June 1982

Division of Research
Graduate School of Business Administration
The University of Michigan

MEASURING THE PRECISION OF STATISTICAL COST ALLOCATIONS

Working Paper No. 305

Roger L. Wright

The University of Michigan

FOR DISCUSSION PURPOSES ONLY

ABSTRACT

This paper examines the use of regression analysis and other statistical methods for allocating indirect costs. When multiple regression is used to estimate the weights of several allocation factors, conventional standard errors and correlation coefficients can be misleading with respect to the statistical precision of the cost allocations. Alternative measures of precision which use linear prediction theory and Bayesian inference are developed. The proposed methods are illustrated in the context of a university indirect cost study.

KEY WORDS: Bayesian inference, Generalized linear model, Joint costs, BLU Prediction

Author's Footnote:

Roger L. Wright is Associate Professor of Statistics, Graduate School of Business Administration, The University of Michigan, Ann Arbor, MI 48109.

The author wishes to thank M. Bryant, J. S. Demski, and W. E. Wecker for helpful suggestions on earlier versions of this paper.

1. INTRODUCTION

This paper examines the use of statistical methods, particularly regression analysis, in cost allocation studies. Regression analysis is widely used by economists and managerial accountants to estimate cost relationships, but there seems to be no satisfactory technique for determining the statistical precision of the allocations that result from these estimates. Conventional measures of precision can be misleading because they fail to take into account certain aspects of the allocation process. A successful technique must also separate the statistical issues from the difficult policy decisions that arise in connection with the allocation of fixed and joint costs. This paper introduces a new approach that is based on the most relevant features of the cost allocation process, and seems to yield a more appropriate measure of the precision of a regression-based cost allocation.

Section 2 of the paper discusses several alternative statistical procedures that can be appropriate when costs are separable, i.e., free of fixed or joint components. This discussion includes an approach based on Goldberger's (1962) theory of prediction under the generalized linear regression model. Section 3 develops the main approach of the paper, which uses Bayesian inference with a class of multiple regression procedures for allocating both separable and joint costs. The Bayesian approach isolates the statistical precision achieved when estimates are used to implement a prespecified allocation policy formulated in terms of the parameters of the underlying cost function. Section 4 illustrates these methods in the context of a university indirect cost study. The rest of the present section gives a brief introduction to the cost allocation problem.

Although the need for cost allocations arises in a variety of settings, a generic example is a firm composed of a home office and several divisions. Each division produces a single product, using its own resources together with some home office services. In this situation the firm will usually allocate part of its home office costs to each division.

Cost allocations generate substantial revenues for firms working under cost-plus contracts or having significant indirect cost charges built into their fees. Cost allocations also play a central role in public utility rate setting. Even unregulated firms with no opportunities for external cost reimbursement usually have elaborate accounting procedures for allocating costs.

A lively and rather analytical literature in accounting and economics has sprung up around the cost allocation problem. Most of the early accounting texts dealt at length with rules for allocating costs within divisions of the firm or over time (for a review, see Solomons (1968)). Eventually many accounting theoreticians and economists such as Thomas (1974) began to argue that "any allocation of joint costs is arbitrary and serves no useful purpose" (Kaplan 1977, p. 52). Others, though conceding the essential arbitrariness of many cost allocations, have recognized that they are often unavoidable. Using mostly game-theoretic approaches, they have tried to develop allocation procedures that are equitable and consistent with beneficial cooperation among managers (e.g., Shubik 1962 and Hamlen, Hamlen, and Tschirhart 1980). Quite recently, Zimmerman (1979) and Demski (1981) have suggested that cost allocations can help coordinate managers' individual utility functions with the interests of the firm.

Most of the recent literature is concerned with the problem of allocating joint costs, or common costs, which are often somewhat circuitously described as "costs which cannot be readily identified with individual products [or divisions]" (Billera, Heath, and Verrecchia 1981, p. 185). Joint costs comprise the fixed costs of a shared facility and any other costs linked to the activities of two or more divisions. Many accountants feel that "when the costs of producing multiple products are separable, they are readily identified with individual products, so no particular problem arises..." (Billera, Heath, and Verrecchia 1981, p. 185).

However, whenever costs are indirect, whether they are separable or joint, statistical methods are usually needed to implement cost allocation procedures (see for example, Benston (1966), Johnston (1960), or chapters 3 and 14 of Dopuch, Birnberg, and Demski (1974)). These statistical techniques use available cost data, collected either administratively or through sampling, to estimate the cost relationships that form the basis for allocations. While most statistical procedures yield estimates that are substantially free of bias, they often differ markedly in terms of their statistical precision. Methods with poor precision will usually be rejected, especially when cost negotiations are involved.

2. ALLOCATING SEPARABLE COSTS

2.1 Proportional Allocation

Home office costs are said to be separable if we can define the costs incurred by each division. This means that the costs incurred by each division are technically observable, although often sufficiently difficult to trace to make measurement impractical. We assume throughout Section 2 that costs are separable, although only total home office costs are observed. In this case, our problem might be called disaggregation; but to preserve continuity with Section 3 we will continue to speak of allocation.

The most common procedure is to distribute home office costs among several divisions in proportion to an observed variable which is regarded as an index of the service provided by the home office. The allocation variable usually measures the activity within each division, often reflecting a relevant factor of production such as labor hours or wages, or sometimes measuring the division's output or cash flow. Special studies may be undertaken to justify the choice of the allocation variable (e.g., Verrecchia 1981), but once this variable is determined the approach is usually straightforward.

A careful formulation of the proportional allocation procedure will facilitate the discussion of more complex approaches. Consider a particular period i; let y_i denote the costs of the home office to be allocated among several divisions, let x_i denote the total value of the allocation variable across all of these divisions ($x_i > 0$), and let x_{0i} denote the value of the allocation variable within a particular division of interest--say, division 0. Under proportional allocation, the portion of the home office cost allocated to division 0 is $a_i y_i$, where $a_i = x_{0i}/x_i$.

The statistical properties of proportional allocation depend on the nature of the underlying cost relationships. Suppose that y_i is related to x_i in the manner defined by a simple regression model with zero intercept:

$$y_i = \beta x_i + u_i, \qquad (1)$$

$$E(u_i) = 0,$$

$$V(u_i) = \sigma^2.$$

Assume also that the actual unobserved cost incurred by the home office in serving division 0 in period i is separable, is denoted y_{0i}, and is similarly related to x_{0i}:

$$y_{0i} = \beta x_{0i} + u_{0i}, \tag{2}$$

$$E(u_{0i}) = 0,$$

$$V(u_{0i}) = \gamma\sigma^2, \quad 0 < \gamma < 1.$$

Finally, assume that u_{0i} and $u_i - u_{0i}$ are uncorrelated so that

$$C(u_{0i}, u_i) = E(u_{0i}u_i)$$

$$= \gamma\sigma^2.$$

Then

$$a_i y_i - y_{0i} = a_i u_i - u_{0i}, \tag{3}$$

$$E(a_i y_i - y_{0i}) = 0,$$

$$V(a_i y_i - y_{0i}) = (a_i^2 - 2a_i\gamma + \gamma)\sigma^2.$$

Thus, given (1) and (2), the allocated cost is an unbiased predictor of the incurred cost with variance proportional to σ^2.

Three points might be made about proportional allocation. (a) The relationship of the prediction variance to σ^2 in (3) helps to explain why the correlation between y_i and x_i is often considered in selecting an allocation variable. (b) The allocation proportion a_i can be written as $\hat{\beta}x_{0i}/\hat{\beta}x_i$, where $\hat{\beta}$ is any estimator of β. Here the choice of $\hat{\beta}$ and its precision is irrelevant, but this suggests a natural extension of proportional allocation to situations involving multiple allocation variables. (c) The unbiasedness of proportional allocation depends critically on the validity of (2). For example, if the β are actually unequal in (1) and (2), the allocation will be biased.

Proportional allocation can also result in bias if the costs are not separable. Suppose, for example, that (1) and (2) are replaced by

$$y_i = \beta_0 + \beta_1 x_i + u_i,$$
$$y_{0i} = \beta_1 x_{0i} + u_{0i}.$$

Then

$$a_i y_i - y_{0i} = a_i \beta_0 + a_i u_i - u_{0i}.$$

Thus, in this case the allocated cost $a_i y_i$ as a predictor of the incurred cost y_{0i} is biased by the allocated fixed cost $a_i \beta_0$.

In order for an allocation procedure to be accepted as equitable, it should be generally regarded as free of bias. In some applications the possibility of bias can be reduced substantially by utilizing multiple allocation variables, or by using statistical sampling to improve the data base.

2.2 Finite Population Sampling

When proportional allocation is suspected of bias, statistical sampling is sometimes used to directly estimate the cost y_{0i} incurred by the home office in serving division 0. For example, many firms allocate the cost of their capital to divisions. As part of this allocation, statistical sampling is often used to estimate the current value of the inventories held by each division (e.g., Newman 1976). Another example is the use of work sampling to study the activities of employees (e.g., Mandel 1971).

In other circumstances, there may be a particular activity variable that would provide an acceptable basis for proportional allocation but which is observable only through sampling. An example occurs in electric utility load research. Most electric utilities would like to allocate certain capacity-related costs to rate classes in proportion to each class's consumption of electricity during certain hours of peak system-wide demand. A common procedure is to estimate each class's consumption during these peak hours by using special metering equipment on a sample of customers. This application is more fully described in Argonne National Laboratory's Load Research Manual (1980).

Statistical sampling is often regarded as the most objective approach to the allocation of separable costs, but it can be expensive. However, sampling procedures can often be made more efficient by taking advantage of the association between the target variable and other auxiliary variables that are known throughout the population. The model-based survey sampling techniques of Cassel, Sarndal, and Wretman (1976), Kaplan (1973), and Royall (1976) offer a useful link to other regression-based allocation methods (Wright 1981).

2.3 A Prediction Theory Approach

Goldberger's (1962) theory of prediction under the generalized linear model yields some useful insights for the use of several allocation variables. Suppose that we observe an N-dimensional vector $y = (y_1 \cdots y_N)'$ which gives the home office costs in N periods. We also observe two N x k matrices X and X_0. X describes the activities of all divisions and X_0 describes the particular activities of division 0. We continue to assume that costs are separable, and we let y_0 be the N-dimensional vector of unobserved costs incurred by the home office in serving division 0. We generalize (1) and (2) by assuming that there exists an unknown k-dimensional vector β, an unknown scalar $\sigma^2 > 0$, and known N x N matrices Ω_{00}, Ω_{01}, and Ω_{11} such that $(X'\Omega_{11}^{-1}X)^{-1}$ exists and

$$\begin{aligned}
y &= X\beta + u, \qquad (4)\\
y_0 &= X_0\beta + u_0,\\
E(u) &= 0,\\
E(u_0) &= 0,\\
V(u) &= E(uu')\\
&= \Omega_{11}\,\sigma^2,
\end{aligned}$$

$$V(u_0) = \Omega_{00}\,\sigma^2,$$

$$C(u_0, u) = E(u_0 u')$$

$$= \Omega_{01}\,\sigma^2.$$

Here Ω_{01} will usually be nonzero, since u_0 is a component of u.

Imagine momentarily that β is known. Given β, the best linear unbiased (BLU) predictor of y_0 is y_{01}, where

$$y_{01} = E(y_0|u) \tag{5}$$

$$= X_0\beta + \Omega_{01}\Omega_{11}^{-1}u.$$

In the spirit of Shillinglaw (1963), we call y_{01} the <u>costs attributable</u> to division 0, and we call $y_0 - y_{01}$ the <u>unidentifiable costs</u> of division 0. The unidentifiable costs are uncorrelated with y and have variance

$$V(y_0 - y_{01}) = (\Omega_{00} - \Omega_{01}\Omega_{11}^{-1}\Omega_{01}')\sigma^2. \tag{6}$$

When β is unknown, we consider the class of unbiased linear predictors of y_0 or y_{01}, i.e., predictors of the form

$$\hat{y}_0 = Ay$$

for some fixed N x N matrix A, such that

$$E(\hat{y}_0) = E(y_0) \tag{7}$$

$$= E(y_{01}).$$

We call $V(\hat{y}_0 - y_0)$ the <u>prediction variance</u> of $\hat{y}_0$, and we call $V(\hat{y}_0 - y_{01})$ the <u>allocation variance</u>.

Several useful results follow from Goldberger (1962) or Fuller (1980).

(a) The unbiasedness condition (7) is equivalent to the algebraic identity

$$AX = X_0. \tag{8}$$

Thus, if b is any k-dimensional vector,

$$\hat{y}_0 = X_0 b + A(y - Xb), \tag{9}$$

so that $\hat{y}_0$ is the sum of a linear prediction of y_0 based on X_0 and b plus an allocation of the associated total residual cost associated with b.

(b) The prediction variance of $\hat{y}_0$ is

$$\begin{aligned} V(\hat{y}_0 - y_0) &= V(Au - u_0) \\ &= (A\Omega_{11}^{-1}A' - \Omega_{01}A' - A\Omega_{01}' + \Omega_{00})\sigma^2, \end{aligned} \tag{10}$$

and the allocation variance of $\hat{y}_0$ is

$$V(\hat{y}_0 - y_{01}) = V(\hat{y}_0 - y_0) - V(y_0 - y_{01}), \tag{11}$$

where $V(y_0 - y_{01})$ is given by (6) and does not depend on A.

(c) $V(\hat{y}_0 - y_0)$ and $V(\hat{y}_0 - y_{01})$ are both minimized subject to (8) by choosing A equal to A*, where

$$\begin{aligned} A^* &= (X_0 - X_{01})C + \Omega_{01}\Omega_{11}^{-1}, \\ C &= (X'\Omega_{11}^{-1}X)^{-1}X'\Omega_{11}^{-1}, \\ X_{01} &= \Omega_{01}\Omega_{11}^{-1}X. \end{aligned} \tag{12}$$

This gives the best linear unbiased allocation

$$\begin{aligned} \hat{y}_0 &= X_0\hat{\beta} + \Omega_{01}\Omega_{11}^{-1}\hat{u}, \\ \hat{\beta} &= Cy, \\ \hat{u} &= y - X\hat{\beta}. \end{aligned} \tag{13}$$

Also,

$$\hat{y}_0 - y_{01} = (X_0 - X_{01})(\hat{\beta} - \beta),$$

$$V(\hat{y}_0 - y_{01}) = (X_0 - X_{01})V(\hat{\beta})(X_0 - X_{01})'$$

where

$$V(\hat{\beta}) = (X'\Omega_{11}^{-1}X)^{-1}\,\sigma^2.$$

(d) Let $\hat{y}_0$ be defined by (13) and let

$$X_c = X - X_0,$$

$$C(u - u_0,\, u) = \Omega_{c1}\sigma^2,$$

$$\hat{y}_c = X_c\hat{\beta} + \Omega_{c1}\Omega_{11}^{-1}\,\hat{u}.$$

Then the BLU allocation is complete in the sense that $\hat{y}_0 + \hat{y}_c = y$.

Thus, prediction theory yields a seemingly optimal procedure for allocating separable costs under the criteria of unbiasedness and minimum variance. As an example, consider (4) with $k = 1$, $x_i > 0$, $\Omega_{11} = \text{diag}(x_i)$, and $\Omega_{00} = \Omega_{01} = \text{diag}(x_{0i})$. Then the BLU predictor gives proportional allocation

$$A^* = \text{diag}(x_{0i}/x_i),$$

and the allocation variance $V(\hat{y}_0 - y_{01})$ is zero.

Nevertheless, the prediction approach has serious practical limitations. It is unlikely that Ω_{01} will be known or satisfactorily estimable when y_0 is unobserved. Moreover, it is not even clear how to define y_0 and hence (4) when the cost function for y includes joint costs.

3. ALLOCATING JOINT COSTS THROUGH REGRESSION

3.1 Bayesian Prediction

When joint costs are involved, nonstatistical considerations dominate the choice of an allocation policy. However, as pointed out by Dopuch, Birnberg, and Demski (1974, p. 581), "even though allocations of fixed costs are necessarily arbitrary, they may produce reasonably equitable results if the allocation method is based on the relative amount of services consumed by the receiving departments."

As an example, suppose that the home office provides two services in the quantities x_{1i} and x_{2i}, and has the cost function

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}.$$

Also assume that division 0 consumes all of the first service and none of the second service. Then, as a matter of policy the firm might define the costs attributable to division 0 to be $a_i(\beta_1, \beta_2)y_i$, where

$$a_i(\beta_1, \beta_2) = \beta_1 x_{1i}/(\beta_1 x_{1i} + \beta_2 x_{2i}).$$

In this case the costs allocated to division 0 could be calculated as $a_i(\hat{\beta}_1, \hat{\beta}_2)y_i$,where $\hat{\beta}_1$ and $\hat{\beta}_2$ are suitable estimates of β_1 and β_2.

To generalize this example, we assume throughout this section that an allocation policy has been specified in terms of the parameters of a regression model for the cost function of the home office. The firm is assumed to implement the desired allocation by estimating these parameters. Our interest is not in the choice of allocation policy but only in the allocation variance induced by these estimates.

Specifically, we assume a model for the total home office cost y as in (4):

$$y = X\beta + u, \tag{14}$$
$$E(u) = 0,$$
$$V(u) = \Omega_{11}\sigma^2.$$

Although we no longer assume that y_0 is defined, we do assume an allocation policy that defines the costs attributable to division 0--say, y_{01}--in terms of a prespecified N x N matrix $A(\beta)$:

$$y_{01} = A(\beta)y. \tag{15}$$

Here each entry of $A(\beta)$--say, $a_{ij}(\beta)$--is assumed to be a continuously differentiable function of β in the region of interest.

The Bayesian approach to statistical inference (e.g., Zellner 1971) is convenient in discussing inference about y_{01} conditional on the observed y. (A more conventional approach might be developed along the lines of McDonald (1981).) In Bayesian inference, β is regarded as a random variable. Assume for convenience that the prior information is relatively vague, that $X'\Omega_{11}^{-1}X$ is not poorly conditioned, and that N is reasonably large. Then the posterior distribution of β is approximately normal, with mean $\hat{\beta}$ and variance $V(\beta)$, where $\hat{\beta}$ is as in (13) and

$$V(\beta) = (X'\Omega_{11}^{-1}X)^{-1}\,\hat{\sigma}^2, \tag{16}$$

$$\hat{\sigma}^2 = (N - k)^{-1}\,\hat{u}'\Omega_{11}^{-1}\,\hat{u}.$$

The posterior distribution of β induces the posterior distribution of y_{01} through (15). For large N, the posterior distribution of y_{01} can be easily approximated through a Taylor's series expansion of $A(\beta)$ around $\hat{\beta}$. Define $D(\beta)$ to be the N x k matrix with entry $d_{ij}(\beta)$ given by

$$d_{ij}(\beta) = \sum_{h=1}^{N} y_h(\partial a_{ih}(\beta)/\partial\beta_j). \tag{17}$$

Here β_j is the jth entry of the k-dimensional vector β, and y_h is the hth entry of y. Then, in the neighborhood of $\hat{\beta}$,

$$a_{ih}(\beta) \doteq a_{ih}(\hat{\beta}) + \sum_j (\partial a_{ih}(\hat{\beta})/\partial\beta_j)(\beta_j - \hat{\beta}_j), \tag{18}$$

so that

$$y_{01} \doteq A(\hat{\beta})y + D(\hat{\beta})(\beta - \hat{\beta}).$$

Thus, the posterior distribution of y_{01} is approximately normal, with mean $\hat{y}_{01}$ and variance $V(y_{01})$, where

$$\hat{y}_{01} = A(\hat{\beta})y,$$

$$V(y_{01}) = D(\hat{\beta})V(\beta)D(\hat{\beta})'. \tag{19}$$

Here $V(y_{01})$ provides a measure of the precision associated with the allocation $\hat{y}_{01}$.

In many applications, $A(\beta)$ may take a fairly simple form. For each period i, let Q_i and Q_{0i} be two known k-dimensional vectors such that $Q_i'\beta > 0$, and define $A(\beta)$ by

$$a_{ii}(\beta) = Q_{0i}'\beta/Q_i'\beta, \tag{20}$$

$$a_{ij}(\beta) = 0, \quad i \neq j.$$

Then the cost attributable to division 0 in period i can be written as

$$a_{ii}(\beta)y_i = Q_{0i}'\beta + a_{ii}(\beta)[(X_i - Q_i)'\beta + u_i],$$

which generalizes (5). Here y_i and u_i are the ith entries of y and u, and X_i' is the ith row of X. Also, if we let $D_i(\beta)'$ denote the ith row of $D(\beta)$, (17) implies that

$$D_i(\beta) = y_i(Q_i'\beta)^{-1}[Q_{0i} - a_{ii}(\beta)Q_i]. \tag{21}$$

Now suppose that we are interested in the total costs attributable to division 0 throughout all N periods--say, $T = e'y_{01}$--where $e' = (1\ldots1)$. Then (19), (20), and (21) imply that the posterior distribution of T is approximately normal, with mean $\hat{T}$ and standard deviation s_T, where

$$\hat{T} = \sum_{i=1}^{N} (Q_{0i}'\hat{\beta}/Q_i'\hat{\beta})y_i \tag{22}$$

$$= \sum_{i=1}^{N} Q_{0i}'\hat{\beta} + \sum_{i=1}^{N} (Q_{0i}'\hat{\beta}/Q_i'\hat{\beta})[(X_i - Q_i)'\hat{\beta} + \hat{u}_i], \text{ and}$$

$$s_T = \sum_{i=1}^{N} D_i(\hat{\beta})\ V(\beta) \sum_{i=1}^{N} D_i(\hat{\beta})',$$

and where $V(\beta)$ is defined as in (16).

Note that if Q_{0i} is proportional to Q_i--say, $Q_{0i} = c_i Q_i$--then $\hat{T}$ is simply $\sum_{i=1}^{N} c_i y_i$ and $s_T = 0$. Proportional allocation is a special case of this.

3.2 Examples

(a) Assume the separable cost framework (4), and let Q_i' and Q_{0i}' be the ith rows of X and X_0, respectively. Then

$$T = \sum_{i=1}^{N} [E(y_{0i})/E(y_i)]y_i$$

$$= \sum_{i=1}^{N} E(y_{0i}) + \sum_{i=1}^{N} [E(y_{0i})/E(y_i)]u_i,$$

$$\hat{T} = \sum_{i=1}^{N} (X_{0i}'\hat{\beta}/X_i'\hat{\beta})y_i$$

$$= \sum_{i=1}^{N} X_{0i}'\hat{\beta} + \sum_{i=1}^{N} (X_{0i}'\hat{\beta}/X_i'\hat{\beta})\hat{u}_i.$$

Thus, the allocation $\hat{T}$ is proportional with respect to $X_{0i}'\hat{\beta}$, and can be written as the sum of a simple regression-based predicted cost and a proportional allocation of the residual.

(b) Suppose that (14) is a simple linear cost function

$$y_i = \beta_1 + \beta_2 x_i + u_i.$$

Perhaps x_i is the total wages paid in all divisions, and x_{0i} is the total wages paid in division 0. If the firm wished to allocate the fixed cost β_1 in proportion to each division's share of the variable cost, we would take

$$Q_i' = (0 \; x_i),$$

$$Q_{0i}' = (0 \; x_{0i}).$$

Alternatively, the firm might decide to allocate one-half of the fixed costs to division 0, using

$$Q_i' = (1 \ x_i),$$

$$Q_{0i}' = (.5 \ x_{0i}).$$

(c) Suppose that home office costs are assumed to be a quadratic function of the wages x_{1i} and x_{2i} in each of two divisions, 1 and 2:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_{11} x_{1i}^2 + \beta_{22} x_{2i}^2 + \beta_{12} x_{1i} x_{2i} + u_i.$$

Suppose further that we want to allocate home office costs in each period i in proportion to the cost components $\beta_j x_{ji} + \beta_{jj} x_{ji}^2$ uniquely associated with each division j = 1, 2. Then, to allocate costs to division 1, use

$$Q_i' = (0 \ x_{1i} \ x_{2i} \ x_{1i}^2 \ x_{2i}^2 \ 0),$$

$$Q_{0i}' = (0 \ x_{1i} \ 0 \ x_{1i}^2 \ 0 \ 0).$$

In these and similar situations, the statistical precision of the allocated costs can be evaluated using (19) or (22). This Bayesian approach is simpler to implement than the predictive approach of Section 2.3, since there is no need to specify the covariance matrix Ω_{01}. The Bayesian approach can be adapted to a wide variety of models and allocating policies, and seems to capture the spirit of accounting practice.

Although superficially similar, the predictive and Bayesian approaches are conceptually quite distinct. The predictive approach is built around a causal model for the costs actually incurred by each division, so that it is not applicable to the allocation of joint costs. The Bayesian approach is based on an allocation policy expressed in terms of the parameters of the home office cost function; this policy may be more or less arbitrary, but nevertheless it is taken to be prespecified. The Bayesian approach moves the focus from the conventional emphasis on the precision of the estimated regression coefficients to the more relevant precision of the allocations determined by these estimates and the given policy.

4. AN APPLICATION TO INDIRECT COST ACCOUNTING FOR UNIVERSITIES

4.1 Background

In determining costs for federal grants and contracts, educational institutions are currently governed by the cost accounting principles established by the Office of Management and Budget's Circular A-21 (1979). These standards prescribe how indirect costs should be distributed among major functions which usually include instruction, departmental research, and organized (contract) research. The principal indirect cost categories are depreciation and use allowances, general administrative expenses, and operation and maintenance expenses. According to Circular A-21 (1979, p. 12371), "the overall objective of the indirect cost allocation is to distribute these indirect costs to the major functions in proportions reasonably consistent with the nature and extent of their use of the institution's resources."

Consider the case of centrally supplied heating and other utilities. The standard approach is to allocate utility expenses to the individual functions performed in each building in proportion to the square feet of space used by each function. However, the cost distribution methodology is not restricted to proportional allocation. Circular A-21 allows cost determinations to be based on special cost analysis studies which take into account multiple allocation variables together with suitable weighting factors. These cost analysis studies must (a) provide adequate documentation for review, (b) give cost distributions in proportion to benefits, (c) be statistically sound, (d) be performed at the institution itself, and (e) be reviewed periodically. The methodology developed in the previous section seems to comply with the provisions of Circular A-21 for special cost studies.

4.2 A Specific Example

In considering the allocation of its utility expenses, one large university felt that proportional allocation was biased because various types of space incurred utility costs at different rates. A special study was undertaken to determine the feasibility of a cost allocation that differentiated four types of space: nonlaboratory space (x_N), including offices, classrooms, and dormitories; laboratory space (x_L); automobile parking space (x_P); and common space (x_C), including entranceways and halls. The study involved N = 187 buildings in regular use at the university.

A data base which described these 187 buildings was assembled. These data included the total number of square feet of each of the four types of space for each building as well as the amount of each type of space used by the function of interest--say, the organized research function. The study data base also included the utility cost of each building during the 1979/80 fiscal year.

The total utility cost of these 187 buildings was $19,175,000. Excluding structural space, these buildings had a total size of 15,045,000 square feet, of which 54% was nonlaboratory, 11% was laboratory, 11% was parking, and 24% was common. The organized research function utilized 1,755,000 square feet, of which 24% was nonlaboratory, 42% was laboratory, and 34% was common. Virtually no parking was assigned to organized research.

Conventional allocation on a total square foot basis is straightforward. Since organized research utilized 11.7% of the total space, the research function might be assigned $2,243,000 of the total utility cost. Alternatively, the utility cost of each building might be allocated on a total square foot basis. Under this procedure, the total organized research allocation would be increased to $3,890,000. This method still utilizes proportional allocation, and no statistical estimation is involved. The substantial increase in the allocation to organized research in going from a total to a building-by-building procedure was caused by the high utility costs in buildings with research activity. This strengthened the conviction that the various types of space were not homogeneous in their use of utilities.

On the basis of preliminary analysis of the data base, the following regression model seemed to be appropriate:

$$y_i = \beta_N x_{Ni} + \beta_L x_{Li} + \beta_P x_{Pi} + \beta_C x_{Ci} + u_i, \qquad (23)$$

$$E(u_i) = 0,$$

$$V(u_i) = \sigma^2 E(y_i)^2,$$

$$C(u_i, u_j) = 0, \; i \neq j.$$

Because of the heteroscedastic specification, an iterative weighted-least-squares algorithm was used to estimate the parameters of this model (Maddala 1977, p. 261). This yielded the estimated regression equation (with standard errors):

$$\underset{}{\hat{y}_i} = \underset{(0.1066)}{0.8223\, x_{Ni}} + \underset{(0.5428)}{3.1961\, x_{Li}} + \underset{(0.0413)}{0.1253\, x_{Pi}} + \underset{(0.3395)}{2.0050\, x_{Ci}}, \qquad (24)$$

$$\hat{\sigma} = .5987.$$

The estimated coefficients were consistent with the expectation that utility costs varied by type of space. The results showed that utility costs averaged about \$0.82 per square foot of nonlaboratory space and about \$3.20 per square foot of laboratory space. The rather high cost of common space, about \$2.00 per square foot, was surprising but seemed reasonable when the nature of the space was considered. The cost of parking space (\$0.13 per square foot) reflected the cost of lighting.

The university felt that it was equitable to allocate utility costs to each function in proportion to the expected costs incurred by its space based on (23). Since no joint costs were involved, the procedure followed example (a) of Section 3.2. Using (22) and (24), the allocation to organized research was $4,302,000.

However, this approach was questioned because of the poor standard errors in (24), all of which exceed 12% of their respective coefficients. In contrast, (22) gave the Bayesian standard deviation of the allocation as $66,000, i.e., only 1.5% of the allocation. This was regarded as much more satisfactory.

A small simulation experiment was carried out to check this analysis. For simplicity, a sampling orientation was adopted rather than a strict Bayesian view. The space variables were considered to be fixed for the 187 buildings, but the utility costs were randomly generated. For this purpose, (24) was taken to be the true model with the disturbances assumed to be normally distributed. The simulated utility costs were used with the assumed coefficients to calculate T. Then the simulated data were analyzed using the preceding estimation procedure, and $\hat{T}$ and s_T were calculated using (22).

This entire procedure was repeated 100 times to get some indication of the sampling relationship between T and $\hat{T}$, and the validity of s_T. The distribution of the relative allocation error was examined in the 100 trials. The coverage of Bayesian confidence intervals for the allocation was also investigated.

The results of this simulation experiment supported the methodology in every respect. The allocation errors were immaterial, as expected. In the 100 trials, all of the relative allocation errors, $(T - \hat{T})/T$, were less than 4%, and 97% of them were less than 3%. Moreover, there was no indication of any bias in the estimated allocation. In fact, the average relative allocation error was less than 0.3% and was not statistically significant.

The simulation experiment also demonstrated that the standard error s_T gives a realistic estimate of the allocation error. This was determined by calculating the observed coverage of various Bayesian confidence intervals, i.e., the proportion of simulation trials for which T was in the interval $\hat{T} \pm z\ s_T$. The observed coverage was compared to the nominal coverage expected under the standard normal distribution. For $\hat{T} \pm 1\ s_T$, the observed coverage was 70%, slightly exceeding the nominal coverage of 68%. For $\hat{T} \pm 2\ s_T$, the observed coverage was 97%, again slightly above the nominal coverage of 95%. The same pattern occurred for the interval $\hat{T} \pm 3\ s_T$, where the observed coverage was 100%. This suggests that, if anything, s_T may be slightly conservative.

5. SUMMARY

This paper has examined the precision of a cost allocation based on a statistically estimated multiple regression model. Our central model characterizes the total costs incurred by a home office in serving several divisions. When the central model can be expanded to characterize separable divisional costs, linear prediction theory gives a specific allocation procedure that is unbiased and has minimum variance among all linear allocations. Alternatively, if costs are joint or separable with unknown covariances, and if a linear allocation policy can be specified in terms of the parameters of the central model, then Bayesian inference yields measures of the precision of the resulting cost allocations.

REFERENCES

Argonne National Laboratory (1980), Load Research Manual, Argonne, Illinois: Argonne National Laboratory.

Benston, G. J. (1966), "Multiple Regression Analysis of Cost Behavior," The Accounting Review, 41, 657-672.

Billera, L. J., D. C. Heath and R. E. Verrecchia (1981), "A Unique Procedure for Allocating Common Costs from a Production Process," Journal of Accounting Research, 19, 185-196.

Cassel, C. M., C. E. Sarndal and J. H. Wretman (1976), "Some Results on Generalized Difference Estimation and Generalized Regression Estimation for Finite Populations," Biometrika, 63, 615-620.

Demski, J. (1981), "Cost Allocation Games," in Joint Cost Allocations, ed. Shane Moriarity, Norman, Oklahoma: The Center for Economic and Management Research, The University of Oklahoma, 142-173.

Dopuch, N., J. G. Birnberg, and J. Demski (1974), Cost Accounting Data for Management's Decisions, 2d edition, New York: Harcourt, Brace, Jovanovich.

Fuller, W. A. (1980), "The Use of Indicator Variables in Computing Predictions," Journal of Econometrics, 12, 231-243.

Goldberger, A. S. (1962), "Best Linear Unbiased Prediction in the Generalized Linear Regression Model," Journal of the American Statistical Association, 57, 369-375.

Hamlen, S W Hamlen, Jr. and J Tschirhart (1980), "The Use of the Generalized Shapley Allocation in Joint Cost Allocation," Accounting Review, 55, 269-287.

Johnston J. (1960), Statistical Cost Analysis, New York: McGraw-Hill.

Kaplan, R. S. (1973), "A Stochastic Model for Auditing," Journal of Accounting Research, 11, 38-46.

________ (1977), "Application of Quantitative Models in Managerial Accounting: A State of the Art Survey," in Management Accounting--State of the Art, Madison, Wisconsin: University of Wisconsin Press, 30-71.

Maddala, G. S. (1977), Econometrics, New York: McGraw-Hill.

Mandel, B. J. (1971), "Work Sampling in Financial Management--Cost Determination in Post Office Department," Management Science (B), 17, 324-338.

McDonald, G. C. (1981), "Confidence Intervals for Vehicle Emission Deterioration Factors," Technometrics, 23, 239-242.

Newman, M. S. (1976), Financial Accounting Estimates through Statistical Sampling by Computer, New York: Wiley.

Office of Management and Budget (1979), Circular A-21, "Cost Principles for Education Institutions," Federal Register, 44, no. 45, pp. 12368-12380.

Royall, R. M. (1976), "The Linear Least Squares Prediction Approach to Two-Stage Sampling," Journal of the American Statistical Association, 71, 657-664.

Shillinglaw, G. (1963), "The Concept of Attributable Cost," Journal of Accounting Research, 1, 73-85.

Shubik, M. (1962), "Incentives, Decentralized Control, the Assignment of Joint Costs and Internal Pricing," Management Science, 7, 325-343.

Solomons, D. (1968), "The Historical Development of Costing," in Studies in Cost Analysis, 2d edition, ed. David Solomons, Homewood, Illinois: Irwin, 3-49.

Thomas, A. L. (1974), The Allocation Problem: Part Two, Sarasota, Florida: American Accounting Association.

Verrecchia, R. (1981), "A Question of Equity: Use of the Shapley Value to Allocate State and Local Income and Franchise Taxes," in Joint Cost Allocations, ed. Shane Moriarity, Norman, Oklahoma: The Center for Economic and Management Research, The University of Oklahoma, 14-40.

Wright, R. L. (1981), "Electric Utility Load Research Using Model-Based Statistical Sampling," Proceedings of the 1981 DOE Statistical Symposium, Brookhaven National Laboratory, Brookhaven, N.Y.

Zellner, A. (1971), An Introduction to Bayesian Inference in Econometrics, New York: Wiley.

Zimmerman J. (1979), "The Costs and Benefits of Cost Allocations," Accounting Review 54, 504-521.