY POLICIES AND NONNORMAL DEMAND)*

JOSÉ A. ALFARO AND CHARLES J. CORBETT

University of Navarra, Faculty of Economics and Business Administration, Campus Universitario—Edificio de Bibliotecas (*Entrada Este*), 31080 *Pamplona* (*Spain*)
The Anderson School at University of California, *Los Angeles*, 110 *Westwood Plaza*, *Suite B*507, *Los Angeles*, *California* 90095-1481, *USA*

Several approaches to the widely recognized challenge of managing product variety rely on the pooling effect. Pooling can be accomplished through the reduction of the number of products or stock-keeping units (SKUs), through postponement of differentiation, or in other ways. These approaches are well known and becoming widely applied in practice. However, theoretical analyses of the pooling effect assume an optimal inventory policy before pooling and after pooling, and, in most cases, that demand is normally distributed. In this article, we address the effect of nonoptimal inventory policies and the effect of nonnormally distributed demand on the value of pooling. First, we show that there is always a range of current inventory levels within which pooling is better and beyond which optimizing inventory policy is better. We also find that the value of pooling may be negative when the inventory policy in use is suboptimal. Second, we use extensive Monte Carlo simulation to examine the value of pooling for nonnormal demand distributions. We find that the value of pooling varies relatively little across the distributions we used, but that it varies considerably with the concentration of uncertainty. We also find that the ranges within which pooling is preferred over optimizing inventory policy generally are quite wide but vary considerably across distributions. Together, this indicates that the value of pooling under an optimal inventory policy is robust across distributions, but that its sensitivity to suboptimal policies is not. Third, we use a set of real (and highly erratic) demand data to analyze the benefits of pooling under optimal and suboptimal policies and nonnormal demand with a high number of SKUs. With our specific but highly nonnormal demand data, we find that pooling is beneficial and robust to suboptimal policies. Altogether, this study provides deeper theoretical, numerical, and empirical understanding of the value of pooling.
(POOLING; STOCK KEEPING UNIT; SKU; RATIONALIZATION; PRODUCT VARIETY; IN-VENTORY MANAGEMENT; SUBOPTIMAL INVENTORY POLICY)

1. Introduction

The complexity of supply chains can be quite staggering, entailing high costs combined with incessant customer service problems. Increasing variety is a critical ingredient of this complexity, and managing this variety is becoming a field of research in its own right, as witnessed by (among others) the recent book edited by Ho and Tang (1998). Roughly speaking, variety can be induced by serving multiple geographic locations or stocking multiple products or stock-keeping units (SKUs). A growing body of literature exists on how

to deal with both types of variety, sometimes focusing on reducing lead times, sometimes on various types of aggregation of uncertainty. Mitigating the impact of variety in turn can occur through geographical "pooling" of inventories or through some form of rationalization of product lines. In both cases, the general mechanism underlying pooling or SKU rationalization is the ability to exploit "statistical economies of scale" by aggregating uncertainties, which reduces the total uncertainty one needs to deal with. The analysis here is presented in terms of SKU rationalization, but the arguments are highly similar in the case of geographical pooling; we use the term *pooling* to refer to both.

Sometimes, pooling inventories is easy to do and (almost) costless; in such situations, little discussion is needed, and companies should seek and exploit such opportunities. A typical example would be shipping kits with a single set of assembly instructions in multiple languages rather than having to distinguish between kits intended for different markets. However, often, the costs involved are more significant. Some people will argue that reducing product variety, for instance by limiting the number of sizes in which toothpaste is sold, leads to marketing disadvantage. In other cases, rationalizing product lines may directly increase unit manufacturing costs, if multipurpose and more costly components or process changes are required. A well-known example of this is the HP DeskJet printer (Lee, Billington, and Carter 1993): the redesigned product could be localized at distribution centers across the world, but final assembly of the power supply had to be made easier because it would no longer be performed at HP's central facility in Vancouver, leading to higher component and assembly costs. However, these added costs were easily outweighed by the inventory and customer service advantages HP gained. Lee and Tang (1997) developed models to assist in this trade-off between delayed differentiation and increased manufacturing costs. In such cases, more careful evaluation of the benefits of pooling and of possible alternative strategies is needed.

Typically, a company's interest in the various types of pooling that are available is sparked by excessive inventory costs and/or constant customer service problems. Although pooling should benefit such a company in most cases, these may merely be symptoms of poor inventory management to begin with. Indeed, inventory policies in practice often are suboptimal, sometimes dramatically so. Should such a company even consider pooling, or should they get their inventory management sorted out first? In this study, we address this dilemma by asking and answering the following questions:

1. How does the value of pooling behave under a suboptimal inventory policy?

2. Should a company suffering high inventory costs start by improving its inventory policy or by implementing pooling? Ideally, one would do both, but given that practitioners are often hard-pressed to find enough time to even implement one of the two, it is important to know where to start.

3. How do correlation and concentration of uncertainty affect the value of pooling?

4. Do the answers to the foregoing questions vary if demand is not normally distributed?

These questions are particularly relevant because the literature has focused on the value of the pooling effect when an optimal inventory policy is implemented and when demand is normally distributed. This study analyzes the pooling effect from a more practical perspective through its focus on suboptimal inventory policies and its use of nonnormal demand patterns. Much of the literature related to the pooling effect assumes either identical or independent demand distributions (or both). We show that, regardless of the underlying demand structure, there is always a range of suboptimal inventory policies within which pooling is better than improving inventory policy. For the multivariate normal case, we show how correlation but also the level of concentration of uncertainty are key drivers of the value of pooling. To answer the fourth question, we perform simulations based on the Johnson system (Johnson 1987) of distributions (multivariate normal, lognormal, logitnormal, and $\sinh^{-1}$-normal), which allow for a wide variety of demand patterns. We find that the value of pooling is relatively robust across different distributions when an optimal inventory policy is in place,

but that, in some cases, pooling actually may increase costs if an equally suboptimal policy is used before pooling (BP) and after pooling (AP). We also use empirical data that follow no recognizable distribution to further study the drivers of the value of pooling.

Throughout, we focus explicitly on the inventory issues; benefits of pooling due to factors such as manufacturing simplification can be (as) significant, but we do not consider those here. Section 2 reviews literature on various types of pooling. Section 3 presents the framework and basic model for analysis and compares the value of pooling with the value of improving the inventory policy for general distributions; Section 4 does the same for multivariate normal demand. Section 5 reports on our simulation study, and Section 6 analyzes the value of pooling using highly erratic but characteristic demand data from a chemical manufacturer. Section 7 contains our conclusions.

2. Literature Review

The number of SKUs is the standard measure of product variety in a company. Each of these SKUs may be stored in multiple locations, leading to geographical variety. Reducing either product variety or geographical variety generally will allow a company to maintain current customer service at lower cost (or improve service without incurring extra costs). The fundamental mechanism underlying both approaches is equivalent, even if there are significant differences between them. In both cases, pooling can be defined as a business strategy that consists of aggregating individual demands (Gerchak and Mossman 1992). The concept of pooling originated with Eppen's (1979) and Eppen and Schrage's (1981) work on multiechelon inventory systems with geographically dispersed stocking locations. A typical scenario includes a central depot that supplies N locations (or retailers or warehouses) where exogenous, random demands for a single commodity must be filled. The assumptions made there and in much subsequent work include that demand at each location i in period t is normally distributed, that all locations have identical linear holding and penalty costs, that the coefficient of variation is low enough that the probability of negative demand can be ignored, that items are allocated to locations so as to equalize probability of stockout, that unsatisfied demand is backlogged, and that material is nonperishable. Eppen and Schrage and many others also assume that all demands are independent. Eppen (1979) was the first to show how centralization of inventory could reduce expected costs in a multilocation single-period newsboy problem. He called this "statistical economies of scale," and concluded that the expected holding and penalty costs in a decentralized system exceed those in a centralized system. If demands are identical and uncorrelated, costs increase with the square root of the number of locations; as demand correlation increases, this effect is reduced. Schwarz (1989) defines this as the "risk pooling incentive" for centralizing inventories.

In Eppen and Schrage (1981), the depot orders from the supplier and allocates products to the warehouse each period or every m periods. They show that the reduction of total inventory from pooling is greater as the depot's review period increases and is proportional to the number of products. Erkip, Hausman, and Nahmias (1990) find that high positive correlation among products and among successive time periods (around 0.7) results in significantly higher safety stock than the no-correlation case. Schwarz (1989) focuses on lead times and defines the price of risk pooling as the cost of the pipeline inventory caused by the internal lead time. Jönsson and Silver (1987) present an exhaustive study on the impact of changing input parameters on system performance; the redistribution system is more advantageous in situations with high demand variability, a long planning horizon, many locations, and short lead times. Gerchak and Mossman (1992) show how the order quantity and associated costs depend on the randomness parameter in a simple and highly interpretable manner. Federgruen and Zipkin (1984) extend Eppen and Schrage's (1981) model in three important ways: finite horizon, other-than-normal demand distributions (including exponen-

tial and gamma), and nonidentical retailers. Corbett and Rajaram (2001) generalize Eppen's model to (almost) arbitrary multivariate dependent demand distributions.

The benefits of delayed product differentiation are quite similar to those of the pooling effect in multiechelon inventory systems. In fact, most work on postponement has drawn on this body of research (Garg and Lee 1999), referring to multiple products instead of multiple locations. Early research studied standardization of components. Collier (1982) and Baker, Magazine, and Nuttle (1986) designed models to minimize aggregate safety stock levels by using component standardization, subject to a service level constraint. More recently, Groenevelt and Rudi (2000) and Rudi (2000) have examined the interactions between the optimal inventory policy, the degree of component commonality, demand variability, and correlation under bivariate distributions. Cattani (2000) showed that the benefits of risk pooling may be sufficient to offset the higher costs caused by selling a universal product. Recent research has also focused on various types of postponement, as in Lee and Tang (1997; 1998) and Kapuscinski and Tayur (1999).

From a practical perspective, this literature suggests several questions. First, if a company's inventory policy is not optimal to begin with, how does that affect the value of pooling? Second, how well does pooling work with nonnormal demand? These are precisely the questions we answer in this study.

3. Framework and Basic Model

The literature on the various manifestations of the pooling effect makes the assumption that the underlying inventory policy is optimal, i.e., that it minimizes total cost. However, in practice, an "optimal" inventory policy can be very hard to determine, partly because the exact costs are very difficult to estimate and demand distributions are not known. A practitioner, faced with high variety and a suboptimal inventory policy, may well ask how useful pooling will be in such a context. Should the practitioner first improve his inventory policy or should he first implement pooling? Of course, doing both is always the most desirable, but it is not clear which of the two approaches is the best place to start.

Assume that the company initially produces N SKUs; pooling means consolidating all products into one, with demand equal to the sum of the demands of all the original SKUs. We consider a single planning period. Initially, there are N SKUs, with demands z_i that follow a distribution $F_i(z_i)$. Given initial inventory y_i for product i, total expected cost (TC) is given by

$$\text{TC}(y_i) = \int_0^{y_i} h(y_i - z_i)\text{d}F_i(z_i) + \int_{y_i}^{\infty} p(z_i - y_i)\text{d}F_i(z_i). \tag{1}$$

We analyze TC before pooling and after pooling, and with a suboptimal and an optimal inventory policy. Total expected cost before pooling is minimized at $y_i = y_i^*$. The optimal inventory level can be found from $\alpha_i^* = F(y_i^*) = p/(p + h)$, where α_i^* is the optimal service level; for normal distributions, this also can be expressed as $y_i^* = \mu_i + k_i^*\sigma_i$, where k_i^* is the optimal safety factor of product i.

We asked whether, starting with no pooling and a suboptimal inventory policy, it was better to implement pooling or to improve inventory policy. To evaluate suboptimal policies, let the current inventory level for product i be y_i, while the optimal level would be y_i^*. Define the "suboptimality factor" γ_i by $\gamma_i = (y_i - y_i^*)/y_i^*$; γ_i indicates how far the optimal inventory level deviates from the current level, so y_i can be written as $y_i = (1 + \gamma)y_i^*$. For instance, if $\gamma_i = -2/3$, current inventory y_i is 2/3 below the optimal inventory y_i^*. Similarly, if $\gamma_i = 1/2$, current inventory y_i is 50% higher than y_i^*. Assume that $\gamma_i = \gamma$ for all products. From this point on, we will often express total costs as a function of inventory level y_i or of suboptimality factor γ interchangeably. The initial costs are given by $\text{TC}_{\text{BP}}(\gamma)$, the two alternatives (pooling and optimizing inventory policy) carry costs of $\text{TC}_{\text{AP}}(\gamma)$ and TC_{BP}^*

$= \mathrm{TC}_{\mathrm{BP}}(0)$. The following key proposition characterizes the trade-off between pooling and optimizing inventory policy.

PROPOSITION 1. *There is an interval* $[\gamma^L, \gamma^U]$, *which contains 0, such that pooling under the current suboptimal inventory policy is better than improving inventory policy without pooling if and only if* $\gamma \in [\gamma^L, \gamma^U]$.

Proof. We need to compare $\mathrm{TC}_{\mathrm{AP}}(\gamma)$ with $\mathrm{TC}_{\mathrm{BP}}(0)$. For $\gamma = 0$ and any realization of demand, the total costs after pooling cannot exceed total costs before pooling under optimal policies, because one could always choose the optimal before pooling inventory level and treat all products as separate after pooling. Therefore, also in expectation, $\mathrm{TC}_{\mathrm{AP}}(0) \le \mathrm{TC}_{\mathrm{BP}}(0)$. Because $\mathrm{TC}_{\mathrm{AP}}(\gamma)$ is convex in γ and minimized at $\gamma = 0$, the equation $\mathrm{TC}_{\mathrm{BP}}(0) = \mathrm{TC}_{\mathrm{AP}}(\gamma)$ must have two (possibly equal) solutions $\gamma^L \le 0 \le \gamma^U$, which completes the proof. □

Figure 1 shows total costs before pooling and after pooling as a function of γ and illustrates the "pooling range" $[\gamma^L, \gamma^U]$. As long as current inventory y_i lies between $(1 + \gamma^L)y_i^* \le y_i^*$ and $(1 + \gamma^U)y_i^* \ge y_i^*$, pooling is preferred over optimizing inventory policy. In general, this interval $[\gamma^L, \gamma^U]$ expands as the number of SKUs before pooling increases. To formalize this, choose any distribution $G(z)$ for aggregate demand Z. Consider any sequence $\{P_N\}_{N=1}^{\infty}$ of portfolios of N products, each with demand $\{Z_i^{(N)}\}_{i=1}^{N}$, jointly making up aggregate demand so that $Z =_d \sum_{i=1}^{N} Z_i^{(N)}$, where the equality is in distribution. Assume that $\mathrm{TC}_{\mathrm{BP}}(P_{N+1}, \gamma)|_{\gamma=0} \ge \mathrm{TC}_{\mathrm{BP}}(P_N, \gamma)|_{\gamma=0}$ for all N, i.e., there are no two portfolios in the series such that the portfolio with the larger number of SKUs has lower expected costs under optimal policies. A simple example of such a sequence of portfolios is that in which the N products are i.i.d. with $Z_i^{(N)} \sim N(\mu/N, \sigma^2/N)$; clearly, aggregate demand $Z \sim N(\mu, \sigma^2)$ for any portfolio P_N in this sequence, and $\mathrm{TC}_{\mathrm{BP}} = \sqrt{N}\sigma[hk + (p + h)I_n(k)]$ is increasing in N. This expression for $\mathrm{TC}_{\mathrm{BP}}$ is discussed in more detail in the next section.
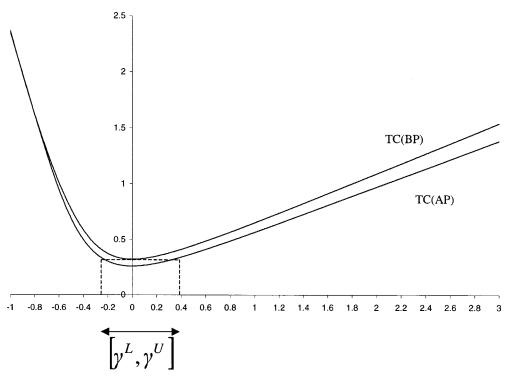

FIGURE 1. Total costs before pooling and after pooling under $\sinh^{-1}$-normal demand, with $\rho = -0.7$, $p = 0.9$, $h = 0.1$, and $\sigma_1 = \sigma_2 = 0.2$.

PROPOSITION 2. *Given any sequence of portfolios $\{P_N\}_{N=1}^{\infty}$ defined previously, the interval $[\gamma^L, \gamma^U]$ expands in N, i.e., $\gamma^L(N)$ is decreasing in N and $\gamma^U(N)$ is increasing in N.*

Proof. To examine $\gamma^U(N)$, define $h(N, \gamma) \equiv \text{TC}_{\text{BP}}(N, 0) - \text{TC}_{\text{AP}}(N, \gamma)$ on $N \times \mathbb{R}_0^+$ and define $\gamma^U(N)$ through $h(N, \gamma^U(N)) \equiv 0$. Existence of $\gamma^U(N) \geq 0$ follows from Proposition 1. Clearly, $h(N, 0) \geq 0$ for all $N \in \mathbb{N}$. We know that $\text{TC}_{\text{BP}}(N, 0)$ is increasing in N, by the assumption on $\{P_N\}_{N=1}^{\infty}$. $\text{TC}_{\text{AP}}(N, \gamma)$ does not depend on N, and is increasing in γ for $\gamma \geq 0$. Therefore, $h(N, \gamma)$ is increasing in N for all γ and is decreasing in γ for all N. The definition of $\gamma^U(N)$ then implies that $\gamma^U(N)$ must be increasing in N. The result for $\gamma^L(N)$ follows analogously. $\square$

4. Pooling or Improving Inventory Policy: The Multivariate Normal Case

We have shown, for general distributions, that there is a unique interval $[\gamma^L, \gamma^U]$ in which implementing pooling is better than improving inventory policy. Here, we focus on the multivariate normal case. First, we theoretically analyze the effects of correlation and of concentration of uncertainty on the pooling effect. Then, we study in depth how the interval $[\gamma^L, \gamma^U]$ behaves under normal demand; in the next section we use simulation to examine the behavior of $[\gamma^L, \gamma^U]$ under nonnormal distributions.

Write $\mu = 1/N \sum_{i=1}^{N} \mu_i$ and $\sigma = 1/N \sum_{i=1}^{N} \sigma_i$, the "average" mean and standard deviation, and define a vector $\mathbf{x}$ of x_i such that $\sigma_i = x_i\sigma$ for all i and $\sum_{i=1}^{N} x_i = N$. If $x_i = 1$ for all i, then all SKUs contribute equally to total uncertainty. Write $\sigma_x^2 = 1/N \sum_{i=1}^{N} (x_i - 1)^2$ for the variance of the x_i, a measure of how total uncertainty is dispersed among the N SKUs. The higher σ_x^2, the more uncertainty is concentrated among fewer SKUs. Covariances are given by $\sigma_{ij} = \sigma_i\sigma_j\rho_{ij} = x_ix_j\rho_{ij}\sigma^2$, so that $\mathbf{\Sigma} = \sigma^2\mathbf{x}'\mathbf{R}\mathbf{x}$ with $\mathbf{R}$ the correlation matrix. Assume the coefficients of variation of demand σ_i/μ_i are small so that the probability of negative demand can be ignored; for simplicity, we assume that all SKUs have the same penalty and holding costs. Total costs for product i in the multivariate normal distribution are then

$$\text{TC}(y_i) = h(y_i - \mu_i) + (p + h) \int_y^{\infty} (z_i - y_i)\mathrm{d}F_i(z_i) = (hk_i + (p + h)I_n(k_i))\sigma_i \qquad (2)$$

where $k_i = (y_i - \mu_i)/\sigma_i$ and where I_n is the unit normal loss function (Eppen and Schrage 1981):

$$I_n(k) = \int_k^{\infty} (z - k) \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\mathrm{d}z = \frac{1}{\sqrt{2\pi}} e^{-k^2/2} - k[1 - \Phi(k)] \qquad (3)$$

The term $I_n(k)$ is convex decreasing in k, and $\lim_{k \downarrow -\infty} I_n(k) = \infty$ and $\lim_{k \uparrow +\infty} I_n(k) = 0$. At optimum, total expected costs for product i are

$$\text{TC}_i^* = [hk_i^* + (p + h)I_n(k_i^*)]\sigma_i = [hk_i^* + (p + h)I_n(k_i^*)]x_i\sigma. \qquad (4)$$

The optimal safety factor is the same for each product: $k_i^* = k^*$. Lemma 1 is well known.

LEMMA 1. *After pooling N SKUs into 1, demand for the consolidated product is $N(\mu_{AP}, \sigma_{AP}^2)$, where $\mu_{AP} = N\mu$ and, letting $\boldsymbol{\iota}$ be the N-dimensional vector with all elements equal to one,*

$$\sigma_{AP} = \sqrt{\sum_{i=1}^{N} \sigma_i^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sigma_i\sigma_j\rho_{ij}} = \sqrt{\boldsymbol{\iota}'\mathbf{\Sigma}\boldsymbol{\iota}} = \sigma\sqrt{\mathbf{x}'\mathbf{R}\mathbf{x}}. \qquad (5)$$

Two special cases help us interpret this expression:
- When $\rho_{ij} = 0$ for all nondiagonal elements, so that $\mathbf{R} = \mathbf{I}$ (the N-dimensional unit matrix), then $\sigma_{AP} = \sigma\sqrt{\mathbf{x}'\mathbf{x}} = \sigma\sqrt{N(\sigma_x^2 + 1)}$, where $\sigma_x^2 = 1/N \sum_{i=1}^{N} x_i^2$

$- (1/N \sum_{i=1}^{N} x_i)^2 = 1/N \sum_{i=1}^{N} x_i^2 - 1$ because $\sum_{i=1}^{N} x_i = N$. Clearly, σ_{AP} increases in σ_x^2, the concentration of uncertainty; it is minimized at $\mathbf{x} = \boldsymbol{\iota}$, in which case $\sigma_x^2 = 0$ and $\sigma_{AP} = \sigma\sqrt{N}$.

- If all SKUs are identically distributed, i.e., $\mathbf{x} = \boldsymbol{\iota}$, then $\sigma_{AP} = \sigma\sqrt{N + \sum_{i \neq j} \rho_{ij}}$. Now, σ_{AP} is increasing in all ρ_{ij} as expected, and $\mathbf{R} = \mathbf{I}$ gives $\sigma_{AP} = \sigma\sqrt{N}$.

In general, $0 \leq \sigma_{AP} \leq N\sigma$. Total costs before pooling and after pooling are also well known.

LEMMA 2. *For any safety factor k, total inventory level and expected cost for the system with N SKUs before pooling are given by*

$$y_{BP} = N\mu + Nk\sigma \tag{6}$$

$$TC_{BP} = N\sigma[hk + (p + h)I_n(k)]. \tag{7}$$

LEMMA 3. *For any safety factor k, total inventory level and total expected cost for the system after pooling N SKUs into 1 are given by*

$$y_{AP} = N\mu + k\sigma\sqrt{\mathbf{x'Rx}} \tag{8}$$

$$TC_{AP} = \sigma\sqrt{\mathbf{x'Rx}}\,[hk + (p + h)I_n(k)]. \tag{9}$$

Note that defining suboptimality in terms of $(k - k^*)$ is not equivalent to our definition of γ. For normal demand, $(k - k^*) = \gamma(\mu/\sigma + 1)$, so the conversion factor between $(k - k^*)$ and γ depends on σ and hence is different before pooling and after pooling. Later, we will see that examining γ highlights some counterintuitive effects that are hidden when using $(k - k^*)$. [The same is true using $(k - k^*)/k^*$ as measure of suboptimality.] We can now quantify the value of pooling under an optimal or suboptimal inventory policy, and the value of improving inventory policy before pooling and after pooling. Actually arriving at the optimal policy is hard to do in practice, so using the optimal policy as a benchmark gives an upper bound on the benefits of improving inventory policy.

PROPOSITION 3. *The effects of optimizing inventory policy before pooling and after pooling are*

$$|y_{BP} - y_{BP}^*| = N\sigma|k - k^*| \geq \sigma\sqrt{\mathbf{x'Rx}}\,|k - k^*| = |y_{AP} - y_{AP}^*| \tag{10}$$

$$TC_{BP} - TC_{BP}^* = N\sigma[h(k - k^*) + (p + h)[I_n(k) - I_n(k^*)]]$$

$$\geq \sigma\sqrt{\mathbf{x'Rx}}\,[h(k - k^*) + (p + h)[I_n(k) - I_n(k^*)]] = TC_{AP} - TC_{AP}^*. \tag{11}$$

Proposition 3 shows that the value of improving inventory policy depends on $(k - k^*)$, the extent to which the safety factor k is above or below the optimal k^*, and on $I_n(k) - I_n(k^*)$, where k^* itself depends on $p/(p + h)$. Because of the convexity of TC, we know that the cost of suboptimality is convex increasing in $|k - k^*|$. The value of improving inventory policy before pooling is also increasing in N and σ, as expected. Finally, by the inequality in (10), the absolute deviation from the optimal inventory level is greater before pooling than after pooling.

PROPOSITION 4. *The effects of pooling under a suboptimal inventory policy are*

$$y_{BP} - y_{AP} = (N - \sqrt{\mathbf{x'Rx}})k\sigma \tag{12}$$

$$TC_{BP} - TC_{AP} = (N - \sqrt{\mathbf{x'Rx}})\sigma[hk + (p + h)I_n(k)] \geq TC_{BP}^* - TC_{AP}^* \geq 0. \tag{13}$$

For the effects of pooling under an optimal policy, substitute k* for k in (12) and (13).

Proposition 3 shows how the value of improving a suboptimal inventory policy depends on the suboptimality $(k - k^*)$, and Proposition 4 shows that the effect of pooling is greater under suboptimal policies. To further interpret Propositions 3 and 4, let us return to the two earlier special cases, which confirm what is known (intuitively or explicitly) about the effects of pooling. If $\mathbf{R} = \mathbf{I}$, then the cost reduction depends on $(N - \sqrt{\mathbf{x'x}})\,\sigma = (N - \sqrt{N(\sigma_x^2 + 1)})\sigma$, which is decreasing in σ_x^2 and maximized at $\sigma_x^2 = 0$, or $\mathbf{x} = \boldsymbol{\iota}$. The more evenly uncertainty is spread out among the SKUs, the greater the benefits of pooling. Similarly, if $\mathbf{x} = \boldsymbol{\iota}$, the cost reduction depends on $(N - \sqrt{N + \sum_{i \neq j} \rho_{ij}})\sigma$, which is

decreasing in ρ_{ij}. The more positive correlation among demands, the lower the value of pooling. One can restate Proposition 1 for the multivariate normal case and show that there always exists a pooling range $[k^L, k^U]$, which contains the optimal safety factor k^* such that pooling is preferred over optimizing inventory policy if and only if the current policy k lies within $[k^L, k^U]$. For the multivariate normal case, one can go further: (13) implies that the value of pooling is convex increasing in $|k - k^*|$, i.e., the value of pooling is greater as the current inventory policy is more suboptimal, when suboptimality is expressed in terms of deviation from the optimal safety factor. Our numerical results in Section 5 show that, with our more general definition of suboptimality, we cannot say that the value of pooling increases with suboptimality, even for normal demand.

5. Comparing Pooling and Improving Inventory Policy for Normal and Nonnormal Multivariate Distributions: Numerical Experiments

In this section, we report on an extensive Monte Carlo simulation study to examine how the trade-off between pooling and improving inventory policy depends on the distribution, the critical fractile $\alpha = p/(p + h)$, the correlation structure, and the concentration of uncertainty. We generate bivariate normal random vectors, and then use Johnson's (1987) translation system to transform them to any of the three related distributions (lognormal, logitnormal, and $\sinh^{-1}$-normal). Let $\mathbf{X}$ have a normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The lognormal is then obtained by applying the transformation $Z_i = \lambda_i e^{x_i} + \xi_i$ for some parameters λ_i and ξ_i; the $\sinh^{-1}$-normal is obtained from $Z_i = \lambda_i \sinh (X_i) + \xi_i$; and the logitnormal is obtained from $Z_i = \lambda_i[1 + e^{X_i}]^{-1} + \xi_i$. See Johnson (1987, Chapter 5) for an extensive discussion of these distributions and their characteristics; they allow a wide range of behaviors, including skewed and multimodal distributions, which are fundamentally different from the normal distribution. For instance, the logitnormal distribution with different variances can lead to multimodal densities. Generating correlated multivariate random vectors with other distributions is a major challenge in simulation (Johnson 1987, pp. 3–4).

Following most of Johnson (1987) we focused on the case $N = 2$, as computation time increased drastically with $N > 2$; in the next section we examine the trade-off between pooling and improving inventory policy using empirical demand data with $N \geqslant 2$. For all four distributions, we set $\mu_1 = \mu_2 = 1$, and used the following scenarios for the other parameters. We used three values of concentration of uncertainty: $\sigma_1 = \sigma_2 = 0.2$ so that $\sigma_x^2 = 0$, $\sigma_1 = 0.1$ and $\sigma_2 = 0.3$ so that $\sigma_x^2 = 0.25$, and $\sigma_1 = 0.05$ and $\sigma_2 = 0.35$ so that $\sigma_x^2 = 0.56$. We used seven values for the correlation coefficient: $\rho_{12} \in \{-0.7, -0.5, -0.3, 0, 0.3, 0.5, 0.7\}$. As in Johnson, we specify the covariance matrix before applying the transformations; Johnson (1987) gives the expressions for computing the corresponding covariance structure after transformation. We used three values for $p/(p + h)$: $p = 0.9$ and $h = 0.1$ so that $p/(p + h) = 0.9$, $p = 0.5$ and $h = 0.5$ so that $p/(p + h) = 0.5$, and $p = 0.1$ and $h = 0.9$ so that $p/(p + h) = 0.1$. This corresponds to an experimental design with $4 \times 3 \times 7 \times 3 = 252$ scenarios. Additional simulations were performed but not reported here because they do not provide further insights. The μ_i and σ_i values make the coefficient of variation low enough to avoid negative demands.

The simulation study was based on a Monte Carlo approach; we generated random demand vectors, determined the costs under each demand realization, and then numerically integrated to calculate expected costs for a given inventory policy. We used MatLab6.0, and every experiment was performed with 120,000 random observations (each is a two-dimensional demand vector) and 2,000 intervals (for the numerical integration). This led to sufficient accuracy for our purposes; some experiments were performed with more observations and intervals, but no relevant differences existed and computational time drastically increased. We used the same set of 120,000 normal random vectors for all experiments to allow more

TABLE 1

Pooling Ranges $[\gamma^L, \gamma^U]$ *for* $p = 0.9$ *and* $h = 0.1$

		Normal		Lognormal		Sinh^{-1}-Normal		Logitnormal	
		γ^L	γ^U	γ^L	γ^U	γ^L	γ^U	γ^L	γ^U
$\rho_{12} = -0.7$	$\sigma_x^2 = 0$	−19.9%	45.9%	−25.4%	59.5%	−29.7%	70.7%	−5.5%	11.6%
	$\sigma_x^2 = 0.25$	−19.6%	40.0%	−24.9%	54.0%	−29.6%	62.8%	−5.2%	10.4%
	$\sigma_x^2 = 0.56$	−18.2%	32.0%	−24.7%	43.2%	−28.2%	50.4%	−4.8%	8.6%
$\rho_{12} = -0.5$	$\sigma_x^2 = 0$	−20.0%	38.8%	−25.8%	52.2%	−29.5%	60.6%	−5.5%	9.9%
	$\sigma_x^2 = 0.25$	−19.2%	34.6%	−25.0%	47.8%	−28.9%	55.5%	−5.3%	9.0%
	$\sigma_x^2 = 0.56$	−17.5%	28.6%	−23.8%	38.6%	−26.6%	45.6%	−4.8%	7.6%
$\rho_{12} = -0.3$	$\sigma_x^2 = 0$	−19.2%	33.2%	−25.2%	47.0%	−29.2%	53.6%	−5.5%	8.6%
	$\sigma_x^2 = 0.25$	−18.3%	30.1%	−24.4%	42.0%	−28.1%	48.9%	−5.0%	7.9%
	$\sigma_x^2 = 0.56$	−16.9%	25.3%	−22.3%	35.5%	−25.9%	40.3%	−4.5%	6.7%
$\rho_{12} = 0$	$\sigma_x^2 = 0$	−17.7%	27.0%	−23.6%	38.7%	−27.1%	43.6%	−4.9%	6.9%
	$\sigma_x^2 = 0.25$	−16.7%	24.3%	−22.6%	35.5%	−25.8%	40.1%	−4.6%	6.6%
	$\sigma_x^2 = 0.56$	−15.0%	20.9%	−20.3%	30.1%	−22.8%	34.0%	−4.1%	5.5%
$\rho_{12} = 0.3$	$\sigma_x^2 = 0$	−15.2%	20.8%	−21.3%	31.0%	−24.3%	33.8%	−4.2%	5.5%
	$\sigma_x^2 = 0.25$	−14.1%	19.0%	−20.1%	28.3%	−22.6%	31.5%	−3.9%	5.1%
	$\sigma_x^2 = 0.56$	−12.8%	16.7%	−17.4%	24.4%	−20.2%	26.9%	−3.5%	4.5%
$\rho_{12} = 0.5$	$\sigma_x^2 = 0$	−12.8%	16.6%	−18.9%	26.2%	−20.9%	28.5%	−3.6%	4.3%
	$\sigma_x^2 = 0.25$	−12.3%	15.4%	−18.0%	22.8%	−19.8%	26.1%	−3.4%	4.1%
	$\sigma_x^2 = 0.56$	−11.1%	13.6%	−15.6%	19.9%	−17.4%	21.8%	−3.0%	3.6%
$\rho_{12} = 0.7$	$\sigma_x^2 = 0$	−10.3%	12.3%	−15.3%	19.0%	−16.9%	21.2%	−2.9%	3.2%
	$\sigma_x^2 = 0.25$	−9.6%	11.5%	−14.0%	17.7%	−16.1%	19.2%	−2.7%	3.1%
	$\sigma_x^2 = 0.56$	−8.5%	10.1%	−12.5%	15.3%	−13.6%	16.2%	−2.4%	2.7%

accurate comparisons. We repeated the experiments with several other sets of random numbers but the results were very similar.

Tables 1–6 summarize the results obtained from the simulation. Tables 1–3 show the $[\gamma^L,$

TABLE 2

Pooling Ranges $[\gamma^L, \gamma^U]$ *for* $p = 0.5$ *and* $h = 0.5$

		Normal		Lognormal		Sinh^{-1}-Normal		Logitnormal	
		γ^L	γ^U	γ^L	γ^U	γ^L	γ^U	γ^L	γ^U
$\rho_{12} = -0.7$	$\sigma_x^2 = 0$	−35.4%	35.5%	−32.3%	36.0%	−43.3%	46.9%	−9.6%	9.4%
	$\sigma_x^2 = 0.25$	−33.7%	34.0%	−29.8%	35.4%	−41.2%	46.0%	−9.3%	8.6%
	$\sigma_x^2 = 0.56$	−30.7%	30.7%	−26.5%	32.8%	−36.8%	42.3%	−8.6%	7.7%
$\rho_{12} = -0.5$	$\sigma_x^2 = 0$	−34.6%	34.6%	−31.9%	34.9%	−42.1%	45.8%	−9.4%	9.1%
	$\sigma_x^2 = 0.25$	−32.7%	32.2%	−28.9%	33.9%	−39.7%	43.6%	−8.9%	8.5%
	$\sigma_x^2 = 0.56$	−29.3%	29.3%	−26.1%	30.4%	−35.7%	40.8%	−8.1%	7.5%
$\rho_{12} = -0.3$	$\sigma_x^2 = 0$	−33.1%	32.8%	−30.2%	33.4%	−40.9%	43.7%	−9.0%	8.7%
	$\sigma_x^2 = 0.25$	−31.1%	30.9%	−28.1%	31.9%	−38.1%	42.2%	−8.5%	8.1%
	$\sigma_x^2 = 0.56$	−27.6%	27.6%	−25.5%	29.3%	−33.8%	36.8%	−7.6%	7.1%
$\rho_{12} = 0$	$\sigma_x^2 = 0$	−29.5%	29.5%	−27.9%	30.2%	−37.5%	38.7%	−8.1%	7.9%
	$\sigma_x^2 = 0.25$	−27.7%	27.8%	−25.8%	29.0%	−34.5%	37.3%	−7.7%	7.3%
	$\sigma_x^2 = 0.56$	−24.3%	24.3%	−23.2%	25.2%	−30.2%	32.8%	−6.8%	6.3%
$\rho_{12} = 0.3$	$\sigma_x^2 = 0$	−25.6%	25.9%	−23.5%	26.0%	−31.6%	34.4%	−7.0%	6.7%
	$\sigma_x^2 = 0.25$	−23.6%	23.9%	−22.2%	23.9%	−30.1%	31.2%	−6.5%	6.2%
	$\sigma_x^2 = 0.56$	−20.9%	20.7%	−20.1%	21.4%	−25.6%	28.8%	−5.7%	5.5%
$\rho_{12} = 0.5$	$\sigma_x^2 = 0$	−21.9%	21.5%	−20.8%	22.1%	−27.8%	28.7%	−6.0%	5.8%
	$\sigma_x^2 = 0.25$	−20.5%	20.0%	−19.6%	20.6%	−26.1%	27.0%	−5.5%	5.4%
	$\sigma_x^2 = 0.56$	−17.7%	18.2%	−17.5%	18.7%	−22.6%	24.3%	−4.8%	4.7%
$\rho_{12} = 0.7$	$\sigma_x^2 = 0$	−17.1%	17.1%	−16.9%	16.9%	−21.8%	23.3%	−4.6%	4.6%
	$\sigma_x^2 = 0.25$	−15.8%	15.6%	−15.1%	16.0%	−20.2%	21.6%	−4.3%	4.2%
	$\sigma_x^2 = 0.56$	−13.8%	13.8%	−13.2%	14.1%	−17.8%	19.0%	−3.8%	3.7%

TABLE 3
Pooling Ranges $[\gamma^L, \gamma^U]$ for $p = 0.1$ and $h = 0.9$

		Normal		Lognormal		Sinh^{-1}-Normal		Logitnormal	
		γ^L	γ^U	γ^L	γ^U	γ^L	γ^U	γ^L	γ^U
$\rho_{12} = -0.7$	$\sigma_x^2 = 0$	−72.0%	31.8%	−44.4%	24.0%	−76.7%	37.0%	−18.2%	7.8%
	$\sigma_x^2 = 0.25$	−67.5%	32.3%	−42.6%	22.0%	−69.7%	34.2%	−16.7%	8.1%
	$\sigma_x^2 = 0.56$	−59.7%	33.5%	−37.0%	19.8%	−64.4%	37.3%	−13.9%	8.1%
$\rho_{12} = -0.5$	$\sigma_x^2 = 0$	−70.6%	36.8%	−41.7%	25.0%	−72.6%	41.1%	−16.7%	8.5%
	$\sigma_x^2 = 0.25$	−64.9%	36.3%	−38.6%	22.4%	−67.3%	37.7%	−15.3%	8.4%
	$\sigma_x^2 = 0.56$	−57.0%	35.6%	−33.9%	20.5%	−61.5%	35.3%	−12.9%	8.0%
$\rho_{12} = -0.3$	$\sigma_x^2 = 0$	−67.3%	39.4%	−37.5%	25.7%	−70.9%	45.6%	−15.4%	8.9%
	$\sigma_x^2 = 0.25$	−62.7%	37.8%	−35.1%	22.7%	−65.9%	40.8%	−14.1%	8.4%
	$\sigma_x^2 = 0.56$	−53.5%	35.8%	−29.5%	21.2%	−56.3%	36.0%	−11.9%	7.9%
$\rho_{12} = 0$	$\sigma_x^2 = 0$	−62.5%	42.6%	−33.1%	24.2%	−64.7%	47.0%	−13.4%	8.6%
	$\sigma_x^2 = 0.25$	−58.1%	39.2%	−30.7%	22.3%	−62.6%	43.2%	−12.3%	8.3%
	$\sigma_x^2 = 0.56$	−49.5%	35.7%	−27.6%	19.0%	−52.0%	36.2%	−10.5%	7.4%
$\rho_{12} = 0.3$	$\sigma_x^2 = 0$	−55.6%	41.5%	−27.1%	22.3%	−57.7%	44.4%	−11.1%	8.0%
	$\sigma_x^2 = 0.25$	−50.5%	37.7%	−25.5%	20.0%	−52.0%	39.1%	−10.2%	7.5%
	$\sigma_x^2 = 0.56$	−43.2%	32.6%	−21.8%	18.4%	−43.3%	36.4%	−8.6%	6.7%
$\rho_{12} = 0.5$	$\sigma_x^2 = 0$	−48.9%	39.6%	−22.8%	19.7%	−52.9%	43.7%	−9.5%	7.4%
	$\sigma_x^2 = 0.25$	−44.0%	34.3%	−21.9%	17.9%	−47.5%	40.1%	−8.6%	6.8%
	$\sigma_x^2 = 0.56$	−37.7%	30.4%	−18.8%	15.2%	−36.6%	31.4%	−7.3%	5.9%
$\rho_{12} = 0.7$	$\sigma_x^2 = 0$	−40.3%	34.0%	−17.5%	15.7%	−42.4%	37.4%	−7.2%	6.1%
	$\sigma_x^2 = 0.25$	−36.2%	29.8%	−16.0%	13.4%	−37.7%	35.4%	−6.6%	5.5%
	$\sigma_x^2 = 0.56$	−30.3%	25.1%	−14.7%	13.9%	−31.7%	26.9%	−5.6%	4.9%

$\gamma^U]$ intervals for each scenario, and Tables 4–6 show how pooling reduces total costs. In the first row in Table 1, $[\gamma^L, \gamma^U] = [-19.9\% + 45.9\%]$ for the normal distribution. This means that for any inventory level between 81.1% and 145.9% of the optimal inventory level,

TABLE 4
Cost Reduction due to Pooling for $p = 0.9$ and $h = 0.1$ under Optimal Policy

		Normal	Lognormal	Sinh^{-1}-Normal	Logitnormal
$\rho_{12} = -0.7$	$\sigma_x^2 = 0$	61.0%	56.8%	59.0%	56.1%
	$\sigma_x^2 = 0.25$	54.0%	45.6%	47.0%	52.8%
	$\sigma_x^2 = 0.56$	43.3%	32.1%	33.4%	45.5%
$\rho_{12} = -0.5$	$\sigma_x^2 = 0$	50.0%	49.1%	50.6%	45.1%
	$\sigma_x^2 = 0.25$	44.8%	40.2%	41.0%	42.5%
	$\sigma_x^2 = 0.56$	36.5%	28.5%	29.5%	37.4%
$\rho_{12} = -0.3$	$\sigma_x^2 = 0$	40.9%	42.5%	42.3%	36.4%
	$\sigma_x^2 = 0.25$	36.9%	34.8%	35.2%	34.6%
	$\sigma_x^2 = 0.56$	30.6%	25.2%	25.5%	30.5%
$\rho_{12} = 0$	$\sigma_x^2 = 0$	29.3%	32.0%	31.9%	25.5%
	$\sigma_x^2 = 0.25$	26.8%	26.5%	26.8%	24.5%
	$\sigma_x^2 = 0.56$	22.4%	19.3%	19.5%	21.9%
$\rho_{12} = 0.3$	$\sigma_x^2 = 0$	19.5%	22.2%	21.9%	16.8%
	$\sigma_x^2 = 0.25$	17.8%	18.4%	18.5%	15.7%
	$\sigma_x^2 = 0.56$	15.0%	13.4%	13.6%	14.6%
$\rho_{12} = 0.5$	$\sigma_x^2 = 0$	13.4%	15.7%	15.4%	11.4%
	$\sigma_x^2 = 0.25$	12.4%	13.3%	12.8%	11.1%
	$\sigma_x^2 = 0.56$	10.7%	9.6%	9.6%	10.0%
$\rho_{12} = 0.7$	$\sigma_x^2 = 0$	7.8%	9.3%	9.2%	6.7%
	$\sigma_x^2 = 0.25$	7.1%	7.8%	7.8%	6.4%
	$\sigma_x^2 = 0.56$	6.2%	5.8%	5.8%	5.9%

TABLE 5

Cost Reduction due to Pooling for p = 0.5 and h = 0.5 under Optimal Policy

		Normal	Lognormal	$Sinh^{-1}$-Normal	Logitnormal
$\rho_{12} = -0.7$	$\sigma_x^2 = 0$	61.2%	53.7%	56.2%	60.0%
	$\sigma_x^2 = 0.25$	53.9%	48.8%	50.1%	53.8%
	$\sigma_x^2 = 0.56$	43.2%	39.5%	39.8%	43.6%
$\rho_{12} = -0.5$	$\sigma_x^2 = 0$	50.2%	44.4%	46.2%	49.3%
	$\sigma_x^2 = 0.25$	44.7%	40.3%	41.7%	44.7%
	$\sigma_x^2 = 0.56$	36.5%	33.2%	33.7%	37.0%
$\rho_{12} = -0.3$	$\sigma_x^2 = 0$	40.8%	36.7%	38.0%	40.4%
	$\sigma_x^2 = 0.25$	37.1%	33.5%	34.4%	36.9%
	$\sigma_x^2 = 0.56$	30.4%	27.7%	28.2%	30.8%
$\rho_{12} = 0$	$\sigma_x^2 = 0$	29.4%	c26.7%	27.3%	29.2%
	$\sigma_x^2 = 0.25$	26.7%	24.4%	24.9%	26.8%
	$\sigma_x^2 = 0.56$	22.4%	20.3%	20.7%	22.9%
$\rho_{12} = 0.3$	$\sigma_x^2 = 0$	19.5%	17.9%	18.1%	19.4%
	$\sigma_x^2 = 0.25$	17.9%	16.3%	16.5%	18.0%
	$\sigma_x^2 = 0.56$	14.9%	13.7%	13.8%	15.3%
$\rho_{12} = 0.5$	$\sigma_x^2 = 0$	13.4%	12.3%	12.4%	13.4%
	$\sigma_x^2 = 0.25$	12.5%	11.3%	11.4%	12.5%
	$\sigma_x^2 = 0.56$	10.5%	9.5%	9.7%	10.6%
$\rho_{12} = 0.7$	$\sigma_x^2 = 0$	7.7%	7.3%	7.2%	7.9%
	$\sigma_x^2 = 0.25$	7.2%	6.6%	6.7%	7.3%
	$\sigma_x^2 = 0.56$	6.1%	5.6%	5.7%	6.2%

pooling is better than optimizing inventory policy. The corresponding entry in Table 4 shows that $(TC_{BP}(y^*) - TC_{AP}(y^*))/TC_{BP}(y^*) = 61.0\%$, the cost reduction achievable by pooling under an optimal policy. We discuss the effects of suboptimal policies, demand distribution, correlation structure, critical fractile α, and concentration of uncertainty sepa-

TABLE 6

Cost Reduction due to Pooling for p = 0.1 and h = 0.9 under Optimal Policy

		Normal	Lognormal	$Sinh^{-1}$-Normal	Logitnormal
$\rho_{12} = -0.7$	$\sigma_x^2 = 0$	61.5%	48.9%	53.5%	62.6%
	$\sigma_x^2 = 0.25$	54.0%	46.3%	49.9%	53.0%
	$\sigma_x^2 = 0.56$	43.2%	41.2%	42.9%	41.2%
$\rho_{12} = -0.5$	$\sigma_x^2 = 0$	49.9%	38.7%	43.0%	52.0%
	$\sigma_x^2 = 0.25$	44.8%	36.7%	40.4%	45.3%
	$\sigma_x^2 = 0.56$	36.6%	33.4%	35.1%	35.7%
$\rho_{12} = -0.3$	$\sigma_x^2 = 0$	40.9%	30.6%	34.6%	43.2%
	$\sigma_x^2 = 0.25$	37.1%	29.6%	32.8%	38.2%
	$\sigma_x^2 = 0.56$	30.6%	27.2%	28.5%	30.3%
$\rho_{12} = 0$	$\sigma_x^2 = 0$	29.2%	21.5%	24.6%	31.5%
	$\sigma_x^2 = 0.25$	26.9%	20.6%	23.3%	28.1%
	$\sigma_x^2 = 0.56$	22.4%	18.9%	20.2%	22.7%
$\rho_{12} = 0.3$	$\sigma_x^2 = 0$	19.2%	14.0%	16.0%	21.0%
	$\sigma_x^2 = 0.25$	17.8%	13.3%	15.1%	19.0%
	$\sigma_x^2 = 0.56$	15.1%	12.4%	13.5%	15.5%
$\rho_{12} = 0.5$	$\sigma_x^2 = 0$	13.4%	9.5%	10.9%	14.8%
	$\sigma_x^2 = 0.25$	12.3%	9.1%	10.2%	13.2%
	$\sigma_x^2 = 0.56$	10.5%	8.5%	9.2%	10.8%
$\rho_{12} = 0.7$	$\sigma_x^2 = 0$	7.7%	5.4%	6.3%	8.6%
	$\sigma_x^2 = 0.25$	7.3%	5.3%	6.0%	7.8%
	$\sigma_x^2 = 0.56$	6.2%	5.0%	5.4%	6.5%

Case 2a: $\rho = 0, \sigma_1 = \sigma_2 = 0.2, p = 0.1, h = 0.9$

Case 2b: $\rho = 0, \sigma_1 = 0.05, \sigma_2 = 0.35, p = 0.1, h = 0.9$

Case 2c: $\rho = -0.7, \sigma_1 = \sigma_2 = 0.2, p = 0.1, h = 0.9$

Case 2d: $\rho = -0.7, \sigma_1 = \sigma_2 = 0.2, p = 0.9, h = 0.1$

FIGURE 2. Total costs before pooling and after pooling under normal demand.

rately. Figure 2 illustrates the effect of changes in correlation and in concentration of uncertainty on the costs before pooling and after pooling and on the pooling range $[\gamma^L, \gamma^U]$ for four scenarios, to highlight the effect of individual parameters (ρ, σ_x^2, and α).

1. *Suboptimal inventory policy.* Clearly, when $\gamma = 0$, $\mathrm{TC_{BP}}(\gamma) - \mathrm{TC_{AP}}(\gamma) \geq 0$, because after pooling one can always replicate the before pooling outcome for any demand realization. However, Figure 2 shows that as γ deviates from 0, it is possible to have $\mathrm{TC_{BP}}(\gamma)$ $- \mathrm{TC_{AP}}(\gamma) < 0$, i.e., total costs increase after pooling. If one defines suboptimality in the more restricted way, as $(k - k^*)$, for normal distributions, we know from Proposition 4 that $\mathrm{TC_{BP}}(k) - \mathrm{TC_{AP}}(k) \geq 0$ for all k, from which one would infer that pooling always reduces costs. By using the, in our view, more appropriate definition of suboptimality defined by γ, such possibilities are not obscured. Although this may sound counterintuitive, one has to realize that this occurs because we are changing the total inventory level in order to keep the suboptimality of the inventory policy constant before pooling and after pooling. For instance, if $\gamma = -0.2$, indicating that current inventory before pooling is 20% lower than the optimal level, and the optimal total inventory after pooling is lower than before pooling, the corresponding suboptimal inventory level after pooling is 20% below an already lower after pooling optimal level.

2. *Demand distribution*. Perhaps the most intriguing finding is how robust the value of pooling is under nonnormal demand distributions. The percentage cost reductions under an optimal policy do not vary dramatically between distributions. The pooling ranges $[\gamma^L, \gamma^U]$ seem to depend more heavily on the distribution. Particularly, the ranges are much smaller for the logitnormal distribution than for the others: for instance, in Table 2, when $\alpha = 0.5$, when there is no correlation and when variances are equal, pooling is better than improving inventory policy only if the current inventory level lies within approximately 8% of the optimal inventory policy. For the other distributions, the range stretches between 27.9% and 38.7% on either side of the optimal policy. To summarize, in determining whether to focus on pooling or on improving inventory policy, the demand distribution does have an impact, while for determining the cost reduction, the distribution has relatively little impact.

3. *Correlation*. As expected, higher correlation leads to lower cost reductions. What is remarkable again is the similarity across distributions. In Table 5, taking $p = 0.5$ and $h = 0.5$ and equal variances, the value of pooling varies from 7.2% to 7.9% at $\rho_{12} = 0.7$ and from 53.7% to 61.2% at $\rho_{12} = -0.7$. Consequently, the pooling ranges also shrink as correlation increases. For instance, in Table 2, with $p = 0.5$ and $h = 0.5$ and equal variances under a normal distribution, the pooling range increases from $\pm 17.1\%$ at $\rho_{12} = 0.7$ to $\pm 35.5\%$ at $\rho_{12} = -0.7$. Simulations with higher (absolute) values of ρ_{12} were consistent with this but are not reported because the accuracy of the simulation results deteriorates as ρ_{12} gets closer to ± 1. Comparing Figures 2a and c, we observe a similar effect.

4. *Critical fractile α*. The higher α is, the more the pooling range is skewed "to the right," and conversely for low values of α. In other words, when α is high, pooling is preferred over improving inventory policy when current inventory is not too much below optimal but it can be significantly above optimal. Compare, for instance, the pooling range for the normal under equal variances and $\rho_{12} = -0.7$; in Table 1, with $\alpha = 0.9$, pooling is preferable whenever current inventory lies within $[-19.9\%, +45.9\%]$ from optimal, while in Table 3 with $\alpha = 0.1$, it can lie within $[-72.0\%, +31.8\%]$ from optimal. This behavior is intuitive: when α is high, costs increase sharply as suboptimality factor γ drops below zero (i.e., as inventory drops below optimal), and increase slowly as γ increases above zero (i.e., as inventory increases above optimal). Note that these percentages are expressed relative to optimal total inventory level, not just the safety stock component, so they would be considerably higher still in cases where higher demand uncertainty induces higher safety stock. Note also that the optimal inventory levels y_i^* are much lower in Table 3 (because α is lower), so the ranges, that are defined based on the suboptimality factor $\gamma_i = (y_i - y_i^*)/y_i^*$, will be much wider. A similar effect may be observed by comparing Figures 2c and d.

5. *Concentration of uncertainty*. The value of pooling is lower when uncertainty is more concentrated. For instance, in Table 4, under normal demand with $\rho_{12} = -0.5$, $p = 0.9$, and $h = 0.1$, the cost reduction due to pooling decreases from 50.0% to 36.5% as uncertainty is more concentrated. Similar behavior occurs for the pooling range: the higher σ_x^2 is, the narrower $[\gamma^L, \gamma^U]$ tends to be. When $\sigma_1 = 0.001$ and $\sigma_2 = 0.399$ so that $\sigma_x^2 = 0.99$, with no correlation and $p = 0.5$ and $h = 0.5$, pooling reduces costs by only about 4%. Therefore, as σ_x^2 increases to its theoretical maximum of 1 for $N = 2$, the value of pooling tends to zero, as expected. Comparing Figures 2a and b reveals a similar effect. In the empirical data with larger N, the effect of concentration of uncertainty is even stronger.

Now that we have examined various nonnormal demand distributions with $N = 2$, let us turn to our empirical demand data with $N \gg 2$.

6. SKU Rationalization in Practice

The objective of this section is to illustrate the results obtained in Sections 4 and 5 using highly erratic demand data and high N values. We use 2 years of demand data from Pellton International, a chemical manufacturer acting as a supplier to automotive suppliers. For

reasons of confidentiality, the true company name is disguised; we have also omitted details of lead times in these illustrations.

6.1. *Pellton International*

A more detailed description of Pellton International can be found in Corbett, Blackburn, and Van Wassenhove (1997, 1999). Pellton supplies rolls of plastic to automotive suppliers, ranging in width from 60 cm to 130 cm wide, in 1-cm increments; moreover, there are several different chemical formulations, colors, etc., leading to a total of over 2,000 SKUs. We focus on their two main formulations, grades S1 and S2, and only on the clear (uncolored) plastic. Within grade S2, there is a special formulation for a key customer, grade S2/P. The plastic is produced in rolls 320 cm wide and slit into rolls of the desired width; this slitting is an integrated part of the production line. Because of the inflexibility of the process, Pellton needs to make to stock. They constantly experienced stockouts and excessive inventory costs and were interested in applying some form of SKU rationalization. There were two options:
- A program called "mastersizes," in which rolls would only be produced in 5-cm increments; the customers would then trim the excess material themselves, which was not a problem, because the customers have to trim the plastic to shape anyway. This would reduce the number of SKUs by approximately a factor 5; this reduction is only approximate, because not all sizes with 1-cm increments are currently demanded.
- Postponement, separating the slitting operation from the production line. This would allow Pellton to produce rolls of 320 cm and to postpone slitting to size until firm orders were received. This option clearly would be preferable but also required more capital investment, whereas the mastersizes program was relatively easy to implement.

We should emphasize that it is not our intention to provide an exact estimate of the value to Pellton of pooling, because more specifics of their processes would need to be taken into account. Rather, we merely use the demand patterns observed by Pellton for illustrative purposes.

6.2. *Demand Structure*

Table 7 shows, among others, the number of SKUs produced and shipped in the 2-year period considered for each grade and several measures of concentration of demand and of uncertainty. The top graph in Figure 3 shows weekly demand for the highest-volume S1 SKU, accounting for 17.4% of total demand for grade S1. The lower graph gives the frequency distribution. Aggregate weekly demand is similarly erratic. Incidentally, this is an excellent illustration of the bullwhip effect: automotive assembly schedules do not vary drastically from week to week, but the graphs show that demand two levels upstream in the supply chain can be totally distorted.

We found that the correlation between SKUs for every product (S1, S2, and S2/P) within weeks is in the $(-0.1, +0.1)$ interval, which can be considered negligible. The demand data could not be explained adequately by any reasonable distribution, after extensive goodness-of-fit tests (using Crystal Ball and Maple). We used three standard tests: the chi-squared and the empirical distribution function–based (EDF) tests, and the Kolmogorov-Smirnov (K-S) and Anderson-Darling (A-D) tests. Stephens (1984) concludes that A-D is the recommended omnibus test statistic for an EDF with unknown parameters, especially when tail behavior of the distribution is important, as when studying inventory policies. The A-D statistic is a variation on the standard K-S test and for n observations is defined as (Anderson and Darling 1954)

$$A^2 = -\frac{1}{n}\left\{\sum_{i=1}^{n}(2i-1)[\ln z_i + \ln(1-z_{n+1-i})]\right\} - n \qquad (14)$$

We also decomposed observed demand into two components: a Bernoulli distribution with empirically determined probability p_i for each SKU i of having positive demand in any given

TABLE 7

Number of SKUs and Concentration of Demand and Uncertainty at Pellton

	S1	S2	S2/P
No. of SKUs before pooling (N)	35	80	34
No. of SKUs with mastersizes	11	34	9
No. of SKUs after postponement	1	1	1
Percentage of total demand	35.7%	48.6%	15.7%
Concentration of uncertainty σ_x^2	1.39	2.70	1.65
$(TC_{BP}(y^*) - TC_{MS}(y^*))/TC_{BP}(y^*)$	34.6%	41.1%	34.6%
$(TC_{BP}(y^*) - TC_{AP}(y^*))/TC_{BP}(y^*)$	77.6%	72.6%	49.4%
$[\gamma^L, \gamma^U]$ For mastersizes	$[-44.0\%, 68.0\%]$	$[-39.0\%, 77.4\%]$	$[-35.2\%, 56.0\%]$
$[\gamma^L, \gamma^U]$ For postponement	$[-51.0\%, 209\%]$	$[-51.8\%, 197\%]$	$[-44.8\%, 81.3\%]$

week, combined with a continuous distribution for the size of that demand. We fitted a wide range of distributions and found that gamma, Weibull, and lognormal performed the best, but even they were rejected more often than accepted and none performed well for a large number of SKUs. We conclude that these empirical distributions allow us to analyze pooling under highly erratic demand data, where concentration of uncertainty is not evenly dispersed and with higher N, which complements the theoretical and numerical analysis of pooling performed so far.

6.3. *Pooling Under Optimal and Suboptimal Inventory Policies*

The stockout and holding costs were estimated by Pellton to be $p = 0.0769$ and $h = 0.00278$ ECU/m^2 per week, respectively, leading to a critical fractile $\alpha = 0.965$. (At the
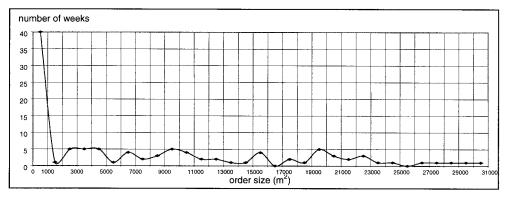

FIGURE 3. Weekly demand pattern and frequency distribution for major product.

time, 1 ECU, or European Currency Unit, was equal to $1.30). Using these costs and the empirical demand data, we followed a methodology similar to that in Section 5. First, we simulated total costs over the 2-year period using the empirical demand data to obtain y_i^*. Second, we repeated this exercise after aggregating weekly demand for all SKUs of the same grade into 5-cm increments (the "mastersizes" program), and after aggregating weekly demand for all SKUs into a single SKU (slitting to order or postponement of customization). In this way, we obtained $TC(\gamma)|_{\gamma=0}$ for the three strategies and the three grades of the product. For some SKUs with very low demand, optimal inventory was zero. The influence of these SKUs on total costs was insignificant; therefore, we set the optimal y_i^* for these SKUs to be small but positive, to avoid division by zero in the definition of the suboptimality factor γ. Third, we simulated total costs, for each of the three strategies, for inventory policies characterized by any suboptimality factor $\gamma \in [-1, 3]$, corresponding to inventory levels from 0 to 3 times the optimal level. This gave us the graphs for total costs for all three strategies, shown in Figure 4 (for grade S1). Fourth, we calculated the cost reductions and the pooling ranges $[\gamma^L, \gamma^U]$, shown in Table 7, by finding those values of γ for which total costs under mastersizes and total costs under postponement for that γ were equal to total costs before pooling when $\gamma = 0$.

The results in Table 7 give us some useful insights. First, we see that the cost reductions from full postponement are much greater than the cost reductions in the bivariate simulations, as one would expect, given that N is much larger here. The impact of concentration of
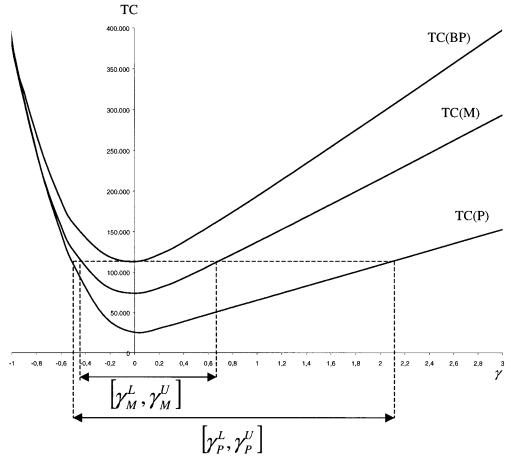

FIGURE 4. Total costs before pooling and after pooling and the "pooling range" for S1, where M = mastersizes and P = postponement.

uncertainty becomes clear by comparing S1 with S2/P, with 35 and 34 SKUs, respectively. The cost reduction for S1 is 77.6%, compared with only 49.4% for S2/P; this is because the concentration of uncertainty is higher for S2/P (at 1.65) than for S1 (at 1.39). The pooling ranges $[\gamma^L, \gamma^U]$ in Table 7 are consistent with those in Table 1 for high α and no correlation: the general behavior is the same, i.e., the range is clearly "skewed to the right," but for the empirical results, the range is wider, because N is higher. This effect can be seen graphically by comparing Figures 1 and 4. Overall, the empirical results are consistent with the theoretical results and the simulations.

7. Conclusions

In this study we have provided a simple framework for comparing the relative value of implementing some form of pooling against that of improving a suboptimal inventory policy. We find that the value of pooling can be negative under suboptimal inventory policies, keeping the same degree of suboptimality before pooling and after pooling. However, there is always a uniquely defined interval within which pooling leads to greater cost reduction than optimizing inventory policy; outside that "pooling range," the reverse is true. This range expands with the number of SKUs. The simulation results show that the pooling range often is very wide, meaning that pooling is more effective than optimizing inventory policy even when current inventory levels are severely suboptimal. The simulations also suggest that the cost reductions achievable by pooling, under an optimal inventory policy, are remarkably robust across a wide range of distributions, but that the sensitivity of the value of pooling to suboptimal policies is less robust. We studied the general multivariate normal case and showed how the combination of correlation and concentration of uncertainty affects the value of pooling.

The managerial implications are significant. This study delves more deeply into how and when one might expect cost reductions to result from pooling. We provide evidence that the value of pooling under an optimal policy is robust across a range of different distributions. In practice, an optimal policy is normally impossible to find, because the demand distributions and cost parameters are not known with any precision. We find that the degree of suboptimality one can have before improving the inventory policy is better than pooling depends heavily on the distribution, and that, in some cases, pooling can increase costs when a suboptimal policy is used. The estimates in Section 6 are very easy to perform in any spreadsheet such as Excel using historical demand data; before implementing pooling, one should, at the very least, conduct an analysis of this type to construct a rough estimate of the potential benefits.

This work poses several interesting questions for future research. For instance, it could be extended to include measures of service performance, and other distributions could be added to the study.[1]

References

ANDERSON, T. W. AND D. A. DARLING (1954), "A Test of Goodness of Fit," *Journal of the American Statistical Association*, 49, 765–769.

BAKER, K. R., M. MAGAZINE, AND H. NUTTLE (1986), "The Effect of Commonality on Safety Stock in a Simple Inventory Model," *Management Science*, 32, 8, 982–988.

CATTANI, K. D. (2000), "Demand Pooling Effects of a Universal Product when Demand is from Distinct Markets," The University of North Carolina at Chapel Hill.

COLLIER, D. A. (1982), "Aggregate Safety Stock Levels and Component Part Commonality," *Management Science*, 28, 11, 1296–1303.

CORBETT, C. J., J. D. BLACKBURN, AND L. N. VAN WASSENHOVE (1997), "Pellton International: Partnerships or Tug of War? Parts A, B, C," Teaching cases, INSEAD.

———, ———, AND ——— (1999), "Partnerships to Improve Supply Chains," *Sloan Management Review*, 40, 4, 71–82.

——— AND K. RAJARAM (2001), "Stochastic Dominance, Multivariate Dependence and Aggregation of Uncertainty," manuscript.

EPPEN, G. (1979), "Effects of Centralization on Expected Costs in a Multi-Location Newsboy Problem," *Management Science*, 25, 5, 498–501.

——— AND L. SCHRAGE (1981), "Centralized Ordering Policies in a Multi-Warehouse System with Leadtimes and Random Demand" in *Multi-Level Production/Inventory Systems: Theory and Practice,* L. B. Schwarz (ed.), North-Holland, New York, 51–67.

ERKIP, N., W. H. HAUSMAN, AND S. NAHMIAS (1990), "Optimal Centralized Ordering Policies in Multi-Echelon Inventory Systems with Correlated Demands," *Management Science*, 36, 3, 381–392.

FEDERGRUEN, A., AND P. ZIPKIN (1984), "Approximations of Dynamic, Multilocation Production and Inventory Problems," *Management Science*, 30, 1, 69–84.

GARG, A., AND H. L. LEE (1999), "Managing Product Variety: An Operations Perspective," in *Quantitative Models for Supply Chain Management,* S. Tayur, R. Ganeshan, and M. Magazine (eds.), Kluwer Academic Publishers, Boston, MA, 467–490.

GERCHAK, Y., AND D. MOSSMAN (1992), "On the Effect of Demand Randomness on Inventories and Costs," *Operations Research*, 40, 4, 804–807.

GROENEVELT, H., AND N. RUDI (2000), "Product Design for Component Commonality and the Effect of Demand Correlation," The University of Rochester, Rochester, NY, manuscript.

HO, T-H., AND C. S. TANG (1998), *Product Variety Management: Research Advances,* Kluwer Academic Publishers, Boston, MA.

JOHNSON, M. E. (1987), *Multivariate statistical simulation,* Wiley, New York.

JÖNSSON, H., AND E. A. SILVER (1987), "Analysis of a Two-Echelon Inventory Control System with Complete Redistribution," *Management Science*, 33, 2, 215–227.

KAPUSCINSKI, R., AND S. TAYUR (1999), "Variance vs Standard Deviation: Variability Through Operations Reversal," *Management Science*, 45, 5, 765–767.

LEE, H. L., C. A. BILLINGTON, AND B. CARTER (1993), "Hewlett-Packard Gains Control of Inventory and Service through Design for Localization," *Interfaces*, 23, 4, 1–11.

———, AND C. S. TANG (1997), "Modeling the Costs and Benefits of Delayed Product Differentiation," *Management Science*, 43, 1, 40–53.

———, AND ——— (1998), "Variability Reduction through Operations Reversal," *Management Science*, 44, 2, 162–172.

RUDI, N. (2000), "Optimal Inventory Levels in Systems with Common Components," The University of Rochester, Rochester, NY, manuscript.

SCHWARZ, L. B. (1989), "A Model for Assessing the Value of Warehouse Risk-Pooling over Outside-Supplier Leadtimes," *Management Science*, 35, 7, 828–842.

STEPHENS, M. A. (1984), "Tests Based on EDF Statistics," in *Goodness-of-fit Techniques,* R. B. D'Agostino and M. A. Stephens (eds.), Dekker, Inc., New York, 97–185.

José A. Alfaro is an Assistant Professor at the University of Navarra, in Pamplona (Spain) since October 2002. He earned his Ph.D. in Business Administration from this same University, in June 1998, after which he was a Visiting Scholar at the Anderson School at the University of California, Los Angeles. From October 1999 to September 2002 he was an Assistant Professor at the University Carlos III of Madrid. His research interests include inventory management and supply chain management. He is especially interested in the study of the bullwhip effect and of traceability, and its consequences for supply chain management. His empirical research is focused on both the automotive and the food industry. He has published other work in the *International Journal of Production Economics.*

Charles J. Corbett is an assistant professor of decisions, operations, and technology management at the Anderson Graduate School of Management at UCLA. Before that, he was a Visiting Scholar at the Owen Graduate School of Management at Vanderbilt University. He received his doctorate degree in operations research from the Erasmus University in Rotterdam and his Ph.D. from INSEAD. His current research focuses on competition and coordination in supply chains, on modeling multivariate dependence, and on environmental issues in international business. His work has appeared in *Sloan Management Review, California Management Review, Operations Research, Management Science, European Journal of Operational Research, the Journal of the Operational Research Society,* and *Environmental and Resource Economics.* Charles Corbett is an associate editor for *Management Science* and guest editor (with Paul Kleindorfer) of a double special issue of *Production and Operations Management* on environmental management and operations.