The Cyclical Behavior of Hospital Utilization and Staffing

Gail Jensen and Richard Kronick

Aggregate monthly data on hospital utilization and staffing are examined to assess the hospital industry's ability to adjust staffing levels to regular monthly cycles in demand. Graphical analysis and linear regression are used to assess the relationship between monthly trends in utilization and full-time-equivalent hospital personnel. We show that although regular seasonal patterns exist in both utilization and staffing levels, these series are largely independent of each other. The staffing level response to cycles in admissions and patient-days is, in fact, small relative to those observed for other industries that face predictable and regular fluctuations in product demand. Staffing levels appear to be more closely related to bed levels than to actual utilization levels. For a typical hospital which does not face effective incentives to control costs, smoother patterns of seasonal utilization probably will not result in lower staffing levels and reduced costs unless accompanied by a slowdown in the rate of increase in hospital bed size.

The utilization of hospital services fluctuates over time in daily, monthly, annual, and perhaps even longer cycles. Much of the short-term fluctuation in nonelective emergent utilization is of a random nature—on some days many people are in accidents, have heart attacks, or need an appendix removed; on other days, not as many. The need to guard against random surges in demand means that hospi-

---

The study on which this article is based was performed pursuant to Contract Number 232-79-0076 with the Bureau of Health Professions, Department of Health and Human Services. Contractor for the study was Abt Associates Inc., Cambridge, MA.

An earlier draft was presented at the 1980 annual conference of the Western Economic Association, San Diego, CA.

Gail Jensen is a Ph.D. student in the Department of Economics, the University of Minnesota, and a Research Analyst at the University's Center for Health Services Research. Address all correspondence and requests for reprints to her at Department of Economics, 1035 Management and Economics, University of Minnesota, Minneapolis, MN 55455. Co-author Richard Kronick is a Senior Analyst at Abt Associates Inc., Cambridge, MA.

tals cannot operate safely at 100 percent of capacity but must have a cushion of empty beds. Recent proposals in the areas of hospital planning [1,2], such as admissions scheduling with call-ins and discharge planning, allow hospitals facing chronic excess demand to operate safely at higher-capacity levels by deliberating offsetting troughs in nonelective utilization with greater elective utilization.

Some of the fluctuation both in emergent and in elective use of hospital services correlates with climate, holidays, and vacation patterns. Episodes related to chronic illness (especially for the aged) and inpatient hospital utilization are highest in the winter; births are highest in the fall; outpatient activity is highest in the summer. This seasonal fluctuation, most of which would not be smoothed by present admissions scheduling practices, may have consequences on the cost and quality of hospital medical care. For example, in hospitals unaccustomed to operating close to capacity, seasonal fluctuations in demand result in a strain on staff and nonlabor resources. As a result of congestion, length of stay may increase due to queuing in ancillary service departments and an overall decline in the efficiency of care provided, physicians may begin rationing admissions, and quality of care may decline. Luke and Culverwell [3] and Rafferty [4] found some evidence to support these hypotheses. In response to seasonal fluctuations in demand, hospitals may hire more part-time staff, who are frequently more expensive and less efficient than full-time staff. Alternatively, hospitals may set input levels high enough to cope with peak demand and thus hold more excess labor and beds for much of the year.

The purpose of this article is to describe the degree of seasonality in the use of hospital services as well as the nature of the hospital industry response to this seasonality in demand and to consider the effects of this fluctuation on hospital industry size and costs. Intramonth cycles, such as regular daily and weekly variations in utilization and staffing, are not examined. The data used in the analyses are the monthly hospital panel data of the American Hospital Association, which cover the period January 1963–August 1979. Seasonal indexes for series describing aspects of hospital utilization and staffing are constructed and then analyzed graphically. The relationship between hospital industry staffing levels and the seasonal fluctuation in demand is then examined further in a linear regression model of monthly full-time-equivalent personnel. We conclude with a summary of the results and a discussion of some policy considerations.

Our empirical work is limited to analyses at an aggregate level for community hospitals in the entire United States and in each of the nine

U.S. census regions. This poses a real problem, because the behavior we are interested in studying—hospital responses to seasonal variation in demand for services—is at a micro level. A more appropriate unit of analysis would be the individual hospital, but monthly data on individual hospitals were unavailable for use in our study. The aggregate level analyses should provide some limited clues, however, about behavior in individual hospitals. In this next section, we consider the effects of seasonal variation in utilization and relate the hospital industry's ability to cope with fluctuation to those of other industries that also face seasonal variation in output.

RESPONSES TO SEASONAL VARIATION IN UTILIZATION

The hospital response to seasonal variation in utilization depends, to some extent, on the incentives underlying hospital behavior. Most researchers agree that hospital goals are complex and multidimensional, and that a realistic specification of a hospital's objectives would account for measures of volume, quality, and the levels of amenities (for both patients and staff).

In the following discussion, we consider three alternative models of hospital staffing. In the first, hospitals follow patterns in other industries and try to adjust staffing levels to fluctuations in demand. In the second, hospitals staff for expected peak utilization, and in the third, hospitals staff for bed capacity.

Suppose that hospitals attempt to adjust the number of person-hours worked in a given month in accordance with their expectations regarding the volume of patient-days and the severity of case mix. Such behavior is plausible under the simplifying assumption of Davis [5] that hospitals maximize net revenue. Fluctuations in person-hours worked should then tend to mimic the pattern of fluctuations in utilization. The amplitude of these fluctuations, however, would generally be smaller than the amplitude of utilization fluctuations, because institutional constraints such as labor contracts force hospitals to treat some part of labor costs as fixed. Nonetheless, we might expect to see a positive correlation between monthly staffing levels and monthly occupancy rates. The amount of inefficiency associated with the degree of seasonality depends, then, on the amplitude of monthly variations in utilization as well as the costs of adjusting person-hours worked.

Adjustment costs differ for increases versus decreases in output. For increases, adjustment costs may include overtime pay for current staff, the higher fees and training costs associated with personnel brought in from temporary employment agencies, and the training costs of new hires. For decreases in output, adjustment costs depend on unemployment compensation and severance pay if layoffs occur, and on the nonpecuniary or morale costs to employees of the threat of layoff or reduced working hours.

Alternatively, if hospitals staff for peak load periods, then during most of the year hospitals employ too much labor given the volume of patients served. Worker productivity is lower in the low-utilization periods than in periods of high demand. Behavior of this kind may be quite rational for hospitals, since the costs of holding excess labor can frequently be passed on to third-party payers.

For nonhospital industries which are unable to "pass through" most cost increases to consumers and which are subject to large seasonal fluctuations in output, changes in the number of workers employed in response to changes in output are reasonably large relative to those in the hospital industry. In Table 1, the percentage change from the trough month to the peak month of the year in output and in the number of workers employed are presented for the hospital industry and compared with data for 17 U.S. manufacturing industries. Data for the manufacturing industries are from Fair [6]. Acknowledging that the hospital industry differs in most basic respects, certainly, from the types of industries listed, it is nonetheless interesting that the rate of employment to output change (M/Y) is far lower for hospitals than for any other industry.

It should be noted, however, that the hospital's environment and the particular characteristics of its labor markets may act to impede the process of labor adjustment when utilization changes. As noted elsewhere [7], hospitals that perceive a shortage of nurses will tend to hire them whenever possible. Since many nurses first enter the labor market in June or July, just after graduation from nursing school, hospitals may hire at those times to build up a reserve of staff, even though utilization is declining during that time. The same argument holds for other types of health professionals. Hiring in the summer months may, in fact, be efficient, if training time for new hires is taken into account. When the need for health professionals is low, the hospital can take time to train new graduates without reducing quality of care.

Physician influence on staffing decisions may limit a hospital's ability to adjust person-hours worked to utilization. In the same way that physicians use a "quality improvement" argument to influence the capital acquisition decisions of administrators [8], physicians may encourage staffing for capacity or overstaffing to guarantee improved quality of care.

Table 1: The Percentage Changes from the Trough Month to the Peak Month of the Year in Output (Y) and Employment (M) for 1964*

Code	*Industry*	*Y*	*M*	*M/Y*
	Hospitals†	10.3	-0.47	-0.05
201	Meat products	24.9	7.9	0.32
207	Confectionery and related products	79.7	17.2	0.22
211	Cigarettes	44.9	3.9	0.09
212	Cigars	79.3	19.2	0.24
231	Men's and boys' suits and coats	30.4	4.4	0.14
232	Men's and boys' furnishing	24.8	6.6	0.27
233	Women's and misses' outerwear	19.2	5.8	0.30
242	Sawmills and planning mills	28.7	10.8	0.38
271	Newspaper publishing and printing	23.3	2.8	0.12
301	Tires and inner tubes	19.9	5.0	0.25
311	Leather tanning and finishing	19.8	7.3	0.37
314	Footwear, except rubber	17.0	4.8	0.28
324	Cement, hydraulic	99.0	15.9	0.16
331	Blast furnace and basic steel products	25.3	13.3	0.53
332	Iron and steel foundries	24.6	7.7	0.31
336	Nonferrous foundries	13.2	4.6	0.35
341	Metal cans	71.8	14.4	0.20

*The data for the 17 manufacturing industries were taken from [6], p. 21.

†The output measure for hospitals is patient-days.

Another reason why hospitals may find it difficult to adjust employment to output is that their staff is essentially composed of many persons with "indivisible" and very specialized skills. In departments with few workers (as in many of the ancillary departments), layoffs or even cutbacks in hours worked may be infeasible when demand slows down, because no part-time workers are available for those departments or because equipment rather than utilization determines staffing needs. In addition, the skills required in some departments may be so specialized that full- or part-time workers in other departments could not be brought in to cover increases in utilization should they occur.

For many hospitals, cost minimization may not be an important objective. In these hospitals, the response to seasonal variation in demand may be to set staff levels high enough to provide "adequate" care at times of peak utilization and to hold onto excess staff during slack periods. In such hospitals, we might expect a reduction in seasonal variation to be accompanied by a slight reduction in average staffing levels. Hospitals that staff to meet average utilization requirements, however, might not adjust staff complements even if seasonal fluctuations are smoothed. Staffing changes for these hospitals follow long-run secular trends in utilization.

The third model is one where hospitals set staffing levels high enough to provide "adequate" care if the hospital is operating at capacity, i.e., close to 100 percent occupancy. The main reason to expect that hospitals staff for bed capacity and not expected peak occupancy is that staffing for bed capacity is the minimax-regret decision for a hospital administrator facing uncertain utilization. That is, it ensures that just enough staff, but no more than that, will be on hand to care for patients in the event that all beds are full. To the extent that a hospital sets staffing levels in response to bed capacity, a reduction in the seasonal variation in demand will probably have no impact on staff or payroll levels unless accompanied by a reduction in bed capacity.

As a practical matter, it is difficult to infer from an examination of data on staffing and utilization patterns whether hospitals are responding to expected peak utilization or to bed capacity when they are setting staffing levels. To make such an inference we need to compare the staffing patterns of two sets of hospitals: one in which bed capacity for each hospital remained constant but expected peak utilization varied, and another for which expected peak utilization remained constant but bed capacity varied. The problems are that expectations are not directly measurable, and in addition, that expected peak utilization and bed capacity are related. Current bed capacity is in part a function of long-range expectations regarding peaks in demand. If such forecasts are fairly accurate, prior expectations will be close to short-run expectations regarding peak loads upon which staffing decisions are made. Further, short-run expected peaks in utilization are almost always bounded by bed capacity.

We turn now to the data and methods of analysis used to examine seasonality in the hospital industry.

DATA SOURCES AND METHODS

The data analyzed are the Hospital Panel Survey Data of the American Hospital Association (AHA). The survey data are monthly, covering the period January 1963–August 1979, at ten aggregate levels: for the entire United States and for each of the nine census divisions. The survey from which the estimates are made covers approximately 2,000 community hospitals (about 34 percent of all AHA-registered community hospitals). The survey is conducted through the mail at the beginning of each month with a follow-up questionnaire sent to nonrespondents. The AHA reports that the response rate is consistently over 70 percent. Because the data are tabulated at an aggregate level, our ability to make inferences is limited with respect to the behavior of individual hospital units, but we can make some generalizations that apply to the industry as a whole.

The staffing variables available from the survey are, unfortunately, crude. Regularly employed full-time and part-time staff are reported separately. From these, the AHA constructs full-time-equivalent personnel. The data include staff who are on vacation and on paid sick leave, and they do not reflect staff overtime hours. They exclude trainees such as interns, residents, nursing and other students, private duty nurses, and volunteer workers. A preferred employment measure would be person-hours actually worked. Similarly, the salary expense variable excludes the fees and wages paid to trainees. The inability to adjust the staffing and payroll data for paid vacation, sick leave, and uncompensated overtime weakens the inferences we can draw regarding the degree to which hospitals adjust real person-hours worked to utilization changes.

Seasonal indexes for selected variables are constructed in a manner similar to that of Phillip and Dombrosk [9]. The seasonal index for each variable is calculated in two steps. First, an index value for each month of each year is obtained by dividing the actual value of the variable by the 13-month moving average around that month. The index shows what proportion the monthly value is of the average for neighboring months. For example, a value of 1.07 for admissions indicates that admissions in a particular month run 7 percent above the 13-month average of nearby months. (The average is based on the monthly value, the 6 months before and the 6 months after the month.) Second, the index for each month is calculated by averaging that month's values across the 16-year period, 1963–1978.

For measures which are sensitive to the number of days in the month, such as admissions or patient-days, the data are standardized by length of the month before calculating the index values. A thorough discussion of the construction of these indexes is contained in Appendix II of Phillip and Dombrosk. The methods of calculation used for this article are the same as the methods reported there, except that in this case, several of the series are standardized for the length of the month, and for every index, all of the available data are used. Phillip and Dombrosk discarded outlying monthly values before calculating their indexes.

SEASONAL VARIATION IN THE UTILIZATION OF HOSPITAL SERVICES

In this section, we present a brief description of the monthly variation in the utilization of hospital services. The description follows closely the more detailed description offered by Phillip and Dombrosk [9]. Table 2 and Figure 1 display seasonal coefficients for the United States for selected series of hospital utilization. Admissions are highest during January and February, generally above average for the first half of the year; then they decline gradually in the second half of the year and drop sharply in December. Length of stay is longest in the winter, when the proportion of over-65 patients is highest [9]; shortest in the summer, when the incidence of illness is low [10]. Inpatient-days is a composite of admissions and length of stay. Inpatient-days, which are highest during the winter, decline to 2.8 percent below average in June, July, and August, and drop sharply in December.

Figure 2 displays the seasonal coefficients for outpatient visits and adjusted patient-days, which is the sum of inpatient-days and adjusted outpatient visits. Adjusted outpatient visits equals outpatient visits times the ratio of revenue per outpatient visit to revenue per inpatient-day. The measure thus expresses outpatient visits in equivalent inpatient units. Outpatient visits are highest during the summer and low in the winter, and thus serve to moderate the seasonal fluctuation in inpatient-days. Adjusted patient-days, the best available measure of total hospital output, are 4.5 percent above average in the winter and slightly below average during the summer, and they drop sharply—by 7 percent—in December.

Table 2: Seasonal Coefficients for Selected Hospital Utilization Series

Month	*Admissions*	*Length of Stay*	*Patient-Days*	*Outpatient-Days*	*Adjusted Patient-Days*
January	1.033	1.014	1.046	0.958	1.038
February	1.031	1.030	1.062	0.999	1.057
March	1.028	1.014	1.043	1.008	1.040
April	1.022	1.020	1.020	1.016	1.020
May	0.989	1.009	0.997	1.022	0.999
June	1.027	0.965	0.990	1.038	0.993
July	0.994	0.973	0.968	1.011	0.971
August	1.005	0.966	0.971	1.022	0.975
September	0.984	0.992	0.997	1.021	0.980
October	0.988	1.009	0.998	1.009	0.998
November	0.979	1.016	0.995	0.972	0.995
December	0.915	1.013	0.927	0.913	0.927

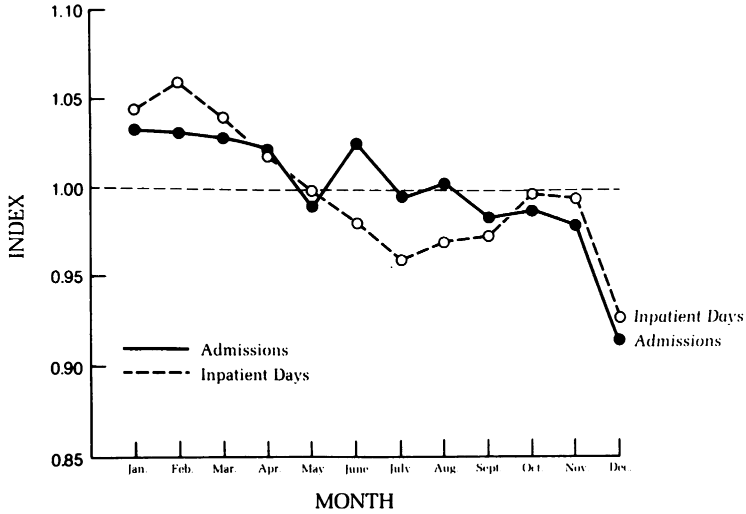

Figure 1: Seasonal Coefficients for Selected Utilization Measures Based on Aggregate Data for the Total United States

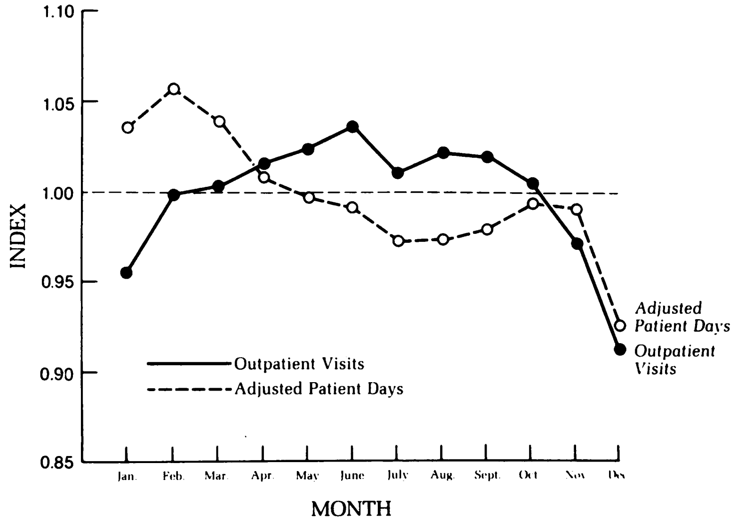

Figure 2: Seasonal Coefficients for Outpatient Visits and Adjusted Patient-Days Based on Aggregate Data for the Total United States

A measure of hospital output adjusted for changes in resource intensity would probably exhibit less seasonal fluctuation than the measure of adjusted patient-days. Lengths of stay are longer during the winter, and resource intensity decreases as length of stay increases. Surgeries are a larger proportion of adjusted patient-days in the summer than in the winter (see below), so hospital resource utilization probably is not as far below average in the summer as the adjusted patient-days series indicates.

If hospitals try to smooth utilization patterns over the year, one can expect that target levels for elective surgical admissions are low in the winter, high in the summer, and high in December. Table 3 and Figure 3 display the seasonal coefficients for surgery. With the exception of the low rate in December, the series follows the preferred pattern. Surgeries are high during June, July, and August, and only slightly above average during the winter. We would like to have a measure of elective and emergency surgeries, but no such measure is available. Because illness levels are highest during the winter, it seems likely that emergency surgeries are higher than average during the winter. Thus, for most of the year, elective surgeries probably counter the seasonal fluctuations in emergency admissions—when emergency admissions are high in the winter, elective surgeries are below average, and when emergency admissions are low in the summer, elective surgeries are high.

The month of December represents an exception to this rule. The December drop in surgery is steeper than the decline in nonsurgical admissions. Low levels of hospital utilization in December are, in large part, a result of low levels of elective admissions, particularly during the week between Christmas and New Year's Day. As Phillip and Dombrosk put it, December is "a month in which consumers, doctors, and hospital personnel would rather be someplace else than in the hospital."

The winter peak in adjusted patient-days is primarily a result of climatic conditions, and is moderated by existing patterns of elective admissions. Nevertheless, a policy to encourage or coerce some of the January, February, and March elective admissions into December or into the summer might reduce the seasonal peak in aggregate hospital industry utilization by 2–3 percent.

The amplitude of seasonal fluctuations in utilization will almost certainly be larger in individual hospitals than in the industry as a whole. Friedman and Pauly [11] present data on the seasonal patterns of admissions for 25 hospitals from monthly reports to the Hospital Administrative Services department of AHA. The amplitude of seasonal fluctuation varies inversely with bed size. In small hospitals (25–50 beds), seasonal fluctuations in admissions are substantial (often ±10 percent from quarter to quarter), while in large hospitals (300–500 beds), quarterly fluctuations average around 3 percent, not much larger than the fluctuations we see at the aggregate level. The aggregate level of our data probably understates the amount of seasonal fluctuation experienced by individual hospitals; but at least for large hospitals, the aggregate data provide a reasonable approximation.

SEASONAL PATTERNS IN HOSPITAL INDUSTRY STAFFING

Next we examine the relationship between hospital industry staffing levels and the seasonal fluctuation in demand faced by the industry. We also consider the implications of this relationship for industry costs. We are primarily interested in two questions:

- Do hospital industry staffing levels or payroll vary with predictable short-term fluctuation in the utilization of hospital resources?
- Do hospitals set their level of staff to cope with utilization or with bed capacity?

Seasonal coefficients for full-time-equivalent personnel (FTEs) are displayed in Figure 4. The most striking characteristic of this series is its small amplitude; the seasonal coefficients for FTEs vary from 0.994 to 1.008 (see Table 4), about one-tenth of the range of the

Table 3: Seasonal Coefficients for Admissions and Their Components

Month	*Admissions*	*Surgical Admissions*	*Nonsurgical Admissions*
January	1.033	0.991	1.073
February	1.031	1.019	1.043
March	1.028	1.033	1.023
April	1.022	1.040	1.055
May	0.989	0.999	0.978
June	1.027	1.089	0.969
July	0.994	1.011	0.979
August	1.005	1.030	0.981
September	0.984	0.969	0.998
October	0.988	0.975	1.000
November	0.979	0.971	0.987
December	0.915	0.872	0.956

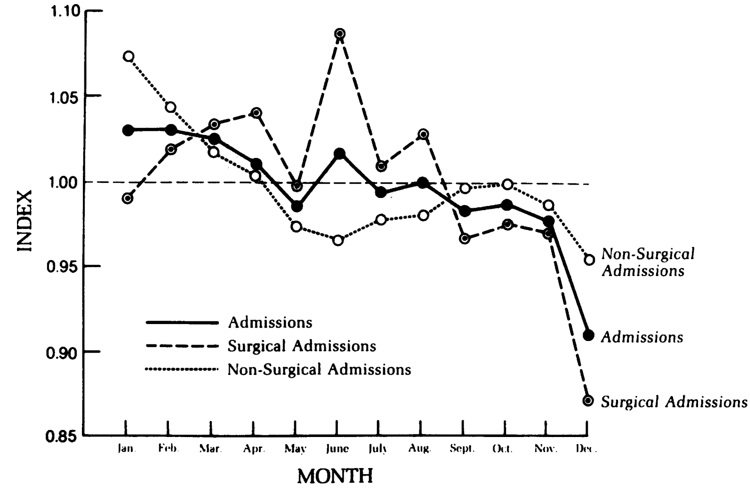

Figure 3: Seasonal Coefficients for Admissions and Components of Admissions Based on Aggregate Data for the Total United States

seasonal coefficients for adjusted patient-days. The fluctuation displayed by the series does not follow the fluctuation in hospital utilization. FTEs are slightly higher in the summer, when a cohort of new graduates enters the market, and do not drop in December or increase during the winter. While neither FTEs nor series for payroll exhibit a drop in December, July, August, or September, it is likely that hours worked in these months are below the annual average, since a larger percentage of the staff is on vacation or using accrued "comp" time. It is likely that some slight relationship exists between utilization and actual hours worked, although it is not reflected in these data. As noted earlier, since the data do not adjust for paid vacation and uncompensated overtime, they cannot capture the full adjustment in actual person-hours worked to changes in utilization. Nevertheless, it appears that the January, February, and March increase in utilization is not accompanied, at the industry level, by an increase in the reported level of payroll or staff.

The apparent February increase in monthly payroll expense is probably an artifact of the adjustment process used to standardize the payroll data for length of month. Since some employees are paid on a monthly (rather than weekly) schedule, adjusting these salaries for length of month will lead to an overadjustment. We think that this has happened here and that the February payroll seasonal coefficient is probably very close to the January and March value.

Table 4: Seasonal Coefficients for Adjusted Patient-Days, Employment, and Payroll Expense

Month	*Adjusted Patient-Days*	*FTEs*	*Payroll*
January	1.038	0.999	0.991
February	1.057	1.000	1.016
March	1.040	1.001	0.994
April	1.020	0.997	0.992
May	0.999	0.998	0.985
June	0.993	1.001	1.010
July	0.971	1.008	1.012
August	0.975	1.004	1.002
September	0.980	0.999	0.999
October	0.998	0.999	0.990
November	0.995	0.999	0.998
December	0.927	0.994	0.993

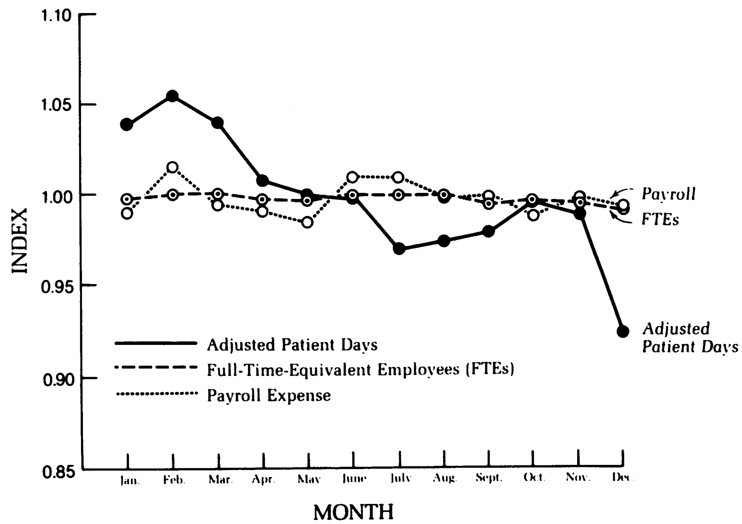

Figure 4: Seasonal Coefficients for Adjusted Patient-Days, Employment, and Payroll Expense Based on Aggregate Data for the Total United States

If hospitals set monthly staffing levels to provide a "minimally adequate" level of care during months of peak demand, as suggested by Harris [8], then a reduction in peak demand through a reduction in seasonal fluctuation in utilization should result in lower staffing levels and associated cost savings. In the long run, a reduction in peak demand should also restrain the growth in hospital bed capacity since capacity decisions depend in large part on expected peak demand. The policy question is whether by reducing the winter peak in utilization at a given hospital, we can expect that hospital staffing levels will be reduced. Much of the aggregate variation in bed capacity comes from the opening and (occasionally) the closing of entire hospitals. The opening of a new hospital leads to an immediate step increase in the number of FTEs and beds in the industry as a whole. Therefore, even if a micro-level relationship exists within a hospital between expected peak utilization and staffing levels, we can expect that at an aggregate level this relationship is dominated by the relationship between newly constructed hospital beds and newly hired FTEs in the industry. Hospital-level data are necessary to control for the staffing discontinuities caused by entry, exit, and consolidation of hospitals.

To determine whether monthly staffing levels vary in response to utilization—controlling for bed capacity, we consider a two-equation model of the following form:

$$FTE^*_{it} = \beta_0 + \beta_1 APD_{it} + \beta_2 BEDS_{it} + \sum_{i=1}^{9} \gamma_i R_i \qquad (1)$$

$$FTE_{it} - FTE_{i,t-1} = \lambda (FTE^*_{it} - FTE_{i,t-1}) + U_{it},\ 0 < \lambda \leq 1 \qquad (2)$$

where:

FTE^*_{it} = desired or optimal number of full-time-equivalent personnel in region i at time t.

FTE_{it} = actual number of full-time-equivalent personnel in region i at time t.

APD_{it} = adjusted patient days in region i at time t.

$BEDS_{it}$ = number of beds in region i at time t.

R_i = dummy variable for census region i (R_i = 1 if the observation is in region i; 0 otherwise).

U_{it} = random error term for region i at time t.

There are two relationships. Equation 1 relates the desired or optimal staffing level in month t to the month's actual utilization level, controlling for bed capacity. Regional dummy variables are included to measure time-invariant fixed effects. A more accurate specification would also include measures of input prices as determinants of optimal staffing levels, e.g., real wages and the relative price of capital. Since monthly input price data are unavailable, they have been omitted in the above specification. Provided input prices do not vary systematically with utilization and bed capacity, we can obtain unbiased estimates of the effects of the latter two variables on staffing levels.

Equation 2 reflects the fact that the observed change in monthly full-time-equivalent personnel may not equal the desired or optimal change in FTEs due to the costs of making such adjustments. It postulates that for month t hospitals will adjust staffing levels only partially to the optimal level for that month. The closer λ is to unity, the greater the adjustment. If adjustment costs are zero, then $\lambda = 1$ and adjustment to the desired level is complete. We expect λ to be relatively small because the time period for adjustment, a month, is short.

Of interest is whether staffing levels vary significantly in response to utilization, controlling for bed capacity and adjustment costs. In terms of the model above, the question is whether β_1, the coefficient of APD_{it}, is significantly different from zero.

Substituting Equation 2 into Equation 1 yields:

$$FTE_{it} = \lambda\beta_0 + (1-\lambda)\, FTE_{i,t-1} + \lambda\beta_1\, APD_{it} + \lambda\beta_2\, BEDS_{it} + (1-\lambda) \sum_{i=1}^{9} \gamma_i R_i + U_{it} \qquad (3)$$

Estimates of $\lambda\beta_0$, $\lambda\beta_1$, $\lambda\beta_2$, $(1-\lambda)$, and $(1-\lambda)\gamma_i$ $(i = 1,2, \ldots 9)$ are obtained by estimating Equation 3. We have data on nine census regions and 200 months beginning with January 1963, for a total of 1,800 observations. Hence, there are 1,791 observations available to estimate Equation 3.

The presence of the lagged dependent variable among the regressors in Equation 3 results in special estimation problems when the residuals are autocorrelated, since ordinary least-squares estimation produces biased and inconsistent parameter estimates [12]. For our data, a test for autocorrelated residuals in Equation 3 indicates the presence of a second-order autoregressive process:

$$U_{it} = \psi_1\, U_{i,t-1} + \psi_2 U_{i,t-2} + e_{it} \qquad (4)$$

The process is stable, provided: $\psi_1 + \psi_2 < 1$, $\psi_2 - \psi_1 < 1$, and $|\psi_2| < 1$.

A two-step procedure proposed by Hatanaka [13] deals with the problem of lagged dependent variables and serially correlated errors. The procedure yields parameter estimates which are both consistent and asymptotically efficient.

Table 5 reports the estimation results. The coefficient of *APD* is statistically significant but extremely small. The relationship between staffing levels and utilization is economically insignificant compared with the relationship between capacity and staffing levels. An increase of 1,000 beds is associated with an increase of 1,641 FTEs. A corresponding increase in adjusted patient-days (1,000 × 30 days or 30,000) corresponds to an increase of 81 FTEs (0.0027 × 30 = 0.081). All three variables in the regression model are measured in units of 1,000.

Our estimate of the partial adjustment parameter, λ, is 0.948, computed from the coefficient of lagged *FTE*s. Estimates of the effects of *BEDS* and *APD* on desired staffing levels are 1.731 and 0.0029, respectively. Here, too, the effect of *APD* relative to *BEDS* is trivial.

These estimates suggest that at the industry level both actual and desired staff size are far more responsive to bed capacity than to actual utilization. This does not deny the possibility suggested in [1] and [3] that individual hospitals set staffing levels in response to expected peak utilization, since the expected peak may be close to bed size. Our analysis suggests that if a reduction in seasonal fluctuation in utilization is to lead to lower hospital costs, it needs to be accompanied by a smaller rate of increase in industry bed capacity. Since bed capacity is a

Table 5: Regression Model for Monthly Full-Time-Equivalent Personnel (FTE_{it}) (Two-Step Efficient Estimator)

Explanatory Variable	*Coefficient*	t-*Statistic*
Intercept ($\lambda\beta_0$)	79.446***	4.353
$FTE_{i,t-1}$ $(1-\lambda)$	0.052**	2.518
APD_{it} $(\lambda\beta_1)$	0.0027***	3.279
$BEDS_{it}$ $(\lambda\beta_2)$	1.641***	18.933
Regional Effects $(1-\lambda)\gamma_i$:		
R_1 New England	-55.169**	-2.354
R_2 Mid-Atlantic	108.191***	4.665
R_3 South Atlantic	66.193***	2.940
R_4 East North-Central	124.506***	5.208
R_5 East South-Central	-30.827	-1.383
R_6 West North-Central	10.180	0.460
R_7 Mountain	-63.476***	-2.754
R_8 Pacific	23.408	1.053
Number of observations	1,791	
R-Squared	0.895	

Note: The residuals follow a second-order autoregressive process given by:

$$U_{it} = 0.481354\ U_{i,t-1} + 0.470028\ U_{i,t-2} + \hat{e}_{it}.$$

**Statistically significant at the α = 0.05 confidence level.
***Statistically significant at the α = 0.01 confidence level.

function of expected peak demand, a reduction in seasonal fluctuation which takes the form of reducing the winter peak in utilization would likely have the greatest potential for cost savings.

SUMMARY AND POLICY CONSIDERATIONS

In the last section we showed that:

- Utilization of hospital resources is highest in the winter, somewhat below average in the summer, and sharply lower in December.
- Surgeries are highest in the summer and around average during the winter. Given that levels of acute illness are highest during winter and lowest in summer, it is likely that elective surgeries are particularly high during the summer and below average during the winter. Thus, elective surgeries may serve to moderate fluctuation in emergency utilization for most of the year. The December drop in utilization, however, is primarily a result of low levels of elective admissions during that month.
- Aggregate hospital staffing levels are primarily related to industry bed capacity. Controlling for bed capacity, a slight relationship appears to exist between monthly hospital industry utilization, measured as adjusted patient-days, and hospital industry staffing levels.

A variety of data problems (inability to distinguish elective from emergent utilization, lack of a measure of case-mix severity, and the lack of a measure of hours worked) preclude us from drawing strong conclusions, even concerning these aggregate relationships.

At the individual hospital level, we can only speculate about the results of seasonal fluctuation in utilization. It seems likely that for a typical hospital which does not face effective incentives to control costs, smoother patterns of seasonal utilization may result in a more uniform level of intensity of care, but will not result in lower staff levels or reduced costs unless accompanied by a slowdown in the rate of increase in hospital bed size. Hospitals that do face incentives to control costs, such as for-profit hospitals, hospitals operated by health maintenance organizations, or hospitals operating under effective prospective reimbursement, might be inclined to reduce average levels of staff in response to a more even seasonal pattern of utilization.

Our macro-level data cannot be used either to support or discredit these speculations concerning micro-level phenomena. The aggregate data, however, are informative for what they do not show. If we had found that staff levels or payroll were higher in the winter and lower in the summer, then our speculation concerning the micro-level relationship would have been different. Similarly, if we had found a macro result that staffing levels, controlling for bed capacity, did respond strongly to utilization, then we would have had some reason to suspect that smoother seasonal patterns of utilization would lead to generally reduced staffing levels.

Policies such as certificate-of-need regulation and prospective reimbursement programs, which have been found to constrain the average bed reserve margin of hospitals [14]—i.e., total beds minus average daily census, may provide an incentive for demand smoothing. By lowering the reserve margin of the hospital such programs, in the absence of any changes in utilization patterns, increase the probability that the hospital will be full and patients turned away. Demand-smoothing policies which lower the winter peak help to decrease this turnaway probability, bringing it closer to a target value. Effective prospective reimbursement programs, in addition to restraining bed capacity, provide incentives for hospitals to contain salary expenses. Demand-smoothing policies may help in achieving this objective by lowering average staffing levels and/or the costs of adjusting staff in accordance with changes in utilization.

The discussion, and to some extent the evidence presented in this article, suggests that policies aimed at smoothing seasonal variation in hospital utilization would have only a limited effect on hospital industry costs through potentially reducing the growth in industry bed capacity. Smoothing seasonal variation in utilization probably would benefit quality of care through a reduction in the strain put on hospital resources in periods of peak demand. It is not clear what kinds of policies can effectively smooth seasonal variation. Admission scheduling systems reduce variation in daily occupancy rates [1] and may increase utilization in December somewhat by delaying some elective admissions in late November and moving some up that otherwise would occur in early January. However, except for areas with a bed capacity shortage, these systems are unlikely to reduce the winter peak in utilization substantially, because patients and their doctors generally will choose to stay out of the hospital in December to the extent that they can.

In other industries (e.g., the airline and telephone industries), peak-load pricing policies have been used as a tool to smooth demand. In the hospital industry, though, several unique barriers forestall implementation of this kind of system. First, in most cases, third-party payers, rather than the patient, pay for the incurred cost of a patient's stay, so that the patient has little incentive to be responsive to price. The real payers, third parties, are not directly involved in the admissions choice. Second, to a large extent, it is the patient's physician rather than the patient who determines the timing of the admission. And physicians, like hospitals, are paid at the same rates by their payers regardless of the hospital's load.

Since the physician generally decides on the date of patient admission, effective policies should be directed in part at altering physician behavior. If the price that physicians were paid varied by month, then desirable shifts in elective admissions might occur. Unfortunately, it is difficult to imagine a differential pricing policy actually being put into place by an insurance carrier or other third-party payer. Not only would it be more difficult to administer than present procedures, but the benefits to the paying party would be somewhat vague and intangible. Health maintenance organizations (HMOs) that either operate their own hospital(s) or have contractual arrangements with hospitals may be an exception. An HMO is able to instruct its physician staff to avoid performing elective surgery in the winter unless the patient has a strong preference for those months. Further, HMOs can realize the cost savings, if any, from smoother demand fairly directly.

In conclusion, this analysis represents only a first step both in measuring and in understanding the effects of seasonality on the industry. Further analysis in the area clearly would benefit from the use of monthly hospital-specific data rather than aggregate-level data of the kind used here. While our findings are probably a reasonable approximation for large hospitals (300–500 beds), they most likely underestimate the amplitude of utilization and staffing cycles which smaller hospitals experience.

ACKNOWLEDGMENTS

The authors wish to thank David Kidder, A. James Lee, Robert Caswell, and Bryan Dowd for helpful comments.

REFERENCES

1. Hancock, W. M., J. B. Martin, and R. S. Storer. Stimulation-based occupancy recommendations for adult medical/surgical units using admissions scheduling systems. *Inquiry* 15:25, March 1978.

2. Brower, F. B. Feasibility of a 5-day medical-surgical unit. *Hospital Administration Currents* 21:1, January 1977.
3. Luke, R. D., and M. B. Culverwell. Hospital bed availability and discharge patterns in the short run. *Inquiry* 17:54, Spring 1980.
4. Rafferty, J. Patterns of hospital use: An analysis of short-run variations. *Journal of Political Economy* 79:154, January-February 1971.
5. Davis, K. P. *A Theory of Economic Behavior in Non-Profit Hospitals*. Unpublished doctoral dissertation, Rice University, 1970.
6. Fair, R. C. *The Short-Run Demand for Workers and Hours*. Amsterdam: North-Holland Publishing Company, 1969.
7. Feldstein, P. J. An Empirical Investigation of the Marginal Cost of Hospital Services. Graduate Program in Hospital Administration, University of Chicago, 1961.
8. Harris, J. E. The internal organization of hospitals: Some economic implications. *Bell Journal of Economics* 8(2):467, Autumn 1977.
9. Phillip, P. J., and S. J. Dombrosk. *Seasonal Patterns of Hospital Activity*. Lexington, Massachusetts: D. C. Heath and Company, 1979.
10. National Center for Health Statistics. *Current Estimates from the Health Interview Survey*. Series 10, Numbers 5, 13, 25, 37, 43, 52, 60, 72, 79, 85, 95, and 100. U.S. Government Printing Office, Washington, DC., 1960–75.
11. Friedman, B., and M. Pauly. Effects of Actual and Forecasted Volume on Hospital Costs: Analysis and Policy Implications. Draft Final Report. Center for Health Services and Policy Research, Northwestern University, April 1980.
12. Johnston, J. *Econometric Methods*. 2nd Edition. New York: McGraw-Hill Book Company, 1972.
13. Hatanaka, M. An efficient two-step estimator for the dynamic adjustment model with autoregressive errors. *Journal of Econometrics* 2:199, 1974.
14. Joskow, P. L. The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital. *Bell Journal of Economics* 11(2):421, Autumn 1980.