Product Costing and Pricing

Rajiv D. Banker
University of Minnesota
John S. Hughes
Duke University

SYNOPSIS AND INTRODUCTION: Recently, management accountants have focussed attention on the appropriate treatment of costs associated with resources committed to support activities which do not vary proportionally to production once initial capacities have been set. In one typical case, it is assumed that costs of committed resources will be incurred irrespective of actual usage, and increasing initial capacities to accommodate unexpected demand involves penalties above normal costs. There are at least two important issues that arise in such a case: (1) how should costs of resources committed to, or subsequently required by, support activities enter into pricing and capacity decisions, and (2) what information should the accounting system provide to marketing and production in order to implement those decisions.

Our objective in this paper is to address these two issues. The basic tension which underlies both pricing and capacity decisions is between the nonrecoverable cost of adding capacity before demand becomes known, and the expected penalty-adjusted cost of doing so after demand becomes known. In regard to the first issue, we find in our setting that only normal cost enters into pricing rules established at the time initial capacities are set. (Penalties for exceeding initial capacities are implicit in that for the marginal unit of each activity, normal cost equals expected penalty-adjusted cost.) Findings on the second issue are that the product costing system can be designed without knowledge of demand parameters; the marketing manager only requires activity-based unit costs provided by that system, along with knowledge of demand parameters, to make pricing decisions which are optimal from a firm-wide standpoint; and the production manager only needs expected demand from marketing, along with knowledge of the distribution of random demand shocks, and cost and production parameters, to make initial capacity decisions.

The economic sufficiency of the activity-based unit cost in pricing decisions is in the spirit of Amershi et al. (1989), while the equivalence of marginal cost, normal cost,
and expected penalty-adjusted cost of under-capacity bear similarity to results in studies by Miller and Buckman (1987), Whang (1989) and Hansen and Magee (1993). The main distinctions between these studies and ours lie in assumptions regarding production functions, the nature of information asymmetries, and our incorporation of pricing decisions as well as capacity decisions.

Key Words: Activity-based unit costs, Normal costs, Cost drivers, Practical capacity, Overhead rates, Decentralization.

The body of the paper is organized as follows: section I provides some brief background and a more in-depth overview of our principal findings; section II sets forth our basic assumptions and definitions for various components of the cost and demand functions; section III examines the optimal pricing and capacity decisions of a centralized monopolist firm; section IV depicts minimal information flows sufficient to implement those decisions in a decentralized firm; and section V concludes the paper.

I. Background and Overview

Reporting product costs to managers faced with pricing decisions is an important task of management accountants. This task often involves assigning costs of support activities which do not vary proportionally with production once initial capacities of those activities have been set. Modern prescriptions for assigning costs of support activities to units of production (such as Cooper and Kaplan’s (1987) call for assignment based on relative use, so-called activity-based costing) do not specify the precise correspondence of costs to prices, the efficacy of relying on assigned costs in making pricing decisions, the information required by accountants in designing product cost systems, or the information required by marketing managers in implementing pricing policies. Accordingly, there is a need for accounting research which addresses these aspects.¹

The setting we employ for examining relationships between support activity costs and prices is akin to the classic newsboy problem. Principal departures have to do with our relaxation of constraints on capacities to meet excess demand and our attention to pricing as well as capacity decisions. In particular, we consider a case in which pricing decisions and initial commitments of resources to support activities are made before uncertain demand is realized. Costs of support activities are fixed in the sense that they are incurred even if committed resources are not fully utilized when production required to meet demand is low.² When production required to meet demand is high, constraints implied by committed resources can be relaxed only at an incremental cost in excess of the normal cost of contracting for them in advance of usage. Such incremental

¹ The National Association of Cost Accountants, a predecessor of the present Institute of Management Accountants, prescribed in an April 1953 bulletin: “Judgment and understanding is required in using marginal costs for pricing purposes. It is necessary to consider long range aspects of such business in order to avoid commitments which cannot be dropped... Supplementary allocation of fixed overhead on normal or other volume base must be made to provide product costs for long range planning decisions.”

² This representation of support activity commitments falls in the category of resources supplied in advance of usage discussed by Cooper and Kaplan (1992) and Kaplan (1993). Key elements are that the costs of committed resources will be incurred irrespective of actual usage, and that obtaining support activity resources in excess of initial capacities involves higher than normal costs. As noted later, the essential feature of costs considered is that they vary with costs driver capacities rather than with actual production.
costs arise because of overtime premiums and other costs of arranging for additional resources on short notice.

The first part of our analysis considers the manner in which the costs associated with support activities are aggregated in the functional relationship between costs and optimal prices. The next part considers whether activity-based unit costs provide sufficient cost and production information to permit decentralization of pricing decisions to marketing managers who know or learn demand functions, but nothing more. We find sufficient conditions for optimal prices to incorporate the fixed costs of resources committed to support activities as if those costs were variable production costs. However, this apparently simple relationship between optimal prices and fixed costs is deceptive in that the relationship between the sunk costs of unused overcapacity from such commitments and the penalties for undercapacity are not evident in the pricing function, even though the tradeoff between these two plays an important role in its determination through concurrent support activity resource commitment decisions. These aspects are brought out in the derivation of this function.²

More notably, we find that the sum of unit support activity costs weighted by the number of units of activity required to support a unit of output is sufficient to convey all information concerning support activities relevant for the pricing decisions. In the parlance of Amershi et al. (1989), this sum is economically sufficient with respect to pricing decisions; Demski and Feltham (1976) refer to it as a sufficient cost statistic.³ Given a decentralized firm, activity-based costing provides an aggregate product cost measure which conveys all the cost and production information that the marketing manager requires to make optimal pricing decisions. The marginal value of more detailed variable cost and normal cost information, conditional on the availability of the activity-based unit cost, is zero for both pricing and capacity decisions.

Miller and Buckman (1987) consider conditions under which full or variable costing better approximate opportunity costs of a service department when the department can be characterized as a queuing system. Allocation of fixed costs is considered (as in our model) in conjunction with a capital budgeting problem. Their formulation also provides for idle capacity due to stochastic demand. Their result, that marginal cost and allocated cost of full capacity are the same when cost of capacity is linear, parallels our result on the equivalence of marginal cost and normal cost in that one can interpret normal cost as the allocated cost of (full) capacity set ex ante. However, the distinguishing aspect of our analysis is that optimal prices depend on penalties for exceeding capacity only indirectly through the equality of normal cost with expected penalty-adjusted cost.

¹Formalization of this relationship between product costs and pricing decisions is also useful in assessing conventional accounting practices for determining support activities' overhead rates. For example, the practice of dividing expected costs by the expected level of support activities results in rates that may differ from the per unit cost of support activities imbedded in optimal prices. Some issues relevant to the ABC debate which we do not address in this study include cost pool identification, tradeoffs between variable and nonvariable cost drivers, uncertainty in input-output mappings (a Leontief structure is assumed throughout) and dynamic behavior in demand.

²It is this economic sufficiency of activity-based unit costs that distinguishes our contribution from an earlier generation of cost allocation studies such as those by Kaplan and Thomp (1971), and Kaplan and Welam (1974). In those studies, suitable allocations were determined as part of the solution to a specific production planning problem that required the knowledge of the detailed information aggregated later in the allocations. Since the detailed information was required to solve the production problem, the computation of the allocations did not create value to the organization by reducing its information requirement. In contrast, we show that an activity-based unit cost aggregates all production and cost information relevant to the pricing decision, a priori, and the aggregation function does not require the solution of either a production or a pricing problem. The aggregation procedure can be specified before learning about the demand, production or cost parameters.
even though initial capacity decisions depend directly on the penalty parameters. There is no analogous result in Miller and Buckman since they do not consider a pricing decision.\(^5\)

Whang (1989) adds an incentive conflict between users of a resource and the resource manager. This conflict arises from an information asymmetry and self-interest of the users and manager. The manager seeks to elicit demand information from users both to set capacity of the resource and to allocate that capacity. Given a linear cost function, as well as an ability of the manager to commit not to use reports by users on the value of the service, he shows that (full) cost allocation induces truthful revelation and \textit{ex post} efficiency in allocating capacity. While our analysis provides for information asymmetries from decentralization, we do not consider the incentive conflicts which might then be present.

Finally, Hansen and Magee (1993) consider a setting in which a manager knows the number of projects which may be submitted by users but not the (random) incremental benefit of each project. The manager must choose capacity and a constant cost rule for charging out units of capacity before observing project submissions. Projects are accepted if their incremental benefits exceed this predetermined cost. When the number of projects is large and capacity costs are linear, this rule is optimal in that it implies that unit costs of capacity correspond to the expected incremental benefits of the marginal project.

Similar to the standard newsboy problem, once capacity is set in these studies, it is not possible to adjust that capacity to meet subsequently determined demand.\(^6\) The environment in our model differs in that capacity can be increased as necessary to meet demand when demand becomes known, albeit at a cost penalty. Since (by assumption) it is always to be optimal to meet demand, there are no lost sales. Rather, the marginal cost of adding initial capacity is equated to the expected penalty-adjusted cost of exceeding initial capacity; i.e., the benefit of adding another unit of initial capacity is in avoiding the penalty associated with adding capacity after demand becomes known. Optimal prices are set simultaneously with initial capacities of support activities, so they need consider only the normal cost of adding capacity. The penalties for exceeding that capacity need not be explicitly considered in pricing. Of course, the results just described are just as relevant for a single product and single activity as they are in the more complex multiproduct, multiactivity production function assumed in our analysis.\(^7\)

\section*{II. Structure of Costs and Demand}

We model a monopolist firm which produces multiple products that require multiple activities to support their production. We consider a decentralized organization comprised of a marketing manager and a production manager. The product costing system is designed first, before detailed information on production, cost or demand parameters is available. Information is aggregated in accordance with the \textit{ex ante} design of the system. The marketing manager chooses product prices and the production manager decides the capacities for support activities, based on information available to them. Figure 1 presents a time line of these events.

\footnote{There is a similarity between their results that the marginal cost of capacity equals the value of incremental productivity and our result that marginal cost equals the expected savings of penalty-adjusted cost when capacity is exceeded. The principal distinction lies in the endogeneity of price in our model.}

\footnote{The issue is moot in Whang since demand information is acquired by the manager before capacity is chosen.}

\footnote{Another way of making the point is to say that the insight is sustained by linear production not limited to single product/activity settings.}

Copyright © 2001 All Rights Reserved
Figure 1
Time Line of Events

The principal question we seek to resolve is whether the aggregation of cost information as in an activity-based cost system will provide the relevant information for the pricing decisions which follow. We address this question by showing how costs enter the optimal solution of both the pricing and capacity problems when all information and decisions are centralized. Subsequently, we show that decentralized decisions using activity-based unit costs result in the same decisions (and therefore, the same expected profit) as those made by a central decision maker having access to all detailed information. Toward this end, we begin by considering a centralized organization in which all information depicted in figure 1 is available to a single individual, who makes all the decisions. We defer the consideration of a decentralized organization with multiple decision makers until section IV.

For now, we assume demand for each product is linear in its price and includes an additively separable stochastic component. Specifically, derived demand functions are defined as follows:

\[ q_j = \alpha_j - \beta_j p_j + \varepsilon_j, \quad j=1,\ldots,J, \]  

where \( p_j \) and \( q_j \) represent prices chosen by the firm and quantities demanded by customers, respectively, \( \alpha_j \) and \( \beta_j \) are positive parameters (\( \alpha_j, \beta_j > 0 \)) whose values are learned after the design of the cost system but before the pricing and capacity decisions, and \( \varepsilon_j \) are random disturbances with \( E(\varepsilon_j) = 0 \) that represent the uncertainty about demand remaining after these decisions are made. The cost and demand parameters, and the lower supports of \( \varepsilon_j \) distributions are assumed to be such that \( \varepsilon_j > 0 \) for all \( j=1,\ldots,J \) under optimal price choices. We will also use the vector notation \( p = (p_1,\ldots,p_J) \), \( q = (q_1,\ldots,q_J) \).

Let \( C(\mathbf{x}, \mathbf{q}) \) denote a cost function which incorporates both variable and fixed costs of production including costs of support activities, where \( \mathbf{x} = (x_1,\ldots,x_I) \) represents a vector of activity cost driver capacities implied by resources committed to activities \( i=1,\ldots,I \). The cost accounting literature is ambiguous about the differentiation between variable and fixed costs in how they relate to activity cost drivers. Our definitions presented below reflect assumptions about commitment and usage of resources that underlie activity-based costing (Cooper and Kaplan 1992).

We define total costs of production as follows:

\[ C(\mathbf{x}, \mathbf{q}) = \sum_{j=1}^{J} \delta_j q_j + \sum_{i=1}^{I} m_i x_i + \sum_{i=1}^{I} \xi_i(x_i, \mathbf{q}), \]  

where
\[
\xi_i(x_i, q) = \begin{cases} 
0, & \text{if } \sum_{j=1}^{l} \mu_j q_j \leq x_i, \\
\theta_i m_i (\sum_{j=1}^{l} \mu_j q_j - x_i), & \text{if } \sum_{j=1}^{l} \mu_j q_j > x_i,
\end{cases}
\]

and

\begin{itemize}
\item \(v_j\) = variable cost per unit of product \(j\),
\item \(m_i\) = normal cost per unit of cost driver for activity \(i\),
\item \(\theta_i\) = factor to reflect penalty for exceeding cost driver resources committed to activity \(i\), \(\theta_i > 1\), and
\item \(\mu_j\) = number of units of activity \(i\) required to support one unit of product \(j\).
\end{itemize}

The first term on the right hand side (RHS) of (2) represents variable costs for resources which are supplied as used; the second term is the fixed cost of committed resources which will be incurred irrespective of whether capacities are fully utilized; and the third term is the penalty incurred contingent on whether initial capacities of support activities implied by those commitments are exceeded in order to meet demand.\(^8\) Note that the three components of \(C(x, q)\) have different cost drivers: variable costs of production are driven by outputs, normal costs of support activities are driven by the levels of cost driver capacities fixed for those activities, and penalties for exceeding initial capacities are driven by both outputs and cost driver capacities.\(^9\) This structure is consistent with Walters’ (1963) econometric survey which indicates that cost functions are approximately linear until capacity is reached, but marginal costs increase sharply once capacity is exceeded.

The pattern of managers committing to make resources available to support production, in advance of knowing precisely how much will be used, characterizes many if not most activities performed in an organization. Such activities include purchasing, receiving, stores, materials handling and moving, scheduling, machine setup, maintenance, janitorial services, quality inspection and engineering support and may constitute a large proportion of total costs (Miller and Vollman 1985; Banker et al. 1992). What we and others term the normal costs of such activity resources are usually labelled as fixed costs in textbooks because they must be incurred once the initial capacity commitments are made regardless of whether such resources remain idle or unused during the period. The essential feature of these costs is that they do not vary with actual production, but rather vary with the cost driver capacities.

For instance, suppose a plant employs ten workers responsible only for performing quality inspections. A fixed proportion of all production is inspected on a sampling basis. Quality inspection costs, however, are not variable costs. Each worker can perform 800 inspections per month. By hiring the ten quality inspectors at fixed monthly wages and benefits of $4,000 per...

\(^8\) The usual activity-based costing model includes both volume-related and batch-related activity costs. While the latter costs are likely step costs, suggesting an integer programming optimization problem, with our well-behaved cost and revenue functions, we assume that optimal solutions obtained treating them as continuous functions yield good approximations for large \(q\) and \(x\). Product-sustaining costs may also be included by defining activities that correspond to individual products. Such costs are not relevant for decisions in our model, however, because we assume \(q > 0\) for all \(j\).

\(^9\) Variable costs correspond to resources whose consumption can be adjusted to match precisely the quantities required by actual production. Direct materials are variable costs because materials are consumed in proportion to current production. Costs of electrical power to operate machines, commissions to sales personnel based on actual sales, and piece-rate direct labor costs are other common examples of variable costs. Direct labor hired on a weekly, monthly or some other time basis, however, must be paid their wages for the period even if they are idled by lack of work. Such direct labor costs do not vary with production in the short run.
worker, the plant commits to spend $40,000 to make available resource capacity for \( x_i = 8,000 \) (\( =800 \times 10 \)) quality inspections a month. The normal cost per inspection is \( m_i = $5 \) (\( =40,000/8,000 \)).

In addition to the normal costs of support activities, there are the costs of additional resources that may be required in case initially committed activity levels are not adequate to support actual production needs. Let \( z_i \) denote the usage of activity \( i \) implied by actual production \( q \). Therefore, the total number of units of support activity \( i \) required for the production of \( q_j \) units of all \( j=1,...,J \) products is

\[
z_i = \sum_{j=1}^{J} \mu_{ij} q_j.
\]

The activity-based costing model assumes that activity cost resource usage \( z \) varies with actual production \( q \), but its total normal cost

\[
\sum_{i=1}^{I} m_i x_i,
\]

varies with the committed cost driver levels \( x \) and not production volume \( q \). Additional definitions which will prove useful for our later analysis include:

\[
\overline{q}_j \equiv E(q_j \mid p_j) = \alpha_j + \sum_{j=1}^{J} \beta_j p_j, \text{ and } y_i = \sum_{j=1}^{J} \mu_{ij} q_j, \text{ so that } z_i = \overline{z}_i + y_i.
\]

Hence, \( \overline{q}_j \) is the expected demand for product \( j \) and \( \overline{z}_i \) is the demand for support activity \( i \) induced by expected demand for all products \( j=1,...,J \). A marginal distribution of each \( y_i \) can be induced as a convolution of the distributions of the \( \epsilon_{ij} \)'s; let \( \phi(y) \) and \( \Phi(y) \) denote the probability density and distribution functions, respectively.

If the actual demand for support activity \( i \) is less than available capacity (i.e., \( z_i < x_i \) or equivalently \( y_i < x_i - \overline{z}_i \)), then the surplus resources expire unused and the total cost of activity \( i \) remains at \( m_i x_i \). If the actual demand for support activity \( i \) exceeds available capacity (i.e., \( y_i > x_i - \overline{z}_i \)), then an additional cost of \( \theta m_i \), \( \theta > 1 \), is incurred for each additional unit of cost driver \( x_i \) that is actually consumed.\(^9\) That is, the additional costs when \( z_i > x_i \) are \( \xi(x_i, q) = \theta m_i (z_i - x_i) \), which exceeds the normal costs of additional resources by \( (\theta-1)m_i (z_i - x_i) > 0 \), and provides the economic rationale for managers to commit in advance rather than waiting to acquire only the precise quantities required after demand for support activities is known with certainty.\(^1\)

More broadly speaking, we can view the premium per unit paid to expand capacities of support activities such as quality inspection as a differential between costs of advance commitments and costs of contracting in the spot market (Cooper and Kaplan 1992). Other penalty costs associated with delays, queues and congestion resulting from actual demand for support activities exceeding committed resources, are less perceptible (Banker et al. 1988), but also important.

Returning to our example of quality inspections, assume that the plant manufactures three products that are inspected on a sampling basis at the rate of one out of 200, 100 and 80 units of the products respectively, so that \( \mu_s = 0.005 \) (\( =1/200 \)), \( \mu_i = 0.010 \) and \( \mu_s = 0.0125 \) inspections per product unit. If actual production is 600,000, 350,000, and 80,000 units respectively, then \( z_i = 7,500 \) (\( =3,000 + 3,500 + 1,000 \)) inspections will be required in that month, notwithstanding that

\(^9\) Linear penalties, while somewhat restrictive, are adequate to produce the desired tension between over and under commitments of resources.

\(^1\) Observe that with multiple products there is a diversification benefit to having one firm supply support activities for all products provided the stochastic components of demand are not perfectly and positively correlated. The relevant measure in determining when penalties will be incurred in meeting joint demand is an average realization. Thus, even though economies of scope are not included directly in our linear, separable cost model, there is a rationale for multiple products to be housed in a single firm. There are externalities implied by the possibility that high demand for one product could be offset by low demand for another product, thereby reducing the prospect of penalties being incurred in meeting total demands.
capacity for $x_i = 8,000$ inspections was available. In other words, $40,000$ of costs would be incurred even though actual production does not require utilization of all resources made available by that expenditure. Alternatively, if actual production for the third product increased to $200,000$, then a total of $z_i = 9,000$ ($= 3,000 + 3,500 + 2,500$) inspections would be required. Assuming that the plant could contract for an additional $1,000$ inspections at $7.50$ per inspection ($5.00$ plus a premium of $2.50$), the actual cost of support activities would be $47,500$ even though the cost would have been just $45,000$ if the plant had committed to a capacity for exactly $9,000$ inspections in advance.

We define the activity-based cost $c_j$ per unit of product $j$ to be

$$c_j = v_j + \sum_{i=1}^{I} m_i \mu_{ij},$$  \hspace{1cm} (3)

where the costs of support activity resources are assigned to the products based on the per unit requirements $\mu_{ij}$ of each product $j$. Clearly, if demand were certain, then the initial capacity commitments would just meet demand and $c_j$ as expressed in (3) would be the cost actually incurred per unit produced. The interesting question is what role $c_j$ plays in pricing and capacity decisions when demand is uncertain.

III. Pricing and Capacity Decisions

Capacity Decision

Suppose the monopolist firm is centralized and chooses the price of each product $j$ and the level of capacity $x_i$ for each support activity $i$. Customers respond, and actual demand $q_i$ is realized in accordance with (1). We assume that it is worthwhile for the firm to fill all realized demand.\footnote{We shall note later that this assumption can be expressed in terms of exogenous conditions on cost and demand parameters that ensure that optimal prices exceed the unit production costs even when demand for support activities exceed their capacities.}

Expected profits conditional on one set of demand and cost function parameters $(\alpha, \beta, m, \theta, v)$ can now be stated as follows:

$$E[\pi(x, p) | \alpha, \beta, m, \theta, v] = \sum_{j=1}^{J} (p_j - v_j)q_j - \sum_{i=1}^{I} m_i x_i - \sum_{i=1}^{I} \theta_i m_i \int_{x_i - z_i}^{\infty} [z_i + y_i - x_i] \phi_i(y_i) dy_i$$  \hspace{1cm} (4)

The three terms on the RHS of (4) are expected contribution margin, fixed costs of committed resources, and expected penalties for exceeding initial capacities of support activities, respectively.

While prices are set before commitments to support activities in the decentralized setting illustrated in figure 1, it is convenient for analytic purposes to first characterize optimal commitments.\footnote{The decisions are interchangeable in terms of the order in which they are addressed, since analytically speaking they are made simultaneously.} Differentiating (4) with respect to $x_i$ and simplifying yields the following first-order conditions for characterizing the capacity levels implied by optimal commitments, $x$ (henceforth called optimal capacity):

$$\frac{\partial E[\pi(x, p)]}{\partial x_i} = m_i + \theta_i m_i \int_{x_i - z_i}^{\infty} \phi_i(y_i) dy_i = 0, \hspace{1cm} i = 1, \ldots, I.$$  \hspace{1cm} (5)

The two terms on the RHS of (5) reflect the tradeoff between the cost of committing resources sufficient to provide capacity for an additional unit of support activity $i$, and the expected cost of
adding capacity should actual demand for outputs induce a demand for a further unit of that activity, respectively. Note that if there were no uncertainty, then total costs would be minimized by choosing \( x_i = z_i \). However, given uncertainty, one would seek to balance overinvestment in capacity when demand is low against penalties incurred when demand is high.

It follows from (5) that

\[
1 - \Phi_i(x_i^* - z_i) = 1/\theta_i, \quad i = 1, \ldots, I.
\]  

(6)

Interpreting (6), we can see that the magnitude of the per unit penalty factor \( \theta_i \) governs the relationship between optimal capacity of the support activity and the requirements of expected demand. To illustrate, assume that the inverse function \( \Phi_i^{-1} \) exists, implying that

\[
x_i^* = \overline{z}_i + \Phi_i^{-1}\left(1 - \frac{1}{\theta_i}\right).
\]  

(7)

The second term on the RHS of (7) represents capacity dedicated to activity \( i \) beyond expected requirements. Since the actual utilization of that capacity depends on the joint realizations of demand for all \( J \) products, it is clear that there is an implicit interdependency across products which, if ignored, would lead to suboptimal decisions.

Further, if \( \phi(*) \) is symmetric, then \( \Phi_i^{-1} \) equals zero only when \( \theta_i = 2 \). Thus, when the penalty for inadequate capacity is sufficiently large (\( \theta_i > 2 \)), the optimal capacity will exceed expected demand for that activity, and when it is sufficiently small (\( \theta_i < 2 \)), the reverse will be true. More generally, we can observe that

\[
\frac{dx_i^*}{d\theta_i} = \left(\frac{1}{\phi_i(x_i^* - \overline{z}_i)}\right)\theta_i^2 > 0,
\]  

(8)

which reflects the intuitively appealing notion that optimal capacity increases with the penalty (\( \theta_i \)) for inadequate capacity.

**Pricing Decision**

Differentiating (4) with respect to \( p_j \) and simplifying, we next obtain the following first-order conditions characterizing the optimal product prices:

\[
\partial E[p(x, p)]/\partial p_j = \overline{q}_j - (p_j - v_j)\beta_j + \sum_i \theta_i m_i \mu_{ij} \beta_j \int_{x_i = \overline{z}_i} \phi_i(y_i)dy_i = 0, \quad j = 1, \ldots, J.
\]  

(9)

As the RHS of (9) illustrates, an optimal price balances the marginal loss in contribution margin from increasing price evident in the first two terms, against the marginal expected reduction in penalties for exceeding optimal capacities of support activities contained in the third term. Alternatively, \( \overline{q}_j - p_j \beta_j \) represents the marginal revenue lost from an increase in price, while

\[
v_j \beta_j + \sum_{i=1}^I \theta_i m_i \mu_{ij} \Phi_i(x_i - \overline{z}_i)
\]

represents the expected marginal cost savings from such a change.

---

\(^{14}\) Given our assumptions that initial committed capacity can be expanded at a penalty and that realized demand is always met, then solutions to both the capacity and pricing decisions are interior and fully characterized by (5) and (9).
If it were certain that demand would not exceed capacities, then (9) would imply an optimal price \( p_j^* = (\alpha_j / 2\beta_j) + (v_j / 2) \). Costs of resources committed to support activities would not enter the determination of prices because of surplus capacity. If the demand were certain to exceed all capacities, then (9) would imply an optimal price

\[
p_j^* = (\alpha_j / 2\beta_j) + (v_j + \sum_{i=1}^{I} \theta_i m_i \mu_{ij}) / 2.
\]

Penalty costs for exceeding all capacities would now enter the optimal prices.

In general, the expected cost of incurring penalties for exceeding capacities must be considered in optimal prices. This expected cost is optimally traded off against the \textit{ex ante} cost of activity capacity as in (5). Therefore, rearranging the terms in (9) and making use of (3), (6) and the definition of \( \bar{q}_j \) yields the following optimal prices:

\[
p_j^* = A_{0j} + A_{1j} c_j, \quad j = 1, \ldots, J,
\]

where \( A_{0j} = \alpha_j / 2\beta_j \) and \( A_{1j} = 1 / 2 \). Thus, given the knowledge of the demand parameters \((\alpha_j, \beta_j)\), the optimal price of each product is a function only of its activity-based cost \( c_j \).\footnote{Since \( q_j = \alpha_j - \beta_j p_j^* = (\alpha_j - \beta_j c_j) / 2 > 0 \), it is clear that prices exceed activity based costs (i.e., \( p_j^* > c_j \)). It can also be seen that \( \beta_j (A_{0j} - A_{1j} c_j)^2 + L_j \gamma > 0 \), where \( L_j \) is the lower support of the distribution of \( c_j \), ensures that \( q_j > 0 \) for optimal price choices. In addition, \( A_{0j} + A_{1j} c_j > v_j + \sum_{i=1}^{I} \theta_i \delta_i m_i \mu_{ij} \) ensures that it is optimal for the firm to fill all realized demand. We further note that prices are increasing in the intrinsic worth of the product to the customer as measured by \( \alpha_j \), and decreasing in \( \beta_j \), the sensitivity of demand to price (Dixit 1979).}

More remarkable is that optimal prices given by (10) are indistinguishable from the monopoly prices which would be chosen if all resources were supplied as used, i.e., if there were no penalties for exceeding initial capacities (i.e., \( \theta = 1 \)).\footnote{Note that although optimal prices are invariant with respect to optimal capacities, optimal capacities depend on prices since they, along with penalties, determine the tradeoff between over-investing in support activities at the outset and incurring extra costs to meet demand if initial capacities are exceeded.}

With no penalties, firms can make all costs variable and avoid fixed costs of overcapacities by committing initial capacities of zero and simply providing the resources needed to meet demand when demand becomes known. Alternatively, if there were no uncertainty, then capacities would be initially set at precisely the levels required of known demand and the penalties would never be incurred. The interesting aspect is that even under uncertainty, when prices and initial capacities are set simultaneously, the penalties do not enter explicitly into optimal prices but rather implicitly through the interdependency of pricing and capacity decisions. The penalties are relevant to setting initial capacities given optimal pricing. However, once capacities are set then prices require no further adjustment to capture costs of exceeding those capacities.\footnote{This would not be true if further demand information were obtained after capacities had been set, but before prices were determined. See Banker et al. (1993) for an analysis of this case.}

\textit{Economic Sufficiency}

The optimal capacity decision represented by the expression in (7) requires additional information on \( \mu_{ij}, \Phi_i \) and \( \theta_i \) to determine the expected demand on activity resources and the optimal tradeoff for adding an incremental capacity unit at the margin. (The capacity decision does not require detailed information on the variable costs \( v_j \) or the normal activity costs \( m_i \).)
Therefore, for the centralized organization we have considered in this section, the activity-based costs \( \{c_i\} \) are **economically sufficient** for the cost parameters \( \{v_j, m_i\} \) with respect to both the pricing and the capacity decisions. The dimensionality of the information required for these decisions is thus reduced from J+I to only J because of the aggregation.

It is useful at this juncture to contrast our results from those in the cost allocation literature in the early 1970s (e.g. Kaplan and Thompson 1971; Kaplan and Welam 1974). These studies begin with a mathematical programming model of the production process, and use the shadow prices obtained from actually solving the mathematical program to determine the cost allocations. The allocated costs thus constructed using shadow prices are found to be consistent with the optimal solution of the mathematical program. A problem with this approach, however, is that the mathematical program must be solved first using all the detailed information to obtain the cost allocations, and therefore no informational saving or value in supporting the production decisions is generated by the computation of these allocations.

Our analysis uses a different model and provides very different results. The aggregation of costs into activity-based product costs \( c_i \) does **not** require that the production (activity resource commitment) or the pricing problem be solved first. Rather, the aggregation only requires knowledge of production and cost parameters. More importantly, the pricing decision does not require any additional cost or production information beyond that contained in the aggregate activity-based product cost data. Furthermore, support activity capacities can be chosen based only on the demand, production, penalty and activity-based unit costs. This decision does not require detailed information about variable costs or normal support activity costs. In other words, the marginal value of variable cost and normal support activity cost information \( \{v_j, m_i\} \), conditional on the availability of the activity-based costs information \( \{c_i\} \), is zero with respect to both pricing and activity capacity decisions.

**IV. Decentralized Information and Decisions**

*Reduction in Information Flows*

We have so far analyzed the pricing and activity capacity decisions only for a centralized organization. Now consider an organization with a production manager and a marketing manager, each receiving limited information and each responsible only for specified decisions as depicted in Figure 2. We assume that two different managers are required for reasons, such as efficiencies that result from functional skill specialization, which are commonly presented in the literature on decentralization. A further assumption that follows naturally is that production and cost parameter information is available initially only to the production manager and demand information only to the marketing manager.

The product costing system is designed first, before the parameters characterizing the cost, production and demand functions are known. It specifies responsibility for information collection, aggregation and communication, and for product pricing and activity resource commitment decisions. Information aggregation and decision rules are specified up front, and are implemented based on the information collected subsequently by the product costing system.

The production manager learns the values of the production parameters \( \mu_j \) measuring the quantity of activity \( i \) required to support the production of a unit of product \( j \), and the values of the cost parameters \( v_j, m_i \) and \( \Theta \) based on production engineering, purchasing, payroll and other pertinent records. Activity-based product costs \( c_j \) are determined using this information in accordance with the expression in (3).

This aggregated product cost information is communicated to the marketing manager, who augments it with information about demand parameters to determine the optimal price \( p_j = (\alpha_j/2\beta_j) \)
Figure 2
Decentralized Information Flows and Decisions

Design
Product Costing
System

Production Manager

Collect information on production and cost parameters $\mu_j$, $v_j$, $m_j$, $\theta_j$.

Determine product costs
$c_j = v_j + \sum_{i=1}^{1} m_i \mu_j$

Communicate $c_j$ to marketing

Determine expected activity demand
$\bar{z}_j = \sum_{i=1}^{j} \mu_j \bar{q}_i$ and its distribution:
$\Phi_j(y_j = \sum_{i=1}^{j} \mu_j \varepsilon_j)$

Communicate resources to activity levels
$x^*_i = \bar{z}_i + \Phi^{-1}_j (1 - \frac{1}{\theta_j})$

Marketing Manager

Collect information on demand parameters $\alpha_j$ and $\beta_j$, distributions $F_j(\varepsilon_j)$.

Choose product prices
$p^*_j = (\alpha_j / 2 \beta_j) + (1/2)c_j$

Communicate expected demand
$\bar{q}_j = \alpha_j - \beta_j p_j$ and distributions $F_j$ to production

Communicate optimal prices $p^*_j$ to customers
+ (1/2)c_{j} for each product j. Having made the pricing decision and learning the distribution of demand, the marketing manager also determines the expected demand \( \bar{q} = \alpha_j \beta p_{j} \) and communicates information about both the expected demand and the demand distribution to the production manager. In our decentralized organization, the marketing manager has the best access to existing and potential customers, and therefore, is also responsible for communicating the product prices to them. Information on penalties, demand distribution, expected demand and production parameters enables the production manager to determine the optimal levels of capacities to commit to each activity in accordance with the expression in (7). The manager also takes the necessary actions to acquire the labor, tools, machinery and other resources required to provide these optimal activity capacities.

It is apparent from the above description of the decentralized information collection and decision making system that it results in the same (optimal) decisions as in the centralized organization we considered earlier. Yet, the decentralized arrangement requires considerably less information be communicated by or to the two manager's responsible for making decisions than the totality of information collected by the organization. Assuming non-trivial costs for the transmission and assimilation of information, such an arrangement and the concomitant reduction of information flows is of value to the organization.

Of particular interest to us is the reduction in the information flows because of the aggregation of detailed cost information into activity-based product costs. Given the organizational separation of the availability of production, cost and demand information, some information must flow from one organizational unit to another for optimal decisions to be made. As we have seen earlier in (10), the optimal pricing decision requires \((I J + J + I)\) dimensions for production and cost information \((\mu_{ij}, \nu_{ij}, m_{ij})\), and \(2J\) dimensions for demand information \((\alpha_j, \beta_j)\). By communicating aggregate activity-based product cost information \((c_{j})\) that is economically sufficient for \((\mu_{ij}, \nu_{ij}, m_{ij})\) with respect to pricing decisions, these information flows are reduced to only \(J\) dimensions. Although optimal capacity decisions require all demand, production and cost information, the only information required to be communicated to the production manager, who is already cognizant of the distribution of random demand shocks and cost and production parameters, is expected demand at the optimal prices. Thus, demand and price information is reduced to only the \(J\) dimensions of expected demand information \((\bar{q}_{ij})\) needed for the capacity commitment decision.

Nonetheless, while there are clearly savings in information flows implied by the economic sufficiency of \((c_{j})\) for \((\mu_{ij}, \nu_{ij}, m_{ij})\) in making pricing decisions, and of \((\bar{q}_{ij})\) for \((\alpha_j, \beta_j, p_{j})\) in making initial capacity decisions, some transmission of information is necessary to implement optimal decisions. In other words, delegation of capacity decisions to the production manager and pricing decisions to the marketing manager does not imply that cost and production parameters are only germane to capacity decisions, and that demand parameters are only germane to pricing decisions. Thus, we have something less than complete decentralization. However, the dimensionality of information regarding those parameters can be reduced for transmission to the other unit without distorting those decisions.

**Determination of Overhead Rates**

Given the economic sufficiency result, we can now meaningfully address the question of determining overhead rates for the purpose of aggregating costs. Recommended accounting practice for product costing, per standard textbooks, calls for determining *budgeted* overhead rates, dividing budgeted activity costs by budgeted volume of the activity base. However, budgeted costs and volumes are generally not defined in precise terms. If the term *budgeted* is
interpreted as \textit{expected}, then an overhead rate applicable to the \textit{i}th activity could be calculated as shown below:

\[
\tilde{m}_i = \left[ m_i x_i^* + \theta_i m_i \int_{x_i^* - \bar{z}_i}^{\infty} (z_i - x_i^*) \phi_i(y_i) dy_i \right] / \bar{z}_i
\]

\[
= m_i + \left[ \theta_i m_i \int_{x_i^* - \bar{z}_i}^{\infty} y_i \phi_i(y_i) dy_i \right] / \bar{z}_i
\]

\[
> m_i. \tag{11}
\]

Hence, we see that the budgeted overhead rate \( \tilde{m}_i \) is greater than \( m_i \), implying that \( \tilde{m}_i \) is inappropriate for pricing purposes. This distortion occurs because the rate \( \tilde{m}_i \) reflects expected penalties for exceeding initial capacities for support activities, whereas optimal prices depend only on the normal costs \( m_i \).\textsuperscript{18}

An alternative definition of budgeted volume is the capacity of support activity that the firm initially makes available at the beginning of the budget period, a definition of budgeted volume consistent with \textit{practical capacity}. If, in addition, expected penalties such as overtime premiums and delay costs are recorded separately in the accounting system, and not included together with the normal costs of activities, then the overhead rate is calculated simply as \( m_i x_i / x_i^* = m_i \). Thus, a normative implication of our analysis is that the activity overhead rates employed for pricing should be calculated based solely on normal costs computed by dividing costs of committed resources by practical capacity. Expected penalties such as overtime premiums or delay costs should be excluded. Cooper and Kaplan (1992) also recommend dividing by practical capacity, but do not discuss the change needed in the numerator to eliminate expected penalties such as overtime premia. (See also Harvard Business School case 191-073 "Micro Devices Division.")

\textit{Generalized Demand}

While we have employed a linear demand function for expositional convenience, our results on economic sufficiency extend to any general, well-behaved demand function. Observe that by using equation (7) for optimal capacities, we can rewrite the expected profits in (4) as:

\[
E[\pi(x^*, p)] = \sum_{j=1}^{J} (p_j - c_j) q_j - \sum_{j=1}^{J} \theta_i m_i \int_{\Delta_i}^{\infty} y_i \phi_i(y_i) dy_i
\]

where

\[
\Delta_i = x_i^* - \bar{z}_i = \Phi_i^{-1} \left( 1 - \frac{1}{\theta_i} \right).
\]

Since the second term in the above expression is independent of product prices and quantities, it is transparent that the evaluation of expected profits requires only the separate consideration of each product's demand function and its activity-based cost. Given our assumption that \( q_j > 0 \) for all \( j = 1, \ldots, J \), expected profits are maximized by prices that maximize expected gross profit for each product, where expected gross profit is computed as gross margin (revenues less activity-based costs) times the expected demand.

\textsuperscript{18} This source of distortion is distinct from that discussed in the activity-based costing literature. The distortion is caused by overstated overhead rates even when all cost drivers are recognized properly.
The correspondence between activity-based costs and optimal prices, therefore, extends to all well-behaved demand functions which are monotone decreasing in price. Analogous to (1), let

\[ q_j = \bar{Q}_j(p_j) + \epsilon_j, \quad j = 1, \ldots, J, \]  

with \( p_j + \bar{Q}_j(\bullet) < 0 \). Writing \( H_j(p_j) = p_j + \bar{Q}_j / \bar{Q}_j \), it follows that optimal prices are given by

\[ p_j^* = H_j^{-1}(c_j), \quad j = 1, \ldots, J. \]  

Although the prices in (14) no longer possess the same linear form, they continue to be a function of the same activity-based costs as before. In other words, activity-based costs are economically sufficient for detailed production and cost information, with respect to the product pricing decision, for any well-behaved downward-sloping demand function. They are also economically sufficient for \( \{ v_j, m_j \} \) with respect to the capacity decision as expected demand can be determined directly using (13) and (14). Thus, we could replace the linear demand and inverse demand functions contained in figure 2 with generalized functions with no qualitative change in our discussion.

V. Concluding Remarks

Starting with a centralized regime, we have characterized how support activity costs enter optimal pricing and initial capacity decisions in a setting related to the classic newsvendor problem. Under conditions wherein capacity costs are committed \textit{ex ante} and demand in excess of capacities is met by incurring penalties, our principal results are that (1) a firm facing uncertain demand at the time of those decisions will find normal activity-based unit costs economically sufficient in arriving at optimal prices; and (2) the firm can achieve the same results when it is decentralized along functional lines. The product costing system can be set in place without benefit of \textit{a priori} demand, production and cost information; marketing managers need only learn activity-based unit costs in order to implement optimal pricing rules; and production managers need only learn expected demand, along with cost and production information, in order to make optimal commitments of resources to support activities. Notwithstanding the limitations of our assumptions—including divisible capacities, piecewise linear costs of resources, and absence of incentive problems which might arise under decentralization—the key role of normal activity unit costs in pricing decisions establishes further economic rationale for the emphasis given to activity-based cost systems.\(^{19}\)

Among the possibilities for continued research along these lines, we note that the situation changes importantly when more can be learned about demand after initial capacities have been set and yet before prices need be determined. The activity-based unit cost aggregate, while appropriate as a benchmark, is no longer economically sufficient. Further demand information leads to an update of expected demand beyond the point in time when capacities of support activities can be expanded at normal costs. Explicit consideration must be given to the expected penalties required to ensure that demand is met. However, doing so is difficult except for certain distributions of demand.\(^{20}\) Likewise, relaxing the restriction that no product is worth dropping, or that demand must be met, can also result in complexities which this paper does not consider. The difficulties posed by these generalizations call for research on the properties of cost-based

\(^{19}\) A recent survey of 566 controllers by the Institute of Management Accountants (Schiff 1991) indicated that 30 percent of the respondents had either already implemented an activity-based system or were considering it.

\(^{20}\) Specifically, Banker et al. (1993) show that closed-form solutions to joint capacity and pricing decisions are feasible for the uniform distribution, but not in general.
heuristics such as the work of Hansen and Magee (1993). Although their model is formulated along different lines, it seems evident that development of such heuristics offers promise for a better understanding of the relationship between accounting measures of activity costs and pricing and capacity decisions. Empirical work in this area could examine plant- and firm-specific cost and production data to test the validity of the assumptions underlying product costing procedures. For example, if the evidence were to indicate significant violations of the Leontief characterizations of production, or other important elements of the structure we impose, then it may be necessary to modify present prescriptions as well as the model upon which they can be based. While much work remains to be done, this study and those cited provide the beginnings of a positive theory of product costing.

21 Any formal model is likely to be a stylized abstraction from reality, and ours is no exception. Accordingly, we would expect minor deviations between the data and the structure of our model such as those captured by noise terms in econometric specifications.

References