MANAGEMENT SCIENCE
Vol. 31, No. 5, May 1985
Printed in U.S.A.

BUDGETING COSTS OF NURSING IN A HOSPITAL*

EDWARD P. C. KAO AND MAURICE QUEYRANNE
Quantitative Management Science Department, University of Houston, Houston, Texas 77004
Faculty of Commerce, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

This paper examines issues in building decision support models for budgeting nursing workforce requirements in a hospital. We determine regular-time, overtime, and agency workforce levels for various skill classes in a budget cycle. We introduce a family of eight models ranging from a single-period, aggregate and deterministic model to a multiperiod, disaggregate and probabilistic model. In a single-period model, we ignore the time-varying nature of demand for nursing hours. Aggregation is done over the nurse skill class mix. For probabilistic models, we consider demand uncertainty. Using empirical data, we evaluate the effects of level of sophistication in model building and in information requirements on their relative performances. The results suggest that ignoring the time-varying nature of demand does not induce gross errors in budget estimates. However, ignoring demand uncertainty produces underestimates (about five to six percent) of budget needs—a consequence of a Madansky (1960) inequality. It also induces added costs to the system due to implementing nonoptimal regular-time workforce levels. Finally, we find that a simple formula using a single-period demand estimate gives excellent approximations to the budget estimates obtainable from the more precise models.
(HEALTH CARE—HOSPITALS; PROGRAMMING—LINEAR, APPLICATIONS, PROGRAMMING—STOCHASTIC)

1. Introduction

Nursing constitutes the largest single cost element in most hospitals today. For hospital administrators, preparing yearly nursing workforce budgets remains a vexing problem. The difficulties in defining and measuring quality patient care and nursing productivity, and the probabilistic nature of demand for nursing care all contribute to the complexity embedded in the budgeting process. Kao and Tung (1980) present a survey of literature on the subject of budgeting nursing costs, and give an account of the nursing staffing process as it relates to budgeting. They also present a linear programming based model for assessing need for permanent staff, overtime pay and contracting temporary help by medical service, nursing skill class and time period (month). In this paper, we consider an important extension of the earlier model—namely, the inclusion of demand uncertainty. This extension renders the model more realistic and its results more accurate. We also consider the possibilities of aggregating decision variables and ignoring the time-varying nature of demand. With these modifications in mind, we introduce a family of eight models as possible alternatives for building the decision support system. Each model is characterized by a triplet XYZ. The first letter signifies whether the model is single or multiple period ($X = S$, or M). In a multiperiod model, demands may vary over time. The second letter indicates whether the model is an aggregate or disaggregate one ($Y = A$, or D). Aggregation is over nursing skill classes. The third letter shows whether demands for nursing hours are deterministic or probabilistic ($Z = D$, or P). Let $c(XYZ)$ denote the minimum expected cost associated with the XYZ model.

*Accepted by Warren E. Walker; received January 14, 1982. This paper has been with the authors 11 months for 2 revisions.

0025-1909/85/3105/0608$01.25

§2 introduces a multiperiod, disaggregate, probabilistic (MDP) model, its aggregate counterpart (MAP), and two deterministic specializations (MDD and MAD). There we show that finding the minimum expected cost c(MAP) is straightforward, whereas solving for the exact value of c(MDP) is difficult. We note also that c(MAP) can be used as an upper bound (UB) for c(MDP) and a subgradient approach can be used to establish the corresponding lower bound (LB). We summarize computation procedures for the multiperiod models. For some medical services, the time-varying nature of demand is not pronounced, and demand distributions can be considered stationary. For these cases, it is natural to consider them as single-period problems. In §3, we present four such models. In §4, we describe a procedure for projecting demand distributions using historical data, and present an illustrative example. In §5, we use a set of actual data to demonstrate the applications of the above models for developing estimates of budgets and workforce levels. Finally, we perform a comparative study of the performances of the various models using historical data for nine different medical services of a hospital. Our goal is to assess the impacts of demand uncertainty, decision-variable aggregation, and the demand-stationarity assumption on the budget estimates, and on the actual costs incurred by implementing the optimal regular-time workforce levels suggested by these models. The results suggest that the time-varying nature of demand does not play a prominent role in affecting budget estimates, whereas demand uncertainty, as expected, does. Ignoring demand uncertainty produces about five to six percent underestimates of budget needs and induces added costs to the system due to implementing nonoptimal regular-time workforce levels. To alleviate the computational burden, we find that a simple formula using a single-period demand estimate gives good approximations to budget estimates.

Starting with Dantzig (1955), using a stochastic programming approach for modeling the type of problems under consideration has a long history (e.g., see references given in Dempster 1980, and Stancu-Minasian and Wets 1976). This paper echoes the central theme made in Kallberg, White and Ziemba (1982)—namely, ignoring stochastic components in linear programming formulations can be costly. Aggregation in linear programming has been studied extensively in the past few years (e.g., see Geoffrion 1976a, 1977 and Zipkin 1979, 1980, 1982). This paper and a companion paper (Kao and Queyranne to appear) investigate aggregation in a two-stage stochastic program. While different in content, our approach is developed in the spirit of Zipkin (1979).

2. Multiperiod Models

For a given medical service, let random variable $\mathscr{d}_t$ denote the demand in total nursing hours in period $t = 1, \ldots, T$. The estimation of demand distributions is discussed in §4. Let decision variable R_i denote the number of regular-time nursing hours of skill class i to be allocated to the service in each period, where $i = 1, 2$, and 3 denote RN (registered nurse), LVN (licensed vocational nurse), and NA (nurse aide), respectively. The above definition presumes a constant regular-time workforce over the budget cycle. In addition, the adjustment of regular-time workforce occurs during the budget lead time. A part of the regular-time workforce is taken up by vacations, holidays, and sick leaves. The fraction of the productive regular-time workforce available in period t is p_t and is assumed to be invariant across skill classes.

In meeting fluctuating demands during the budget cycle, we may employ overtime as well as temporary help through nursing manpower agencies. These will be denoted by O_{it} and A_{it}, the number of overtime and the number of agency nursing hours of skill class i allocated to the service for period t. The maximum available overtime workforce is bounded by a factor g times the productive regular-time workforce. The average costs per hour of regular-time, overtime, and agency nurse of skill class i are denoted by $\tilde{r}_i$, o_i, and a_i, respectively. Let $r_i = \tilde{r}_i T$, where r_i gives the corresponding T-period regular-time cost. The cost parameters satisfy

Condition C. (i) $\tilde{r}_i < o_i < a_i$ for all i,

(ii) $\tilde{r}_i < \tilde{r}_{i-1}$, $o_i < o_{i-1}$, and $a_i < a_{i-1}$, $i = 2$ and 3,

(iii) $r_i / \sum p_t \leqslant o_i$ for all i.

The last inequality implies that the T-period regular-time cost even when adjusted by the sum of all productivity factors is still lower than the overtime cost.

The proportions of nursing hours of different skill classes allocated to a given service depend on the nature of the service (e.g., the type of medical specialties it serves, and the level of care, etc.). While the demand for nursing service is expressed in aggregate hours (over all skill classes and sources), the patient care requirement unique to the service is reflected in terms of the constraint that the number of nursing hours of skill class i allocated to the service cannot exceed a factor b_{i-1} times the corresponding figure associated with skill class $i-1$. This corresponds to the allowance of one-way substitution in a hierarchical structure of skill classes.

The MDP Model

At the budget time, the hospital administrator determines the regular-time workforce levels R_i for each skill class and the total payroll budget for the entire budget cycle. As actual demands occur each month, shortages in workforce are met by overtime and agency personnel. The problem can thus be formulated as a *two-stage stochastic program with recourse* (see Dantzig 1955, or Dantzig and Madansky 1961). First-stage decisions deal with the sizing of regular-time workforce and second-stage decisions determine the extent of use of *recourse* in each period—overtime, and if needed, the more expensive external help provided by nursing agencies.

Let Ω denote the sample space for all possible realizations of demand. Let $d_t = d_t(\omega)$, where $\omega \in \Omega$ represents one such sample path. For a given $R = (R_1, R_2, R_3)$, let $R' = \sum R_i$. For period t with a given d_t, the second-stage optimization problem is

$$c_t(R, d_t) \equiv \text{minimize} \quad \sum_i \{ o_i O_{it} + a_i A_{it} \} \qquad \text{subject to:}$$

$$\begin{aligned} \sum_i \{ O_{it} + A_{it} \} &\geqslant d_t - p_t R', \\ -O_{it} &\geqslant -g p_t R_i, \qquad i = 1,2,3, \\ b_{i-1} O_{i-1,t} - O_{it} + b_{i-1} A_{i-1,t} - A_{it} &\geqslant -b_{i-1} p_t R_{i-1} + p_t R_i, \qquad i = 2,3, \\ O_{it}, A_{it} &\geqslant 0 \qquad \text{for all } i. \end{aligned} \tag{1}$$

In (1), the first constraint requires that the total workforce be sufficent to satisfy the demand in each period, the second set of constraints places upper bounds on the available overtime, and the third set stipulates the nursing staff composition requirements. For a given R, the second-stage optimization can be carried out independently for each period. The two-stage stochastic program is given by $c(\text{MDP}) = \min\{ E[c(R)] : R \geqslant 0\}$, where

$$E[c(R)] = \left\{ \sum_i r_i R_i + \sum_t E[c_t(R, \mathscr{d}_t)] \right\} \tag{2}$$

and E denotes the mathematical expectation with respect to $\mathscr{d}_t$. While the model captures the essence of the underlying process, solving the two-stage problem is a nontrivial task. Before we describe a procedure for obtaining upper and lower bounds for the minimum expected cost (2), we first introduce the following simplified variant.

The MAP Model

In practice, aggregating decision variables and/or constraints may be desirable if it does not induce gross errors. The authors (to appear) proposed an approach for aggregating the decision variables in the MDP model. Let R', O_t', and A_t' be the aggregated decision variables over all skill classes. The corresponding (aggregated) coefficients are

$$r' = \sum \mu_i r_i, \qquad o' = \sum \mu_i o_i, \qquad a' = \sum \mu_i a_i, \tag{3}$$

where $\mu_i = \lambda_i/\Lambda$, $i = 1,2,3$, $\Lambda = \sum\lambda_i$, $\lambda_1 = 1$, and $\lambda_{i+1} = b_i\lambda_i$, $i = 1,2$. For the aggregate model, the minimum expected second-stage cost, denoted by $E[c_t'(R',d_t)]$, is

$$E\left[c_t'(R',d_t)\right] = \int_{p_tR'}^{(1+g)p_tR'} o'(d_t - p_tR')\,dF(d_t) + \int_{(1+g)p_tR'}^{\infty}\left[a'd_t - (a'(1+g) - o'g)p_tR'\right]dF(d_t), \tag{4}$$

where F denotes the distribution function of d_t. In (4), the two integrals give the expected (minimum) overtime cost and agency-time cost, respectively. Moreover, $E[c_t(R',d_t)]$ is convex in R'. The problem reduces to $c(\text{MAP}) = \min\{E[c'(R')] : R' \geqslant 0\}$, where

$$E\left[c'(R')\right] = r'R' + \sum_t E\left[c_t'(R',d_t)\right]. \tag{5}$$

Solving the above one-variable convex minimization problem is straightforward (e.g., see Wagner 1975). The formulas for disaggregation are

$$R_i = \mu_i R' \qquad \text{for all } i, \tag{6a}$$

$$O_{it} = \mu_i O_t' \quad \text{and} \quad A_{it} = \mu_i A_t' \qquad \text{for all } i \text{ and } t. \tag{6b}$$

Equation (6a) will be used to disaggregate the regular-time workforce at the budget time, whereas (6b) will be used in each period as the demand for the period unfolds. Under (6a) and (6b), $R_i = b_{i-1}R_{i-1}$, $O_{it} = b_{i-1}O_{i-1,t}$, and $A_{it} = b_{i-1}A_{i-1,t}$ for all i. Moreover, $\bar{r}'$, o', and a' have the interpretations of being the average hourly regular-time, overtime, and agency-time costs, respectively, in the sense that $\sum r_iR_i = r'R'$, $\sum o_iO_{it} = o'O_t'$, and $\sum a_iA_{it} = a'A_t'$.

Approximating the Solution of the MDP Model

Solving the stochastic program of the MDP model *exactly* is difficult in that finding the values of the objective function requires the parametric solution of the second-stage problem and numerical integration. A way to circumvent this difficulty is to establish tight lower and upper bounds of the stochastic objective function, as in Mangasarian and Rosen (1964) and Huang, Ziemba and Ben-Tal (1977). We now summarize a procedure developed in Kao and Queyranne (to appear) for obtaining an upper bound (UB) and a lower bound (LB) for $c(\text{MDP})$. Since the optimal solution for the MAP model is feasible to the MDP model, $\text{UB} = c(\text{MAP})$. To come up with a lower bound, for any R with $R' = \sum R_i$ we first define $P_{1t}(R') = p_t \cdot \text{Prob}\{p_tR' < d_t \leqslant p_t(1+g)R'\}$, $P_{2t}(R') = p_t \cdot \text{Prob}\{d_t > p_t(1+g)R'\}$, and $P_j(R') = \sum_t P_{jt}(R')$, $j = 1$ and 2. For any R satisfying (6a), we call it $\bar{R}$. We note that $E[c(\bar{R})] = E[c'(\bar{R}')]$, i.e., the expected costs of the two models are the same at $\bar{R}$. For such $\bar{R}$, a subgradient vector $\alpha = (\alpha_1,\alpha_2,\alpha_3)$ of the objective function (2) is given by $\alpha_i = r_i - o_iP_1(\bar{R}') - [(1+g)a_i - co_i]P_2(\bar{R}')$, $i = 1,2$, and 3. Define $u^*(\bar{R}') = \min\{\sum_{i=1}^{k}\lambda_i\alpha_i/\Lambda : k = 1,2,3\}$ and let $\alpha' = \sum \mu_i\alpha_i$. (Note that u^* depends on $\bar{R}'$ through α.) Consider the case in which a total of K such $\bar{R}$ have been considered. We let $\bar{R}_{(k)}$ denote the kth such $\bar{R}$, and define

$$C_k = c'\left(\bar{R}'_{(k)}\right) - \alpha' \bar{R}'_{(k)}, \tag{7}$$

$$T_k = u^*\left(\bar{R}'_{(k)}\right). \tag{8}$$

In Kao and Queyranne (to appear), it is shown that a lower bound $\underline{c}(R')$ on $E[c(R')]$ is given by

$$\underline{c}(R') = \max\{C_k + T_k R' : k = 1, \ldots, K\}. \tag{9}$$

Hence a lower bound on $c(\text{MDP})$ is given by $\text{LB} = \min\{\underline{c}(R') : R' \geqslant 0\}$. In §6, we show that, for all cases examined, the differences between UB and LB are relatively small. Thus they provide a good approximation for $c(\text{MDP})$.

One direction of model simplification is to ignore the demand uncertainty in the two multiperiod formulations. This results in the following two models.

The MAD Model

In the multiperiod aggregate deterministic model, we assume that disaggregation is done by (6). Thus for a given d_t, the solution is at least feasible with respect to constraint set (1). Let $c'_t(R', d_t)$ denote the minimum total cost incurred in period t for a given aggregate regular-time workforce level R' and d_t. Define $d'_t = d_t/p_t$, and $d''_t = d_t/((1+g)p_t)$. Then it is easy to verify that

$$c'_t(R', d_t) = \begin{cases} \tilde{r}'R' & \text{if } d'_t \leqslant R', \\ \tilde{r}'R' + o'(d_t - p_t R') & \text{if } d''_t \leqslant R' < d'_t, \\ (\tilde{r}' + o'gp_t)R' + a'\left[d_t - (1+g)p_t R'\right] & \text{if } R' < d''_t, \end{cases} \tag{10}$$

where $\tilde{r}' = r'/T$, and r', o', and a' are defined by (3). For every fixed d_t, *Condition C* suggests that the slope of c'_t is nondecreasing in R', and hence we conclude that c'_t is convex in R'. Since the total cost under R' is $c'(R') = \sum c'_t(R', d_t)$, c' is piecewise linear and convex in R'. Let set $I = \{d'_t, d''_t : t = 1, \ldots, T\}$. I represents the set of (at most) $2T$ breakpoints of the convex function $c'(R')$. Finding $c(\text{MAD})$ reduces to solving $c(\text{MAD}) = \min\{\sum_{t=1}^{T} c'_t(R') : R' \in I\}$. This can easily be done by comparing the total costs of the model with the value of R' set at each one of the $2T$ breakpoints.

The MDD Model

The multiperiod disaggregate deterministic model is a linear program (LP) containing T sets of constraints (1) with R' replaced by $\sum R_i$ (one set for each period) and a deterministic objective function (2). The LP is almost identical to that given in Kao and Tung (1981). For a yearly budget cycle with $T = 12$, the LP has 75 variables and 72 constraints—a problem of modest size. Since the aggregate MAD model gives a feasible solution to the LP, $c(\text{MAD})$ can be used as an upper bound for $c(\text{MDD})$. In the example shown in §5, we see that the simple algebraic procedure for obtaining $c(\text{MAD})$ yields an extremely close approximation for $c(\text{MDD})$. Thus for practical applications, there is no need to set up an LP solution.

3. Single-Period Models

For some medical services, the seasonal fluctuations in patient loads are not as pronounced as others. The time-varying nature of demand could or should (as in some of the uncertainty cases) be ignored. This leads to the single-period models in which we assume that demands (or demand distributions) are stationary and single-period solutions are to be applied to the T periods in the budget cycle. For the deterministic models, let $\bar{d}$ denote the single-period demand. When the demand per period is a random variable $\mathscr{d}$, let F denote its distribution. Let $\bar{p}$ denote the productivity factor per period.

The SAD Model

The most simple and naive model of the eight models is the single-period, aggregate model with a deterministic demand. *Condition C* implies that $R' = \bar{d}/\bar{p}$ and $c(\text{SAD}) = r'R'$.

The SAP Model

An extension of the SAD model is to include demand uncertainty. Equation (4) suggests that the expected single-period cost $E[c_s'(R',\mathscr{d})]$ is

$$E\left[c_s'(R',\mathscr{d})\right] = \tilde{r}'R' + \int_{\bar{p}R'}^{(1+g)\bar{p}R'} o'(\mathscr{d} - \bar{p}R')\,dF(\mathscr{d}) + \int_{(1+g)\bar{p}R'}^{\infty}\left[a'\mathscr{d} - (a'(1+g) - o'g)\,\bar{p}R'\right]dF(\mathscr{d}). \tag{11}$$

Again, $E[c_s'(R',\mathscr{d})]$ is convex in R'. Solving (11) is straightforward, and $c(\text{SAP}) = T \times \min\{E[c_s'(R',\mathscr{d})]: R \geqslant 0\}$. To simplify the computation further, we rewrite (11) as

$$E\left[c_s'(R',\mathscr{d})\right] = \tilde{r}'R' + \int_{\bar{p}R'}^{\infty} o'(\mathscr{d} - \bar{p}R')\,dF(\mathscr{d}) + (a' - o')\int_{(1+g)\bar{p}R'}^{\infty}\left[\mathscr{d} - (1+g)\bar{p}R'\right]dF(\mathscr{d}). \tag{12}$$

In (12), the last integral represents the expected shortage in nursing hours to be made up from using agency workforce. This term is generally expected to be small and can be ignored. This results in a "newsboy" type of solution (Johnson and Montgomery 1974)

$$F(\bar{p}R') = (o' - \tilde{r}')/o'. \tag{13}$$

Solving (13) only requires the use of the distribution of the one-period demand.

The SDD Model

Let O_i and A_i denote the single-period overtime and agency nursing hours of skill class i to be allocated to the service, respectively. The single-period disaggregate model with a certain demand is given by the following LP:

$$\text{minimize} \quad \sum(\tilde{r}_i R_i + o_i O_i + a_i A_i) \qquad \text{subject to:}$$

$$\sum(\bar{p}R_i + O_i + A_i) \geqslant \bar{d}$$

$$O_i \leqslant g\bar{p}R_i, \qquad i = 1,2,3,$$

$$\bar{p}R_i + O_i + A_i \leqslant b_{i-1}(\bar{p}R_{i-1} + O_{i-1} + A_{i-1}), \qquad i = 1,2,3, \quad \text{and}$$

$$R_i, O_i, A_i \geqslant 0 \qquad \text{for all } i.$$

Again, *Condition C* implies that in an optimal solution $O_i = A_i = 0$ for all i, and consequently, $\sum R_i = \bar{d}/\bar{p} = R'$ and $R_i = \mu_i R'$. Thus we have $c(\text{SDD}) = r'R'$ and the solution is identical to that of the SAD model.

The SDP Model

This model is a one-period specialization of the MDP model. The problem seems nearly as difficult as the MDD model. However, c(SAP) could be used as an upper bound for c(SDP), and a lower bound can be derived using the same subgradient technique as for the MDP model.

4. Demand Forecasts

Demand Forecast for Multiperiod Models

For the multiperiod models, assume that each period corresponds to a month in a yearly budget cycle, so $T = 12$. As in Abernathy *et al.* (1973), we approximate the distribution of $\mathscr{d}_t$ by a normal distribution $\mathscr{N}(m_t, \sigma_t^2)$. To estimate m_t and σ_t^2, let γ_t = the average number of daily admissions in month t, $\sigma_{N,t}^2$ = the variance of daily admissions in month t, W = the average length of stay per admission, σ_X^2 = the variance of length of stay per admission, and L_t = the patient-days generated by all patients admitted on a given day in month t. Then we see that $E[L_t] = \gamma_t W$, and $\mathrm{Var}[L_t] = \sigma_{N,t}^2 W^2 + \gamma_t \sigma_X^2$.

To convert patient-days to nursing-hour requirements, we use the conversion factor e given in Table 3 of Kao and Tung (1981); namely, eL_t gives the nursing hours needed for serving L_t patient-days. Assume L_t are independent random variables for each day in month t, then $m_t = N_t e E[L_t]$ and $\sigma_t^2 = N_t e^2 \mathrm{Var}[L_t]$, where N_t denotes the number of days in month t. In Kao and Tung (1980), a forecasting system using autoregressive integrated moving average (ARIMA) time series models forms the basis for projecting average daily admissions by month. To find an estimator for $\sigma_{N,t}^2$, we choose to equate it with the variance of forecast error, i.e., we let $\sigma_{N,t}^2 = \sigma_\epsilon^2[1 + \sum_{j=1}^{t-1} \Psi_j^2]$, $t = 1, \ldots, 12$, where Ψ_j's are the weights obtained as a by-product of forecasting endeavor (e.g., see Box and Jenkins 1976). For the two multiperiod deterministic models, we use m_t as our estimates for d_t.

We now illustrate the use of the above approach for finding m_t and σ_t^2. In Kao and Tung (1980), using the 1973–1977 data, an ARIMA $(0,0,1) \times (0,1,1)_{12}$ model was established for forecasting monthly admission rates γ_t of surgical patients for 1978 (see Table 1). In addition, we find that $e = 4.96$, $W = 6.75$, $\sigma_X^2 = 299.87$, $\sigma_\epsilon^2 = 1.215$, $\Psi_1 = 0.3206$, and $\Psi_j = 0$ for $j > 1$. Based on these figures, we can compute m_t and σ_t for 1978 shown in Table 1. There we also give direct sample estimates of $\hat{\sigma}_t$ based on 1973–1977 data (i.e., $\hat{\sigma}_t^2 = \sum_{i=1973}^{1977}(d_{it} - \bar{d}_t)^2/4$, where d_{it} = actual demand in nursing hours for month t in year i, and $\bar{d}_t = \sum_{i=1973}^{1977} d_{it}/5$), and the actual d_t for 1978.

TABLE 1

Demand Forecast for Surgical Patients in 1978.†

t	γ_t	m_t	d_t	σ_t	$\hat{\sigma}_t$	p_t
1	11.5376	11975	12385	1637	1418	0.8943
2	12.5239	11740	12335	1621	1315	0.8917
3	11.7249	12169	14512	1652	1535	0.8948
4	13.0746	13132	11363	1714	1537	0.9086
5	13.0313	13525	13644	1740	1077	0.9032
6	12.5431	12598	12657	1680	1575	0.8842
7	13.0105	13503	14448	1738	756	0.8513
8	13.6510	14168	12533	1780	3000	0.8798
9	12.5465	12602	12414	1680	2094	0.8871
10	11.3756	11807	12161	1627	988	0.9043
11	11.2851	11335	11477	1595	1261	0.8606
12	10.0296	10410	13540	1530	1319	0.8341

†Input data based on the study reported in Kao and Tung (1980, 1981).

Demand Forecast for Single-Period Models

For single-period models, one may employ a simple exponential smoothing method for forecasting and use yearly averages per period as inputs. To generate estimates for use in the next section, we average the multiperiod forecasts instead. This gives $\bar{p} = 0.8828$, $\bar{d} = 12414$, and $\sigma = 1666$.

5. An Example

Using demand forecasts obtained earlier, we now illustrate the development of budget estimates for 1978 for the surgical service of the hospital studied in Kao and Tung (1980, 1981). The hospital, supported primarily by tax revenues, serves the medically indigent residing in a growing metropolis in the Sunbelt. From Kao and Tung (1981) we find that $\{\bar{r}_i\} = \{7.03, 4.53, 3.44\}$, $\{o_i\} = \{9.59, 6.18, 4.69\}$, $\{a_i\} = \{11.70, 9.95, 5.78\}$, and $g = 0.2$. With $b_1 = 0.6$ and $b_2 = 2.0$, we have $\lambda_1 = 1$, $\lambda_2 = 0.6$, $\lambda_3 = 1.2$, and $\Lambda = 2.8$, and consequently, $\mu_1 = 0.3571$, $\mu_2 = 0.2143$, and $\mu_3 = 0.4286$. Using (3), we obtain the aggregate cost factors $r' = 59.4669$, $o' = 6.7591$, and $a' = 8.7877$.

The MAD Model

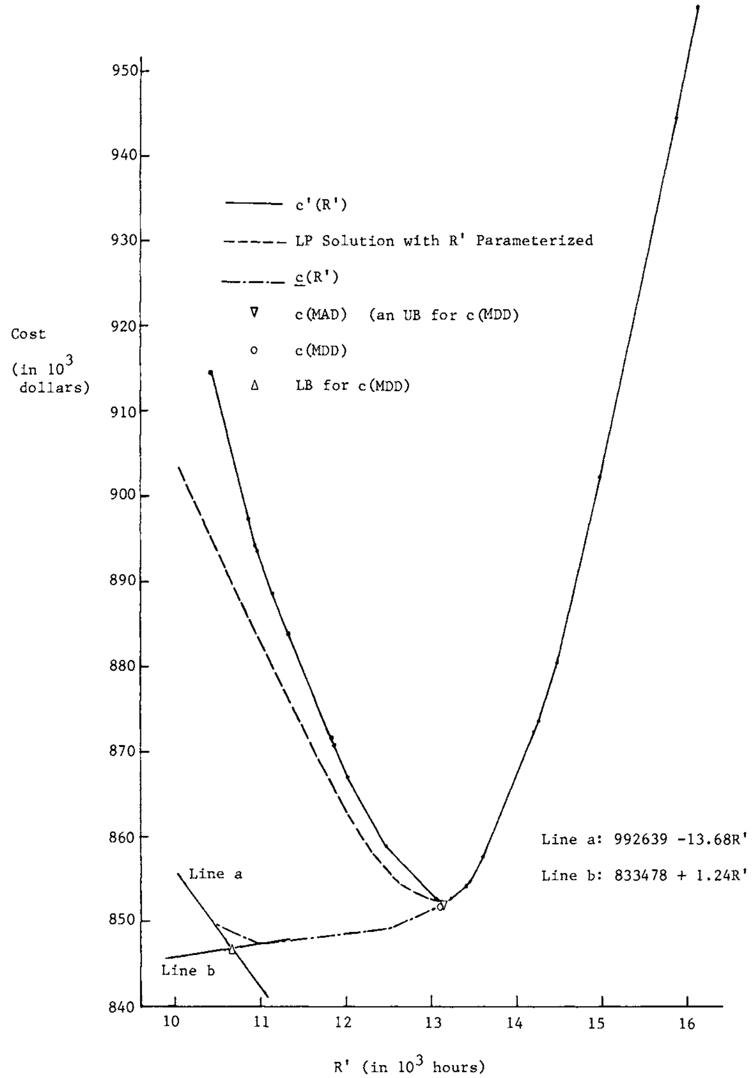
Using the d_t given in Table 1, we first compute the set I comprising the 24 breakpoints. With R' set at each one of these breakpoints, we compute the single period cost $c_t'(R', d_t)$ using (10). For each such R', summing c_t' over all t yields $c'(R')$. In Figure 1, we plot $c'(R')$ for all $R' \in I$, and connect these points to form a piecewise linear (convex) function. For this model, $c(\text{MAD}) = \$852{,}250$ with $R' = 13166$. Using the disaggregation factors, we obtain $R = (4702, 2821, 5643)$.

The MDD Model

Solving the LP, we find $c(\text{MDD}) = \$852{,}214$ and $R = (4718, 2831, 5617)$. It is interesting to note that (i) under this solution $R' = 13166$, which is identical to that obtained under the aggregate model, (ii) the optimal regular-time workforce mix by skill class is similar to that generated by applying the disaggregation factors, and (iii) the minimum cost $c(\text{MDD})$ is only \$36 lower than its upper bound $c(\text{MAD})$. This seems to suggest that for deterministic demands the simple algebraic method can give good approximations. In the LP solution, the agency time is used only to meet extreme demand surges occurring in July and August—for certainty models, this is to be expected.

Computing the LB for $c(\text{MDD})$ does not seem to have any practical value due to the ease with which one can solve the LP directly. However, we include it here for illustrating the procedure which will eventually be applied to the uncertainty models. When demands are deterministic, $P_{1t}(R')$ and $P_{2t}(R')$ are simple constants—either p_t or 0. By considering $R' \in I$, we construct the upper convex envelope formed by $\underline{c}(R')$ shown in Figure 1. An approximate lower bound on $c(\text{MDD})$ is given by LB $= \min\{\underline{c}(R') : R' \in I\} = \$847{,}010$. A more precise LP occurs where $992639 - 13.68R'$ intersects $833478 + 1.24R'$ (the intercepts and slopes of the two lines correspond to those computed at $\bar{R}' = 12481$ and 13166, respectively). Equating the two lines (cf. Figure 1) yields LB $= \$846{,}706$ with $R' = 10668$. The *relative gap*, $(\text{UB} - \text{LB})/\text{LB} \times 100\%$, is 0.65%.

In Figure 1, we also plot the optimal value of the objective function as a function of R' (with $R' = \sum R_i$) using the parametric LP. We see that the difference in minimum total cost between the parametric LP solution and the aggregate solution diminishes as R' increases and becomes negligible when $R' > 13166$. This implies that only at lower levels of regular-time workforce an LP solution can do better than the aggregate



15

FIGURE 1. Cost Computations for the MAD and MDD Models.

solution. The saving achieved by an LP solution (over an aggregate solution) is attributable to the room available for manipulating the overtime and agency workforce over skill classes.

The MAP Model

Evaluating (4) can be carried out either by numerical integration or through the use of the normal loss integral as suggested in Kao and Queyranne (to appear). The minimum cost solution occurs at $R' = 12708$ with $c(\text{MAP}) = \$885{,}874$. Similar to (4), we can compute the second moment of the one-period cost by numerical integration and estimate the variance of yearly cost $\text{Var}[c'(R')]$. If we plot the standard deviation of yearly cost as a function of R', we see that the variation in cost decreases rapidly as R' increases. At optimality with $R' = 12708$, if we use the two-standard-deviation criterion to establish a confidence interval for the actual cost in 1978, we have

885,874 ± (2)(35,258), or approximately, $c'(R') \in [815,358; 956,390]$. In practice, this confidence interval estimate will perhaps be more useful.

The MDP Model

For the range of R' shown in Figure 1, we use 200 equally spaced points $\bar{R}'_{(k)}$ to generate C_k and T_k, and construct the upper convex envelope formed by $\underline{c}(R')$. (The choice of K is arbitrary; a choice of 200 points is extravagant but almost costless.) Let $\text{LB} = \min\{\underline{c}(R') : \bar{R}'_{(k)}, k = 1, 2, \ldots, 200\}$, we obtain LB = \$877,050 at $R' = 10106$. (We could fine-tune our estimate of LB using the line intercept approach shown earlier. This would yield LB = \$877,024—a difference of little significance. However if we had used fewer points, then the improvement could have been somewhat more.)

The SAD and SDD Models

In an optimal solution, both models call for $R' = \bar{d}/\bar{p} = 14061$ and $c(\text{SAD}) = c(\text{SDD}) = \$836,195$.

The SAP Model

For a normal demand distribution, equation (12) can be reduced to

$$E\left[c_s'(R', \mathscr{d})\right] = \bar{r}'R' + a'\bar{d} + \left[o'g - a'(1+g)\right]\bar{p}R' + o'\sigma\phi\left[z_1(R')\right] + (a' - o')\sigma\phi\left[Z_2(R')\right] + o'W_1(R')\Phi\left[Z_1(R')\right] + (a' - o')w_2(R')\Phi\left[Z_2(R')\right], \tag{14}$$

where $W_1(R') = \bar{p}R' - \bar{d}$, $W_2(R') = (1+g)\bar{p}R'$, $Z_1(R') = W_1(R')/\sigma$, $Z_2(R') = w_2(R')/\sigma$, and ϕ and Φ denote the standardized normal density and distribution functions, respectively. Since $E[c'(R', \mathscr{d})]$ is convex in R', applying a search algorithm yields $c(\text{SAP}) = \$877,810$ and $R' = 12825$. Turning to the approximation for the SAP model, (13) reduces to $\Phi(\bar{p}R') = 0.2668$ and

$$R' = (\bar{d} - 0.622\sigma)/\bar{p}. \tag{15}$$

Thus we find $R' = 12888$ with $c(\text{SAP}) \simeq \$877,844$.

Finally, we remark that since the computational simplicity expected of a single-period model is not shared with the SDP model, we exclude it from our illustration.

6. A Comparative Study

We now consider the applications of the budgeting models to the hospital studied earlier in Kao and Tung (1981). The objective is to gain insights about the various issues involved in model building (cf. Geoffrion 1976b)—especially with regard to the roles played by the aggregation of decision variables, the uncertainty, and the time-varying nature of demands. The hospital under study has nine different medical services each with its own parameter sets and each prepares its own budget. The conclusions drawn from this sample of nine data sets, while by no means definitive, are indicative of what are to be expected from cases under similar settings. Based on the 1973–1977 data, the forecasted means for the monthly demands in nursing hours for 1978 and the coefficients b_1 and b_2 for each department are given in Kao and Tung (1981). The actual monthly demands and the estimated standard deviations of demands for 1978 are given in Kao and Queyranne (1982). Along with the cost coefficients and the coefficients g and p_t given in the last two sections, they constitute the complete data set.

TABLE 2

Bounds for c(MDP)

Medical Service	R' Optimizing c(MAP)	(I) c(MAP) [UB for c(MDP)]	$\sigma[c'(R')]^{\dagger}$	R' Yielding LB	(II) LB for c(MDP)	Relative Gap 100% × (I − II)/II
GYN	3453	206844	6835	2710	202103	2.35
MED	11682	809456	33594	9688	800978	1.06
NEU	2199	160182	7895	1773	158772	0.89
NRS	5327	456625	17508	5327	456625	0.
OPH	5833	390303	13637	4585	386712	0.93
PBT	9531	668733	23886	7683	664341	0.66
PSY	3334	235692	12551	2458	233689	0.86
SUR	12708	885874	35204	10106	877050	1.01
PLS	2528	188882	13520	1996	185986	1.56
Summary	56595	4002591		46326	3966256	0.92

†Using the R' that minimizes c(MAP).

The Bounding Procedure for c(MDP)

Since the MDP model is the most detailed (and most accurate) model of the eight and finding c(MDP) is difficult, we empirically evaluate the adequacy of the bounding procedure introduced in §2. First we solve the aggregate MAP models. The minimum costs c(MAP), the minimizing R', and the respective standard deviations $\sigma E[c'(R')]$ are tabulated in Table 2. Following the procedure presented earlier, we find the LB's and the minimizing R' shown in the table. Since the relative gap represents the maximum relative reduction in the expected cost possible by *not* following the (proportional) disaggregation formulas (6), the figures shown in the table suggest that the maximum amounts of improvement possible by considering a disaggregate model are in most cases in the ballpark of one percent. For the case in which the largest relative gap of 2.35% occurs, it is interesting to note that the service has the largest b_1 factor and the largest coefficient of variation (CV = 10%) of mean demands (specifically, with $s = \sum_{t=1}^{12}(m_t - \overline{m})^2/11$ and $\overline{m} = \sum_{t=1}^{12} m_t/12$, we have $\mathrm{CV} = s/\overline{m}$). The larger the b_i's the more flexible it becomes to manipulate the skill class mix. The larger the CV the more room a decision maker will have in making such manipulations. Following these observations, we also see that the relative gap associated with NRS is zero. There b_1 is very small, b_2 is zero, and CV is only 3%.

Evaluating the MAP Model

The above discussion suggests that c(MAP) is adequate for estimating the budget need. This estimate is on the conservative side in the sense that c(MAP) will always be an overestimate of c(MDP). For decision making under uncertainty, a good decision will not necessarily yield a good outcome—especially when outcomes hinge on the realization of *one* sequence of random phenomena. Nevertheless, to satisfy our curiosity we used the 1978 actual demands to compute the various cost figures. First we obtained the c(MAD) shown in Table 3, the minimum costs obtainable if we know the demands for certain. Comparing these figures with the recommended budgets c(MAP), we have a total underestimate of 6.39% (cf. Table 3). This is an *error in budget estimate*. In actuality, the *error in total cost* resulting from actually implementing the R' minimizing c(MAP) is given by $c(\mathrm{MAP}) - c'(R')$, where $c'(R')$ represents the total cost under R' associated with the MAD model when the actual demands for 1978 are in effect. In Column II of Table 3, we tabulate these $c'(R')$. Column IV of the table shows that if we propose a budget based on c(MAP) and implement the

TABLE 3

Expected Costs Under Actual Demands

Medical Service	(I) c(MAD) Use Actual Demands[§]	(II) $c'(R')$[†] Use Actual Demands[§]	(III) (c(MAP)* − I)/I × 100%	(IV) (c(MAP)* − II)/II × 100%	(V) (II − I)/I × 100%
GYN	172599	187044	19.84	10.59	8.37
MED	811360	812187	− 0.23	− 0.34	0.10
NEU	135760	140458	17.99	14.04	3.46
NRS	544910	550688	− 16.20	− 17.08	1.06
OPH	528985	580091	− 26.22	− 32.72	9.66
PBT	668875	672726	− 0.02	− 0.59	0.58
PSY	303826	326700	− 22.43	− 27.86	7.53
SUR	882253	890709	0.41	− 0.54	0.96
PLS	227087	241850	− 16.82	− 21.90	6.50
Summary	4275655	4402453	− 6.39	− 9.08	2.97

[§]These are the only two cases in which we use the actual 1978 demands shown in Kao and Queyranne (1982) for computing the expected costs. In all other cases, we use the forecasted demands for 1978.

[†]This represents the total cost if we implement the optimal R' obtained from the MAP model while the actual demands for 1978 are in effect.

*From Column (I) of Table 2.

corresponding R', then we will actually get a nine percent underestimate of total cost. This gross underestimate of cost is caused primarily by a substantial underestimate of demands incurred in 1978. For the two largest services (MED and SUR), the forecasts are comparatively accurate, and there the errors both in budgets and actual costs are minimal.

Comparing Columns I and II of Table 3, we find that, if the R' minimizing c(MAP) are in use, then the overall *error in actuality* is only 2.97% (Column V). This is much smaller than the other two types of error considered earlier.

Assessing Other Aggregate Models

We now use the MAP model as the benchmark and assess the adequacy of using the other three aggregate models. The forecasted demands for 1978 are the inputs. Taking the MAD model as an example, we define two criteria for evaluation. The first, the *percent of nominal error* (PNE), is defined as

$$\text{PNE} = 100\% \times (c(\text{MAD}) - c(\text{MAP}))/c(\text{MAP}).$$

This measures the relative error in the proposed budget caused by ignoring the uncertainty of demand. Recall $E[c'(R')]$ denotes the expected cost when R' is implemented in the MAP model. Whichever aggregate model one uses, when the corresponding (minimizing) R' is implemented a more accurate estimate of expected cost is then given by $E[c'(R')]$. This induces a different type of error called the *percent of actual error* (PAE), where $\text{PAE} = 100\% \times (E[c'(R')] - c(\text{MAP}))/c(\text{MAP})$. In Table 4, we give the minimizing R', the expected minimal costs, the expected actual costs $E(c'(R')]$, and the PNE's and PAE's for the other three aggregate models.

The figures in the table indicate that the two deterministic variants are both inadequate in that they tend to give lower budget estimates while in actuality they cost a little more than those produced by c(MAP). On the other hand, the results of SAP models suggest that ignoring the time-varying nature of demand is acceptable so long as one takes demand uncertainty into account. Finally, the figures in Table 5 suggest that the simple formula (15) for approximating R' for the SAP model performs well.

TABLE 4

Assessing the Relative Performances of Aggregate Models

	Medical Service	Optimal R'	c(MAD)	$E[c'(R')]$	PNE	PAE
The MAD Model	GYN	3525	201979	207009	− 2.35	0.08
	MED	12472	773299	815230	− 4.47	0.71
	NEU	2343	151445	161091	− 5.45	0.57
	NRS	5309	442146	456631	− 3.17	0.00
	OPH	6154	377443	392442	− 3.29	0.55
	PBT	10003	643929	671651	− 3.71	0.44
	PSY	3495	221286	236266	− 6.11	0.24
	SUR	13166	852250	887557	− 3.80	0.19
	PLS	2760	170716	190235	− 9.62	0.72
	SUMMARY		3834493	4018116	−4.20	0.39
	Medical Service	Optimal R'	c(SAP)	$E[c'(R')]$	PNE	PAE
The SAP Model	GYN	3455	203630	206846	− 1.55	0.00
	MED	11708	807835	809459	− 0.20	0.00
	NEU	2200	159833	160182	− 0.22	0.00
	NRS	5340	455311	456634	− 0.29	0.00
	OPH	5850	388979	390312	− 0.34	0.00
	PBT	9555	666406	668747	− 0.35	0.00
	PSY	3340	234551	235691	− 0.48	− 0.00
	SUR	12825	877810	885978	− 0.91	0.01
	PLS	2530	188611	188881	− 0.14	− 0.00
	SUMMARY		3982966	4002733	−0.49	0.00
	Medical Serivce	Optimal R'	c(SAD)	$E[c'(R')]$	PNE	PAE
The SAD Model	GYN	3740	196005	209382	− 5.24	1.23
	MED	12830	765810	821714	− 5.39	1.51
	NEU	2445	149462	162810	− 6.69	1.64
	NRS	5838	434111	463683	− 4.93	1.55
	OPH	6376	373034	396351	− 4.42	1.55
	PBT	10443	638382	679423	− 4.54	1.60
	PSY	3751	217792	239692	− 7.59	1.70
	SUR	14061	836195	900724	− 5.61	1.68
	PLS	2880	168481	191996	− 10.80	1.65
	SUMMARY		3779272	4065779	−5.58	1.58

TABLE 5

*Evaluating the Approximation Formula for Finding the Optimal R' and the SAP Model**

Medical Service	(I) Approximate R'	(II) Exp. Cost under R' $E[c_s'(R',\mathscr{A})]$	(Opt. R' − I)/ Opt R' × 100%	(II − c(SAP))/ c(SAP) × 100%
GYN	3485	203661	0.87	0.02
MED	11692	807838	− 0.14	0.00
NEU	2183	159849	− 0.77	0.01
NRS	5360	455326	0.37	0.00
OPH	5902	389043	0.89	0.02
PBT	9648	666512	0.97	0.02
PSY	3309	234577	− 0.93	0.01
SUR	12888	877844	0.49	0.00
PLS	2412	188961	− 4.66	0.19
	SUMMARY	3983611		0.02

*The optimal R' and c(SAP) are shown in Table 4.

The approximate R' are all very close to the optimizing R' (with perhaps the exception of PLS).[1]

[1]This research was supported by a grant from the Center for Health Management, University of Houston. The authors are grateful to Philip Patrick Sun, Executive Director of Planning, the Harris County Hospital District, Houston, Texas, for his continued support and assistance for the study.

References

ABERNATHY, W. J., N. BALOFF, J. HERSHEY AND S. WANDEL, "A Three-Stage Manpower Planning and Scheduling Model—A Service-Sector Example," *Oper. Res.*, 21 (1973), 693–711.

BOX, G. E. P. AND G. M. JENKINS, *Time Series Analysis: Forecasting and Control*, (revised ed.), Holden-Day, San Francisco, 1976.

DANTZIG, G. B., "Linear Programming under Uncertainty," *Management Sci.*, 1 (1955), 197–206.

——— AND A. MADANSKY, "On the Solution of Two-Stage Linear Programs under Uncertainty," *Proc. Fourth Berkeley Sympos. on Math. Statist. and Probab.*, Vol. I, J. Neyman, (Ed.), University of California, Berkeley, 1961, 165–176.

DEMPSTER, M. A. H., (Ed.), *Stochastic Programming*, Academic Press, New York, 1980.

GEOFFRION, A. M., "Customer Aggregation in Distribution Modelling," Working Paper No. 259, Western Management Science Institute, University of California at Los Angeles, 1976a.

———, "The Purpose of Mathematical Programming is Insight, Not Numbers," *Interfaces*, 7 (1976b), 81–92.

———, "A Prior Error Bounds for Procurement Commodity Aggregation in Logistics Planning Models," *Naval. Res. Logist. Quart.*, 24 (1977), 201–212.

HUANG, C. C., W. T. ZIEMBA AND A. BEN-TAL, "Bounds on the Expectation of a Convex Function of a Random Variable: With Applications to Stochastic Programming," *Oper. Res.*, 25 (1977), 315–325.

JOHNSON, L. A. AND D. C. MONTGOMERY, *Operations Research in Production Planning, Scheduling, and Inventory Control*, Wiley, New York, 1974.

KALLBERG, J. G., R. W. WHITE AND W. T. ZIEMBA, "Short Term Financial Planning Under Uncertainty," *Management Sci.*, 28 (1982), 670–682.

KAO, E. P. C. AND G. G. TUNG, "Forecasting Demands for Inpatient Services in a Large Public Health Care Delivery System," *Socio-Econom. Planning Sci.*, 14 (1980), 97–106.

——— AND ———, "Aggregate Nursing Requirement Planning in a Public Health Care Deliver System," *Socio-Econom. Planning Sci.*, 15 (1981), 119–127.

——— AND M. QUEYRANNE, "Aggregation in a Two-Stage Stochastic Program for Manpower Planning in the Service Sector," to appear in *TIMS Studies in the Management Sciences—Delivery of Urban Services*, A. J. Swersey, (Ed.).

——— AND ———, "Budgeting Costs of Nursing in a Hospital," Working Paper, Department of Quantitative Management Science, University of Houston, 1982.

MADANSKY, A., "Inequalities for Stochastic Linear Programming Problem," *Management Sci.*, 6 (1960), 197–204.

MANGASARIAN, O. L. AND J. B. ROSEN, "Inequalities of Stochastic Nonlinear Programming Problems," *Oper. Res.*, 12 (1964), 143–154.

STANCU-MINASIAN, I. M. AND M. J. WETS, "A Research Bibliography in Stochastic Programming," *Oper. Res.*, 24 (1976), 1008–1019.

WAGNER, H. M., *Principles of Operations Research*, (second ed.), Prentice-Hall, Englewood Cliffs, N.J., 1975.

ZIPKIN, P. H., "Transportation Problems with Aggregated Destinations when Demands are Uncertain," Research Working Paper No. 275A, Graduate School of Business, Columbia University, 1979.

———, "Bounds on the Effect of Aggregating Variables in Linear Programs," *Oper. Res.*, 28 (1980), 403–418.

———, "Exact and Approximate Cost Functions for Product Aggregates," *Management Sci.*, 28 (1982), 1002–1012.