HC

PRELIMINARY; PLEASE DO NOT CITE

Within-Year Variation in Hospital Utilization and its Implications for Hospital Costs

Laurence Baker a/b/
Ciaran Phibbs c/a/
James Reynolds a/
Dylan Supina d/

a/ Stanford University
b/ NBER
c/ Palo Alto Veterans Administration
d/ University of Vienna

Abstract

Variability in demand for hospital services may have important effects on hospital costs, but data on within-year variations in hospital use have not previously been available for large samples of hospitals. We use data on California and Florida hospitals to examine variability in hospital utilization at the daily level. We demonstrate that there is substantial within-year variation, and that there are large differences across hospitals. We then examine the impact of variation on hospital costs, showing that increases in variance increase total operating expenditures and labor expenditures. Finally, we discuss the determinants of variations in daily census, focusing on hospital size and managed care effects. We find some evidence that large hospitals have less variability and are better able to respond to variations than smaller hospitals, and that hospitals with higher proportions of HMO patients appear to have lower variances.

We gratefully acknowledge support for this research from the Robert Wood Johnson Foundation. We also acknowledge helpful research assistance from Cassandra Guarino and Susan Wheeler.

I. Introduction

Many authors have noted the importance of variability in demand for health care as an influence of hospital costs. Unlike many goods, health care cannot be stored, and the demand for health care is by its very nature often unpredictable. If hospitals operate with the goal of keeping the probability of having to turn away patients below some threshold level, the amount of capacity the hospital requires will be a function of demand variability. High variance hospitals will have to keep more capacity on hand to meet high demand peaks than low variance hospitals, but this capacity will often go unused. Providing this kind of "standby capacity" generates costs for hospitals.

Variability in demand may also lead to inefficiencies in production. For example, staffing costs may increase if extra workers are regularly called in or excess workers are sent home due to an inability to accurately forecast the number of patients in the hospital at any given time. Costs for other items that require accurate scheduling, like food and perishable supplies, may also be higher with highly variable demand.

A number of previous papers have examined hospital costs, estimating a range of cost functions. Some work has focused on the effects of excess capacity on hospital costs (e.g., Keeler and Ying, 1993; Gaynor and Anderson, 1995). Other work has investigated the effects of variation in demand for hospital services on costs (e.g., Gaynor and Anderson, 1995; Joskow, 1980). But, virtually all of this work has been undertaken with annual data on hospital utilization. Since hospital demand varies considerably from day to day, or even hour to hour, annual data do not capture much of the important variation, which severely hampers efforts to understand variation in demand. Joskow (1980) and Lynk (1995) both present some data on

within-year demand variations, but each uses samples from very small numbers of hospitals for limited time periods.

The first objective of this paper is to compile daily utilization data for a large number of hospitals. Using detailed discharge records, we have computed the midnight hospital census for every hospital in California for every day between January 1, 1983 and December 31, 1995. We have also computed the midnight census for every hospital in Florida every day between January 1, 1988 and December 31, 1995. We use this data to characterize day-to-day variation in hospital utilization, and we show that there is substantial variation in hospital utilization, and that patterns of utilization variation vary across hospitals.

The second main objective of this paper is to use the detailed data on variance that we have derived to investigate the effects of utilization variance on costs. We build on the theoretical work of Gaynor and Anderson (1995), and estimate hospital cost functions that include measures of variance. We find that variance can have important effects on hospital costs.

Given that variance can influence costs, it is important to investigate the determinants of variance. We focus on two particular issues. The first is whether hospital mergers can reduce variance, and thereby contribute to economies of scale. If larger hospitals have smaller variations in utilization, they may have lower costs. It may also be that larger hospitals can more effectively absorb variations in utilization, so that variance has a smaller effect on costs. We investigate the relationships between hospital size, utilization variance, and costs. Second, we examine the possibility that managed care can smooth variations in hospital use, reduce variance, and thus reduce costs. If managed care plans are effective at taking advantage of slack in the system, they may be able to work with hospitals to do things like more effectively schedule

elective admissions, which could help reduce variation and lower costs. We examine the relationship between managed care activity, variance, and costs.

II. Daily Variation in Hospital Utilization

We begin by characterizing day-to-day variation in hospital utilization. This section describes the data and measures that we use.

II.A. Data

To compute estimates of variation in hospital utilization, we obtained discharge abstracts for each patient discharged from a California hospital between January 1, 1983 and December 31, 1995 and for each patient discharged from a Florida acute-care hospital between January 1, 1988 and December 31, 1995. This amounted to 46.9 million records for California and 14.5 million records for Florida. The California Office of Statewide Health Planning and Development (OSHPD) supplied the California data. The State of Florida Agency for Health Care Administration (AHCA) supplied the Florida data. From the original data, we excluded discharges from non-acute facilities and discharges associated with lengths of stay greater than 1 year. This removed 3.8% of the original California records and 0.003% of the Florida records.

Each record indicated the admission and discharge date. We used this information to compute the number of patients in a given hospital at 11:59:59 PM of each day (i.e. census at the midnight that ends each day). For example, a person admitted on January 1 and discharged on January 2 would be included in the census count on January 1. This definition is commonly used

in administrative counts and other activities in which census data are computed. We term this figure the "daily census."

Individuals admitted and discharged on the same day are not included in the daily census counts since they were not in the hospital over any midnight. We did compile separate data on the number of individual admitted and discharged on the same day. There are relatively few "same day" patients. For example, in 1994, there were 111,013 total same day patients in California, relative to 14.6 million overnight patient days.

There are two potential sources of mismeasurement in our daily census counts. First, some discharge abstracts may be inaccurate. Because hospitals in both California and Florida are required by law to submit discharge data to the state agencies, we expect virtually complete records of hospital discharges for the periods studied. But there may be inaccurate records and miskeyed data. There may also be differences across hospitals in the interpretation of reporting guidelines or in the degree of compliance with guidelines. We found evidence, for example, that one California hospital sometimes reported "long term care" patients as acute care patients until instructed to correct this practice by OSHPD. Overall, we believe mismeasurement from this source to be minimal.

The second type of mismeasurement stems from the fact that data is not available to us about discharges occurring subsequent to the sample period. Patients in the hospital during the study period but discharged after December 31, 1995 are therefore not counted. To illustrate, a patient admitted on December 23, 1995 and discharged on January 5, 1996 would not be included since we have no records on patients discharged after December 31, 1995. This will cause us to undercount daily census in the days leading up to the end of 1995. We estimate that

about 1% of patients in California had a length of stay of 31 days or more. Examining data from 1983, 1993, 1994, and 1995, we found no patients with lengths of stay over 61 days. We thus eliminate December 1995 from all subsequent calculations. For our analyses of costs, we eliminate all of 1995, which removes this source of mismeasurement entirely.

These data represent about 500 hospitals per year in California and about 220 hospitals per year in Florida. Consistent with general industry trends, the number of hospitals in the samples declines over time. For example, the California sample contains 572 hospitals in 1984 but only 471 hospitals in 1994. The Florida sample contains 226 hospitals in 1988 and 210 hospitals in 1995.

II.B. Hospital occupancy

We use our data to summarize occupancy in several ways. First, we simply compute average daily census across hospitals for each year. Figure 1 plots the mean daily census across all California and all Florida hospitals in each year. Consistent with previous literature, there are declines in average daily census in both states beginning in the early 1990s.

II.C. Within-year variation in hospital occupancy

There is considerable variation in daily census within years. Figure 2 plots mean daily census by month for selected years. Mean daily census is highest in January and February and falls over the course of the year, reaching a minimum in the late summer and fall. We do not plot the data here, but analyses by week of the year reveal similar patterns, and also show drops in daily census in the weeks near the Thanksgiving, Christmas, and New Years holidays. Figure 3

plots mean daily census values for each day of the week. Daily census is highest during the week and lowest on weekends.

These data are consistent with the types of variations one would expect, given that at least some hospital utilization is elective and short term, and can be scheduled to coincide with convenient times for patients and physicians. But, it is also instructive to note that variation across months and days of the week accounts for a relatively small share of total occupancy. For example, the difference between the mean Sunday census and the mean Thursday census is only about 10% of the average daily census. Much hospital utilization is not elective or short term.

These figures also illustrate the inherent capacity demands that face hospitals. Even without accounting for unexpected shocks in demand, variations across months of the year and days of the week force hospitals to have beds that will not be filled at all times. To meet Thursday's demand, the average hospital needs about 10 more beds than it will need on Sunday.

II.D. Characterizing day-to-day variation for hospitals

We can use our data to characterize the variation in daily census facing each sample hospital. In many cases, the distribution of daily census is approximately normal, at least to a first approximation. Figure 4 plots the distribution of daily census for each of four randomly selected California hospitals. In all four of these cases, the daily census distribution has a characteristic normal shape, although, in each case, the distribution fails formal statistical tests for normality. At least part of the failure appears to result from abnormalities in the tails of the distribution. Generally similar patterns are observed in other hospitals.

Table 1 reports summary statistics for the variance and standard deviation of daily census across California hospitals for some sample years. The first row of each section contains the average variance and the average standard deviation across hospitals in 1984, 1990, and 1994. Below the average across hospitals are the 10th, 50th, and 90th percentiles across hospitals. In 1990, the average variance was 228, and the average standard deviation was about 13. There is wide variation across hospitals: in 1990 the 10th percentile of variance was 10, and the 90th percentile was 503. A difficulty with the use of variance and standard deviation is their scale dependence. To adjust for this, we computed coefficients of variation:

$$CV_{h,y} = (\sigma_{h,y} / \bar{c}_{h,y}) * 100 \qquad \textbf{(1)}$$

where σ denotes the standard deviation, and c denotes the daily census. Subscripts h and y denote hospital h and year y. The CV normalizes by the mean and expresses the amount of variation as a percent of the mean. The CVs in column 4 of Table 1 indicate that, on average, the standard deviation is about one fourth of the mean. The CVs also indicate substantial differences across hospitals in the degree of variation in daily census.

To the extent that the distributions of daily census values are non-normal, mean and variance may not fully characterize the distributions. Percentile distributions provide a non-parametric way of summarizing variation. Table 1 reports the difference between the 75th and 25th percentile and the difference between the 90th and 10th percentile for hospitals. As above, these indicate sizable variations in daily census within years and substantial differences across hospitals.

Daily census variance declined over time. The top panel of Figure 5 plots the average variance and various percentiles of variance across hospitals for California hospitals for 1983-1995. At all levels, variance declines between 1983 and 1995, with the average variance across hospitals falling by about 30 percent. Much of the decline in variances observed between 1983 and 1995 is associated with declines in average daily census. The bottom panel of Table 5 plots average coefficients of variation. No decline in the coefficient of variation is evident between 1983 and 1995, although sizable differences remain across hospitals.

II.E. Decomposing variance into predictable and unpredictable variance

Some portion of the variation in daily census can be anticipated by hospital administrators. Clear patterns can be discerned across days of the week and months of the year, and there may be other factors that can be anticipated as well. One goal of this paper is to investigate the effects of variation in utilization on costs. It may be that variation that can be anticipated has a different impact on costs than variation that cannot be anticipated.

To obtain some information about this, we attempted to predict daily census for each day in a given year using data on day of week and week of year patterns from the three preceding years, and a trend variable. We estimated regression equations of the form:

$$c_{h,y,d} = \beta_0 + \beta_1 \text{ DAY OF WEEK} + \beta_2 \text{ WEEK OF YEAR} + \beta_3 \text{ YEAR} + \epsilon_{h,y,d} \qquad (2)$$

for each hospital in the sample independently, where subscript d denotes day (d=1,...D_y where D_y is the number of days in year y). DAY OF WEEK is a set of 6 dummy variables accounting for variation across days of week Sunday through Saturday, and WEEK OF YEAR is a set of 51 dummy variables accounting for variation across the 52 weeks of the year. YEAR is a linear

time trend. We then used the estimated coefficients to predict daily census for each day of each year after 1986, and computed the variance of the predicted values. This will be an estimate of the amount of variation that could be anticipated by an administrator familiar with previous trends in his or her hospital over days and weeks. It will, of course, underestimate the total amount of variation that might be anticipated by an administrator aware of additional sources of variation.

The rightmost columns of Table 1 present the results from this exercise for 1990. On average, the "predictable" variance is 111, about half of the total variance. Across hospitals, the average ratio of predictable variance to total variance is 0.48. As above, however, there is considerable variation across hospitals.

III. Variance in occupancy and hospital costs

In this section, we use the variance measures discussed in the previous section to examine the effects of variance on hospital costs. We estimate trans-log cost functions that incorporate measures of variance. We also directly explore the implications of variance for labor costs.

III.A Model specification

We begin our model specification building on previous work by Gaynor and Anderson (1995), who have discussed the theoretical formulation of hospital cost functions in the presence of uncertain demand. They assume a hospital with production function $y_t = f(x_{vt}, x_{qt}, k)$ where y is output in period t, x_v is a vector of variable inputs, x_q is a vector of quasi-fixed inputs, and k is capital. Demand in period t, d_t, is a random variable with (invertible) distribution $G(d)$, which in

principle could be a distribution conditional on information known prior to the start of period t. Hospitals also operate with the goal that demand can exceed supply only α percent of the time. That is, each hospital has a target level for the probability that demand will exceed the available bed supply and the hospital will be forced to turn away patients. They show that, under these conditions, the variable cost function can be written

$$C_v = C_v(y_t, G^{-1}(1-\alpha), w_v, w_q, k_t) \tag{3}$$

where C_v is variable costs, G^{-1} is the inverse of the demand distribution function, and w_v and w_q are the factor prices for variable and quasi-variable inputs. In this specification, $G^{-1}(1-\alpha)$ captures the standby capacity that the hospital believes it will need to meet peak demand, subject to maintaining the expected turnaway probability below α.

We wish to empirically estimate a version of (3). In practice, the model estimated varies somewhat from the theoretical specification. Since we focus on inpatient costs, we include the number of patient days as our measure of output. To capture other aspects of output, like the complexity of cases or patient severity, we include a case mix measure, measures of hospital teaching status, measures of the percent of patients covered by HMOs, and measures of for-profit and public ownership. We include bed size to capture the effects of size on costs; bed size may also be correlated with severity or case complexity.

The theoretically appropriate measure for $G^{-1}(1-\alpha)$ is the capacity that the hospital believes will keep its turnaway probability at or below α. However, we cannot measure α for each hospital. Rather, our approach is to describe the distribution of demand. We begin here with the assumption that $G(d)$ is approximately normal, and can be adequately summarized with a mean and a variance. We also incorporate some non-parametric measures in our estimation

below. We assume that hospitals form opinions about the distribution of demand in one year based on the actual mean and variance in the previous year. More sophisticated expectations-generating processes are clearly possible, but have not yet been implemented. Thus, for $G^{-1}(1-\alpha)$ in equation (3), we substitute the average daily census in the preceding year and the variance of daily census in the preceding year. Since our estimation below will include hospital fixed effects, the implicit assumption is that α is fixed for each hospital over time, and thus that increases in variance imply increases in the standby capacity needed by the hospital.

There is a potential pitfall with using the observed census as a measure of demand: we see only patients actually admitted to the hospital, rather than the number of patients who might have arrived at or otherwise wish to have been served by the hospital. Most of the hospitals in our sample operated well below capacity on virtually all of the days of the year, and we suspect that using the daily census as a measure of demand will not cause serious problems for these cases. There are, however, a limited number of hospitals that operated at or near capacity, and we will perform our analyses both including and excluding these hospitals to ensure that they do not cause substantial bias.

The cost function theory developed by Gaynor and Anderson incorporates variance in daily census only through its effect on standby capacity. But, we hypothesize that variance may also influence the efficiency with which hospitals operate, for example with respect to labor scheduling, and we wish to incorporate this effect in the model as well. To do this, we note that standby capacity needs should be a function of lagged variance, but operating inefficiencies induced by variance should be caused by current variance. We thus add a measure of current variance to the model to capture these effects. One question this raises concerns the

autocorrelation between variance, and our ability to separately identify effects of both current and lagged variance. Using our data, we estimate that the correlation between variance in year t and variance in year t-1 is about 0.59, which should still allow separate identification of effects.

The model we estimate is

$$C_v = C_v(y_t, E(y_t \mid y_{t-1}), Var(y_t \mid y_{t-1}), Var(y_t) w_t, k_t; \theta_t) \quad (4)$$

where y_t is the average number of patients in the hospital on any given day, $E(y_t \mid y_{t-1})$ is the expected number of patient days in year t given patient days in year t-1, $Var(y_t \mid y_{t-1})$ is the expected variance of patient days in year t given patient days in year t-1, $Var(y_t)$ is the variance of daily census in year t, w represents input prices, k represents fixed factors, and θ represents other characteristics like case mix, teaching status, and ownership. In practice, we use the actual average daily census (patient days/number of days in year) in year t for y_t, the average daily census in year t-1 as $E(y_t \mid y_{t-1})$, the variance in average daily census in year t-1 for $Var(y_t \mid y_{t-1})$.

Since theory does not provide guidance as to the specific form of the cost function, we follow much of the previous literature in this area and estimate a trans-log cost function (Gaynor and Anderson, Keeler and Ying). The trans-log has a number of favorable properties. It has the usual properties of a cost function, and also allows for flexible substitution, and locally approximates any arbitrary function. The specific model we estimate is:

$$\begin{aligned} \log C_{v,t} = {} & \beta_0 + \beta_1 \log y_t + \beta_2 (\log y_t)^2 + \beta_3 \log y_{t-1} + \beta_4 (\log y_{t-1})^2 + \beta_5 \log w_t \quad (6) \\ & + \beta_6 (\log w_t)^2 + \beta_9 \log b_t + \beta_{10} (\log b_t)^2 + \beta_{11} \log y_t * \log y_{t-1} \\ & + \beta_{12} \log y_t * \log w_t + \beta_{13} \log y_t * \log b_t + \beta_{14} \log y_{t-1} * \log w_t \\ & + \beta_{15} \log y_{t-1} * \log b_t + \beta_{16} \log w_t * \log b_t + \beta_{17} \log V_{t-1} + \beta_{18} \log V_t \end{aligned}$$

$$+ \beta_{19} \log(\text{casemix}) + \beta_{20} \text{hmo} + \beta_{21} \text{TEACH} + \beta_{22} \text{OWNERSHIP}$$

$$+ \beta_{23} \text{YEAR} + \mu + \epsilon_t$$

where V is the variance of daily census, b is available beds, casemix is the average DRG weight per discharge, hmo is the percent of discharges for which HMO is the primary source of payment, TEACH is a vector of dummies for teaching status indicating academic medical centers and other hospitals with residency programs ("intermediate teaching hospitals"), OWNERSHIP is a vector of dummies for ownership indicating church owned, private non-profit, for-profit, and government hospitals, YEAR is a vector of year dummies, μ is a hospital fixed effect, and ϵ is an error term. For parsimony and for clarity in the presentation of results, we force the variance term to enter linearly. More complex specifications with variance interactions produced qualitatively similar results.

It is possible that some of the variables in the model are endogenous. For example, average daily census may respond to operating expenditures. We have not yet investigated this possibility, but plan to follow previous literature and instrument for utilization levels if necessary.

III.B. Data

We estimate equation (6) using financial data on California hospitals from the California OSHPD. The OSHPD requires all hospitals in California (with some minor exceptions) to submit quarterly and annual financial reports.[1] These reports contain information on total

[1]The annual reports filed by hospitals pertain to each particular hospital's fiscal year. The starting and ending points of fiscal years vary from hospital to hospital, and some cases hospitals

operating costs, bed size, wages, ownership, teaching status, and a range of other variables. We obtained total inpatient operating costs by summing inpatient operating costs from each quarter from the quarterly financial filings of each hospital. We obtained available beds as of January 1 each year from the annual data. Available beds are defined as beds that exist and are available for patient use, but need not be staffed. We obtained data on the average daily census during the year from the quarterly reports. We obtained the average census and the daily census variance for the previous year from our analyses of variance described in the previous section. We obtained wage data from the annual filings. The wage variable we use represents an average wage for six categories of labor, including registered nurses, licensed vocational nurses, aids and orderlies, management and supervisory employees, technicians and specialists, and clerical workers, weighted by the total hours worked by each category. The inclusion of other (rarer) categories of labor did not affect the conclusions, nor did using wages for the various categories of labor separately.

Our original sample contained 6,357 California hospital-year observations. We include only hospitals reported to provide short-term general care. Because we use lagged data, we excluded all observations that represented the first year a hospital appeared in the data set. We also excluded observations for which all of the necessary financial and other data were not available, observations reporting zero inpatient operating expenditures for any year or zero census or variance in the previous year. In the end, the data set we use has 4,386 hospital-year observations. This is a panel data set, with many hospitals observed multiple times. The panel is

---

change their fiscal years over time. To allow for standardized analysis, we constructed data pertaining to calendar years (January 1 through December 31) from the annual reports. The quarterly reports are all filed for calendar quarters, so conversion to consistent dates was not necessary for the quarterly data.

unbalanced since some hospitals do not appear in the data set for all years. Means and standard deviations of the variables used in the analysis are shown in Table 2.

III.C. Results

Results from estimation of equation (6) are presented in Table 3. The first column presents the main specification. Most of the variables have the expected signs and are statistically significant. We focus on the variance terms, which are both positive and significant. 10% increases in lagged variance are associated with 1.8% increases in costs. This is consistent with the hypothesis that increases in variance contribute to the need for additional standby capacity and drive up costs. As noted by Gaynor and Anderson (1995), the significance of lagged variance and lagged daily census are consistent with the importance of including expected utilization in hospital cost functions. The coefficient on the current variance term indicates that increases of 10 percent in current variance are associated with 2.3% increases in costs, consistent with the hypothesis that variance may lead to operational inefficiencies in hospitals.

If the daily census distributions are non-normal, variance could be an insufficient measure of dispersion and we could misestimate the effects of variability on costs. To investigate this further, we substitute two non-parametric measures of dispersion: the difference between the maximum and minimum daily census, and the difference between the 90th and 10th percentiles of the daily census distribution. We focus on the tails of the distribution since this is where we suspect most of the action is for costs. Results are shown in Table 2. The results are consistent with the findings so far, although the statistical significance is not as strong. Three of

the four coefficients are positive, and the one statistically significant coefficient suggests that increases in variation are associated with increases in costs.

III.D. Variance and labor costs

We now turn to a direct examination of labor expenditures. Highly variable demand may impair a hospital's ability to efficiently schedule workers and may frequently occasion either summoning additional workers or sending unneeded staff home. Both of these would tend to increase staffing costs. To investigate this, we obtained data on the total expenditures of hospitals for five categories of labor: nurses (including registered nurses and licensed vocational nurses), aids and orderlies, management and supervision, technical and specialists, and clerical workers. Total expenditures are computed as the average wage for each category of labor multiplied by the hours worked by each category. We estimated models in which the dependent variable is the log of total annual spending on each category of labor as a function of variance in the current year, and the other variables included in the models above (for simplicity, we use simple linear models here, not trans-log; we may change this as we move along and work on the theoretical underpinnings of these models.)

Results are shown in Table 4. For nursing, management and supervision, and clerical staff, the coefficient on variance is positive and significant. For technicians and specialists, the coefficient is positive but not significant. For aids and orderlies, the coefficient is negative and insignificant. In general, we interpret this as evidence in favor of the hypothesis that variance is associated with labor costs.[2]

---

[2]It is important to note that contract workers are not included in these data. Given that a common response to upward demand shocks is to bring in registry nurses and other temporary

All told, we interpret the results of this section as evidence that variance is associated with hospital costs in important ways. We infer that variance probably influences costs both through the need for standby capacity, and through inefficiencies in staffing.

IV. The Determinants of Variance

In this section, we examine whether hospital size and the presence of managed care are associated with variance. Although, in principle, there may be many determinants of variance, we focus on these since they may have significant policy implications. Hospitals interested in merging often claim that economies of scale will result from forming a larger institution. One source of economies of scale might be either the ability to reduce variance, or, separately, the ability to more efficiently adapt to variation. If large hospitals either have lower variances or they can better accommodate fluctuations, they hospitals may be able to operate with lower costs than small hospitals. Managed care plans may also have important impacts on hospital costs through manipulation of variation. If managed care plans can smooth variation in utilization in predictable ways, they may be able to help reduce hospital costs. Better understanding of this dynamic might help illuminate the implications of growth in managed care for hospitals and may also contribute to better understanding of the effects of managed care on competitiveness in the hospital sector.

Two questions can be raised here. The first is whether or not there is evidence that large hospitals or hospitals with larger amounts of HMO activity[3] have less variation than other

---

personnel, contract labor expenditures are an important area of further inquiry.

[3]We do not wish to equate HMOs with managed care here, but data constraints force us to rely on a measure of HMO discharges as our managed care measure. There are a wide range of

hospitals. To begin to examine this question, we divided hospitals into four size groups and four HMO groups based on the quantile distribution of hospitals. We then computed the average variance and average CV for hospitals in each group each year.

Results for CV are shown in Figure 6. There are substantial differences between large and small hospitals–in the largest hospitals, the standard deviation is about 10 percent of the mean on average, while it is about 40 percent of the mean in the smallest hospitals. There are similar effects across HMO groups. Hospitals with the highest HMO shares have CVs around 15, while those with the lowest shares have average CVs ranging from 25 to 55.

The second question that could be raised concerns the extent to which large hospitals and HMO dominated hospitals are better able to handle demand variations than other hospitals, so that their costs are less responsive to variance. To test this, we estimated two additional versions of the cost functions discussed above (equation 6), one including terms interacting variance with bed size and one including terms interacting variance with the percent of discharges covered by HMOs. Results are shown in columns 4 and 5 of Table 3. For bed size, the interaction terms are negative and significant, consistent with the hypothesis that large hospitals are less influenced by variance than smaller hospitals. For percent discharges HMO, the interaction terms are not significant, which leads us to believe that hospitals with large fractions of HMO patients are not necessarily better at responding to demand shocks.

With respect to bed size, the results in this section suggest that larger hospitals have less variance and are better able to respond to variance than smaller hospitals. This finding will be

---

different kinds of managed care organizations, all of which might impact variance. We do expect, though, that HMOs are one of the strongest forms of managed care, are likely to have strong effects on hospital use patterns, and thus provide a useful variable for analysis.

explored further in the context of mergers as we move the paper forward. For HMOs, we do find that hospitals with lots of HMO business tend to have less variation as a percent of their mean daily census, which may lead to cost savings, but that these hospitals do not have any advantage in their ability to respond to variability in utilization.

V. Conclusions

This paper used data on California and Florida hospitals to examine variability in daily census in hospitals. Variation in demand is an important component of hospital cost modeling, but data on within-year variations has not previously been available for a large sample of hospitals. We demonstrate that there is substantial within-year variation, and that there is also considerable variability in utilization variation across hospitals. Some of this variation should be predictable by hospital managers, but it is unlikely that all of it is.

We then examine the impact of variation in daily census on hospital costs, showing that increases in variance increase total operating expenditures and labor expenditures. Finally, we begin to discuss the determinants of variations in daily census, focusing on hospital size and managed care effects. We find some evidence that large hospitals enjoy advantages over small hospitals in this area, and that HMOs are also associated with lower variance.

In many ways, this paper is still being developed. We plan to add Florida data to the cost models, and to further refine the cost models with additional data on capital costs and other related variables. We also envision much additional work to further explore determinants of variance.

References

Gaynor, Martin and Gerard F. Anderson, "Uncertain demand, the structure of hospital costs, and the cost of empty hospital beds." *Journal of Health Economics*, 14:3 (August, 1995) 291-317.

Joskow, Paul L., "The effects of competition and regulation on hospital bed supply and the reservation quality of the hospital." *Bell Journal of Economics*, 11:2 (Autumn, 1980) 421-447.

Keeler, Theodore E. and John S. Ying, "Hospital costs and excess bed capacity: a statistical analysis," University of California at Berkeley Working Paper in Economics, No. 93-209, April, 1993.

Table 1: Some Summaries of Variation in Daily Census in California Hospitals

	N	Daily Census	Daily Census Variance	Daily Census Standard Deviation	Daily Census Coef. of Variation	Daily Census 75th pctl - 25th pctl	Daily Census 90th pctl- 10th pctl	"Predictable" Variance	Ratio of Predictable to Total Var
1984									
Mean across hospitals	524	100.0	258.6	12.8	25.1	17.5	32.8	---	---
10th percentile across hospitals	524	10.8	14.0	3.7	8.9	5	9	---	---
50th percentile across hospitals	524	60.5	102.5	10.1	16.5	14	26	---	---
90th percentile across hospitals	524	231.3	591.9	24.3	37.2	33	63	---	---
1990									
Mean across hospitals	472	100.8	228.1	12.1	27.7	16.3	30.7	111.0	0.48
10th percentile across hospitals	472	7.3	9.6	3.1	8.6	4	8	2.5	0.16
50th percentile across hospitals	472	63.4	101.1	10.1	15.5	14	26	36.9	0.39
90th percentile across hospitals	472	235.4	502.5	22.4	47.7	31	57	291.5	0.90
1994									
Mean across hospitals	471	85.1	202.5	11.4	37.7	15.0	28.5	---	---
10th percentile across hospitals	471	4.1	7.4	2.7	9.8	3	6	---	---
50th percentile across hospitals	471	56.5	97.2	9.9	17.0	13	25	---	---
90th percentile across hospitals	471	205.1	433.6	20.8	51.0	27	53	---	—

Table 2: Means and Standard Deviations of Cost Function Variables

Variable	Mean	Standard Deviation
Inpatient operating expenses (thousands)	37351.032	48459.031
variance in daily census, year t-1	240.474	566.425
log(variance in daily census, year t-1)	4.606	1.416
variance in daily census, year t	232.470	554.614
log(variance in daily cen, year t)	4.571	1.424
mean daily census, year t	102.339	114.035
log(mean daily census, year t)	4.066	1.211
mean daily census, year t-1	104.825	116.085
log(mean daily census, year t-1)	4.100	1.195
max-min census, year t	68.868	50.683
log(max-min census, year t)	3.991	0.731
90th-10th pctl, year t	31.589	21.849
log(90th-10th pctl, year t)	3.221	0.724
max-min census, year t-1	68.874	50.688
log(max-min census, year t-1)	3.991	0.731
90th-10th pctl, year t-1	31.592	21.851
log(90th-10th pctl, year t-1)	3.221	0.724
available beds	193.982	162.690
log(available beds)	4.952	0.839
mean wage	15.224	3.917
log (mean wage)	2.724	0.230
case mix index	0.991	0.180
log(case mix index)	-0.025	0.177
intermediate teaching facility	0.037	0.188
academic medical center	0.018	0.131
% discharges HMO	0.105	0.141
Year=1984	0.100	0.299
Year=1985	0.098	0.298
Year=1986	0.096	0.299
Year=1987	0.093	0.290
Year=1988	0.093	0.291
Year=1989	0.091	0.288
Year=1990	0.088	0.283
Year=1991	0.088	0.283
Year=1992	0.087	0.281
Year=1993	0.085	0.279
Year=1994	0.082	0.275
Sample size	4386	

Table 3: Main Cost Model Regression Results
dependent variable: log (total inpatient operating expenditures)

	(1)	(2)	(3)	(4)
log(daily census variance, year t-1)	**0.018****	—	**0.093***	**0.013**
	(0.006)		**(0.042)**	**(0.007)**
log(daily census variance, year t)	**0.023****	—	**0.095***	**0.025****
	(0.006)		**(0.038)**	**(0.007)**
log(max-min, year t)	—	**-0.018**	—	—
		(0.016)		
log(90th-10th pctl, year t)	—	**0.066****	—	—
		(0.016)		
log(max-min, year t-1)	—	**0.014**	—	—
		(0.014)		
log(90th-10th pctl, year t-1)	—	**0.014**	—	—
		(0.014)		
log(daily cen variance, yr t-1)*log(avl beds)	—	—	**-0.014**	—
			(0.008)	
log(daily cen variance, yr t)*log(avl beds)	—	—	**-0.014**	—
			(0.007)	
log(daily cen var t-1)*% discharges HMO	—	—	—	**0.045**
				(0.032)
log(daily cen var t)*% discharges HMO	—	—	—	**-0.014**
				(0.032)
log(average daily census, year t)	0.774**	0.763**	0.730**	0.767**
	(0.102)	(0.102)	(0.106)	(0.102)
log(avg. daily census year t)^2	0.078**	0.080**	0.078**	0.079**
	(0.009)	(0.009)	(0.009)	(0.009)
log(avg. daily census, year t-1)	0.410**	0.404**	0.327**	0.419**
	(0.106)	(0.106)	(0.112)	(0.106)
log(avg. daily census, year t-1)^2	0.024**	0.025**	0.023**	0.024**
	(0.006)	(0.006)	(0.006)	(0.006)
log(available beds)	-0.464**	-0.460**	-0.511**	-0.438**
	(0.124)	(0.123)	(0.125)	(0.125)
log(available beds)^2	0.282**	0.281**	0.290**	0.280**
	(0.018)	(0.018)	(0.019)	(0.018)
log(average wage)	0.412*	0.441*	0.422*	0.430*
	(0.180)	(0.180)	(0.180)	(0.181)
log(average wage)^2	-0.039	-0.037	-0.040	-0.035
	(0.034)	(0.034)	(0.034)	(0.034)
log(avg. daily cen t)*log(avg daily cen t-1)	0.052**	0.050**	0.055**	0.051**
	(0.011)	(0.011)	(0.012)	(0.012)
log(avg. daily census t)*log(avl. beds)	-0.291**	-0.289**	-0.280**	-0.292**
	(0.021)	(0.021)	(0.022)	(0.021)
log(avg. daily census t)*log(avg. wage)	0.143**	0.139**	0.137**	0.145**
	(0.041)	(0.041)	(0.041)	(0.041)
log(avg daily census t-1)*log(avl. beds)	-0.146**	-0.143**	-0.132**	-0.144**
	(0.022)	(0.022)	(0.023)	(0.022)
log(avg daily census t-1)*log(avg. wage)	-0.009	-0.008	-0.004	-0.013
	(0.040)	(0.040)	(0.040)	(0.040)

continued

Table 3, continued

	(1)	(2)	(3)	(4)
log(avl. beds)*log(avg. wage)	-0.109**	-0.116**	-0.110**	-0.116**
	(0.034)	(0.034)	(0.034)	(0.034)
log(case mix index)	0.382**	0.379**	0.383**	0.381**
	(0.030)	(0.030)	(0.030)	(0.030)
% discharges HMO	0.130**	0.128**	0.129**	-0.021
	(0.025)	(0.025)	(0.025)	(0.114)
intermediate teaching hospital	0.294**	0.296**	0.299**	0.298**
	(0.063)	(0.063)	(0.064)	(0.064)
academic medical center	0.762**	0.762**	0.784**	0.766**
	(0.069)	(0.069)	(0.069)	(0.069)
church owned	-0.003	-0.004	-0.004	-0.003
	(0.026)	(0.026)	(0.026)	(0.026)
private non-profit	0.013	0.014	0.013	0.014
	(0.019)	(0.019)	(0.019)	(0.019)
for-profit	-0.015	-0.016	-0.015	-0.014
	(0.025)	(0.025)	(0.025)	(0.025)
government	-0.085**	-0.083**	-0.083**	-0.084**
	(0.025)	(0.025)	(0.025)	(0.025)
intercept	12.684**	12.612**	12.724**	12.602**
	(0.357)	(0.356)	(0.357)	(0.361)
hospital fixed effects	Yes	Yes	Yes	Yes
year fixed effects	Yes	Yes	Yes	Yes
N	4377	4376	4377	4377
R^2	0.991	0.991	0.991	0.991

Note: standard errors in parentheses.

** $p<0.01$
* $p<0.05$

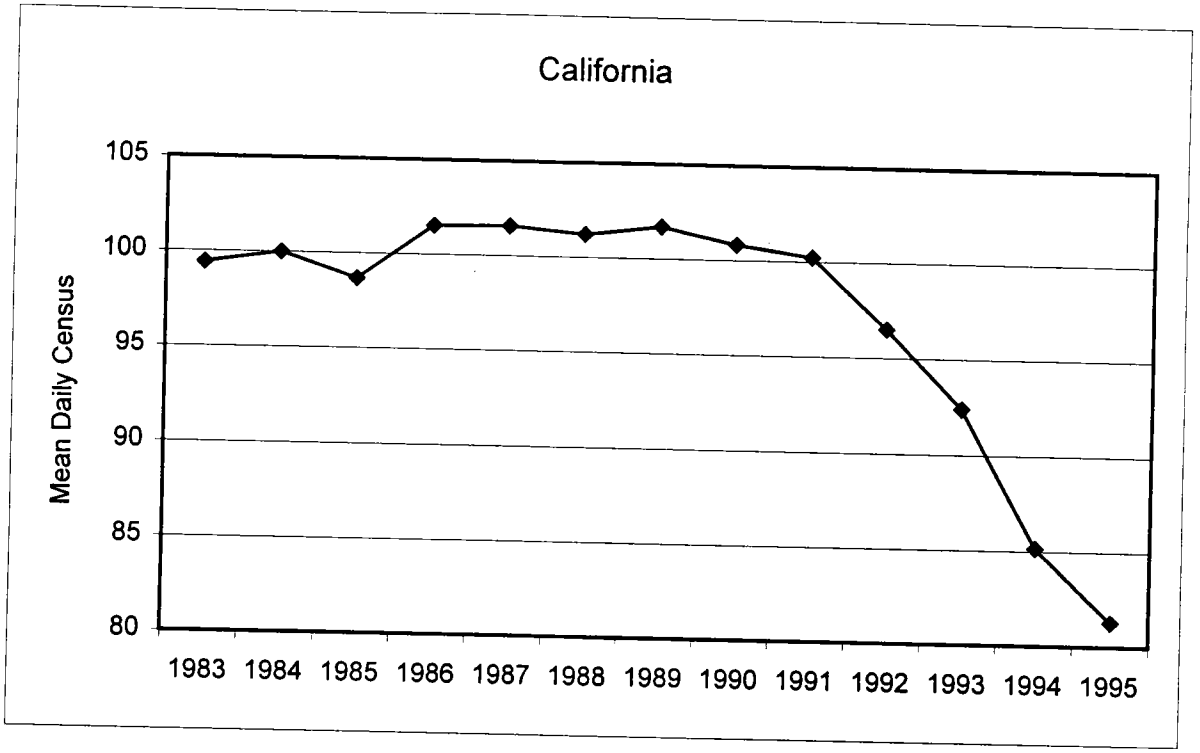
Table 4: Labor spending models

	dep var: log of total expenditures for:				
	Nursing RN+LVN	aids/ orderlies	management/ supervision	technician/ specialist	clerical
log(daily cen variance, t)	**0.038****	**-0.013**	**0.053****	**0.017**	**0.040***
	(0.008)	**(0.021)**	**(0.016)**	**(0.011)**	**(0.019)**
log(mean daily census, t)	0.246**	0.627**	0.179**	0.178**	0.313**
	(0.013)	(0.034)	(0.028)	(0.017)	(0.032)
log(mean daily census, t-1)	0.141**	0.029	0.067**	0.086**	0.199**
	(0.012)	(0.033)	(0.024)	(0.016)	(0.028)
log(available beds)	0.096**	0.303**	0.271**	0.179**	0.138**
	(0.022)	(0.059)	(0.047)	(0.030)	(0.053)
log(case mix index)	0.089	-0.032	0.138	0.090	0.424**
	(0.039)	(0.107)	(0.083)	(0.053)	(0.095)
% discharges HMO	0.212**	0.030	0.087	0.007	0.044
	(0.032)	(0.088)	(0.067)	(0.044)	(0.078)
Intermediate teaching hospital	0.413**	-0.263	2.022**	0.914**	0.519**
	(0.074)	(0.201)	(0.154)	(0.101)	(0.176)
Academic medical center	1.137**	0.900**	2.856**	1.642**	1.713**
	(0.075)	(0.205)	(0.158)	(0.103)	(0.179)
Church owned	0.040	0.145	0.078	0.157**	-0.044
	(0.030)	(0.081)	(0.064)	(0.040)	(0.075)
private not for profit	0.032	0.056	0.098	0.202**	-0.039
	(0.027)	(0.073)	(0.058)	(0.036)	(0.069)
for-profit	-0.037	-0.080	-0.025	0.181**	-0.195*
	(0.032)	(0.087)	(0.070)	(0.043)	(0.088)
Intercept	13.321**	8.951**	9.454**	12.211**	9.997**
	(0.113)	(0.308)	(0.243)	(0.155)	(0.276)
hospital fixed effects	Yes	Yes	Yes	Yes	Yes
year fixed effects	Yes	Yes	Yes	Yes	Yes
N	4082	4075	3991	4110	3959
R^2	0.983	0.889	0.919	0.972	0.941

Note: standard errors in parentheses

** p<0.01
*p<0.05

Figure 1: Average Daily Census by Year



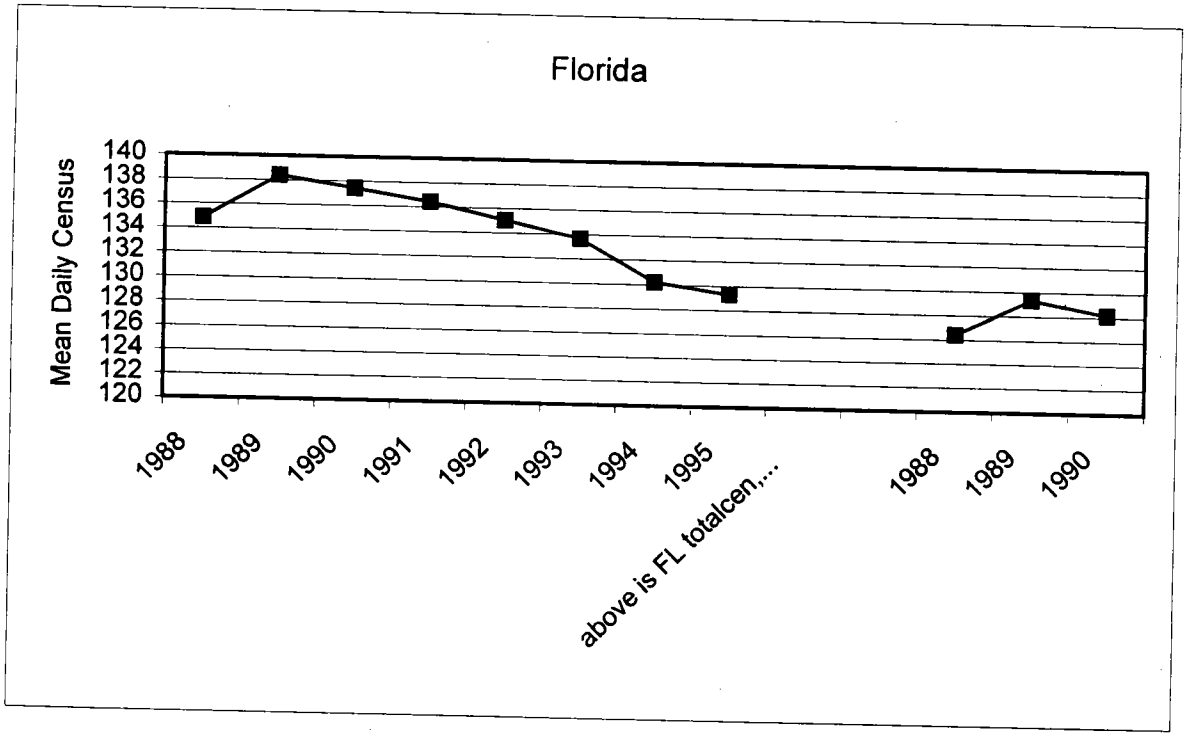
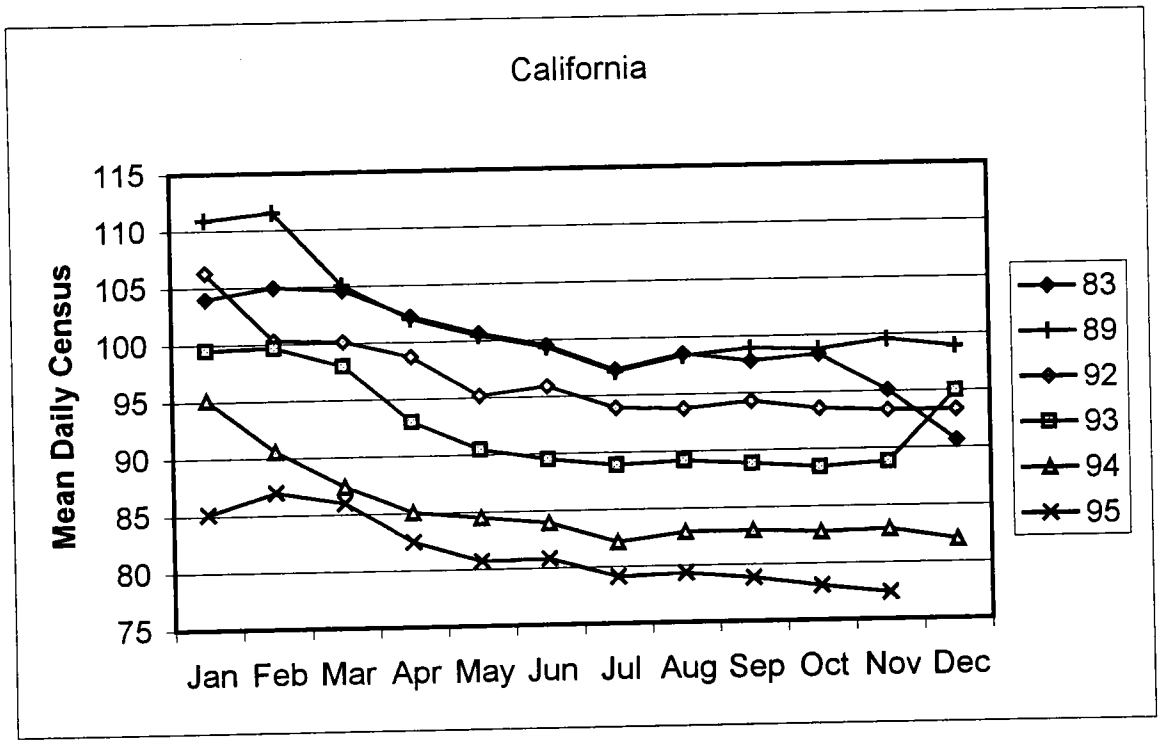


Figure 2: Average Monthly Census



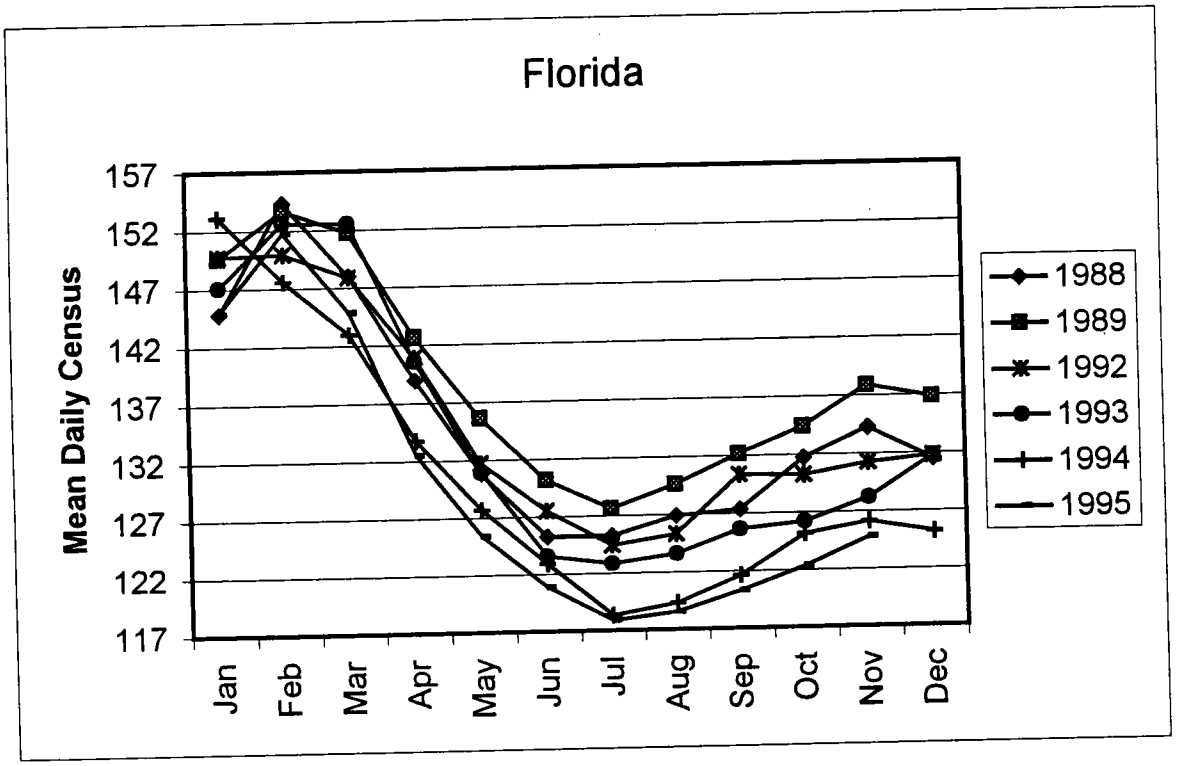
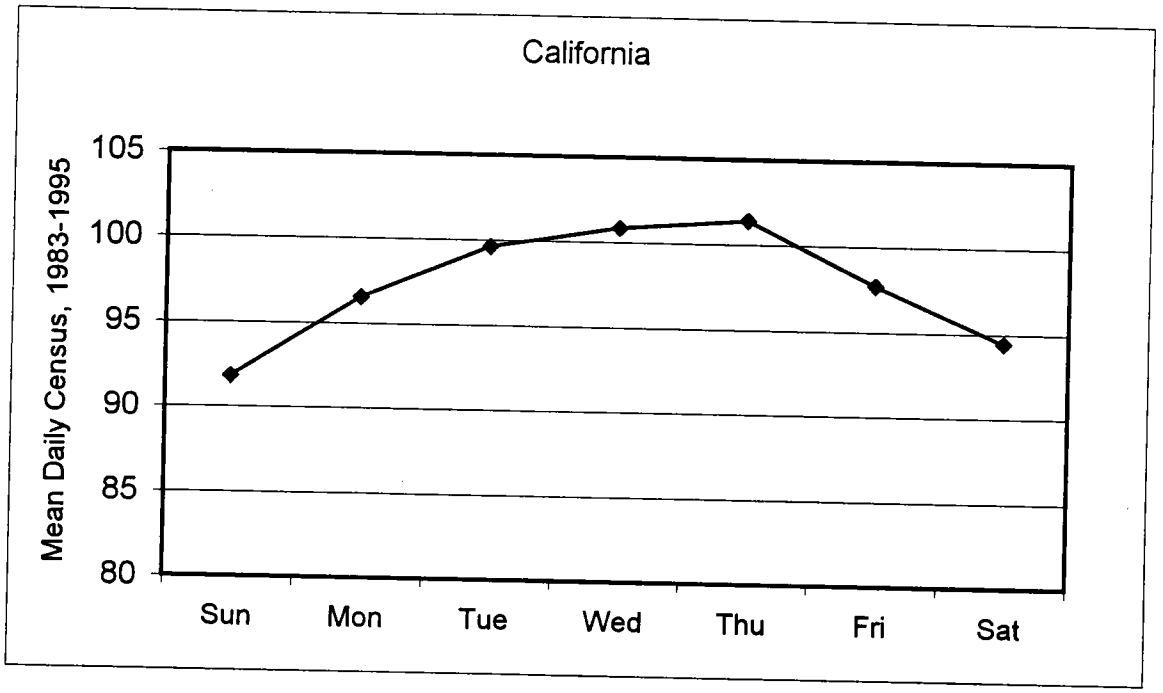


Figure 3: Mean Daily Census by Day of Week



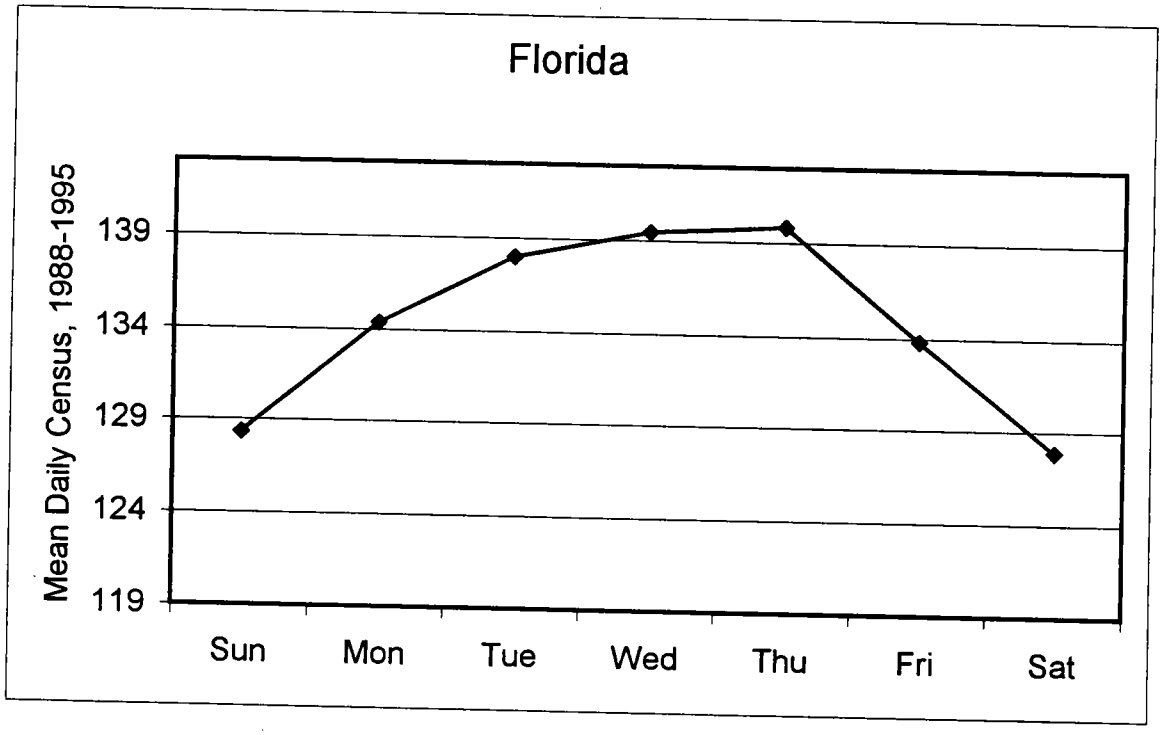


Figure 4: 1990 Daily Census Distributions for Four Selected California Hospitals

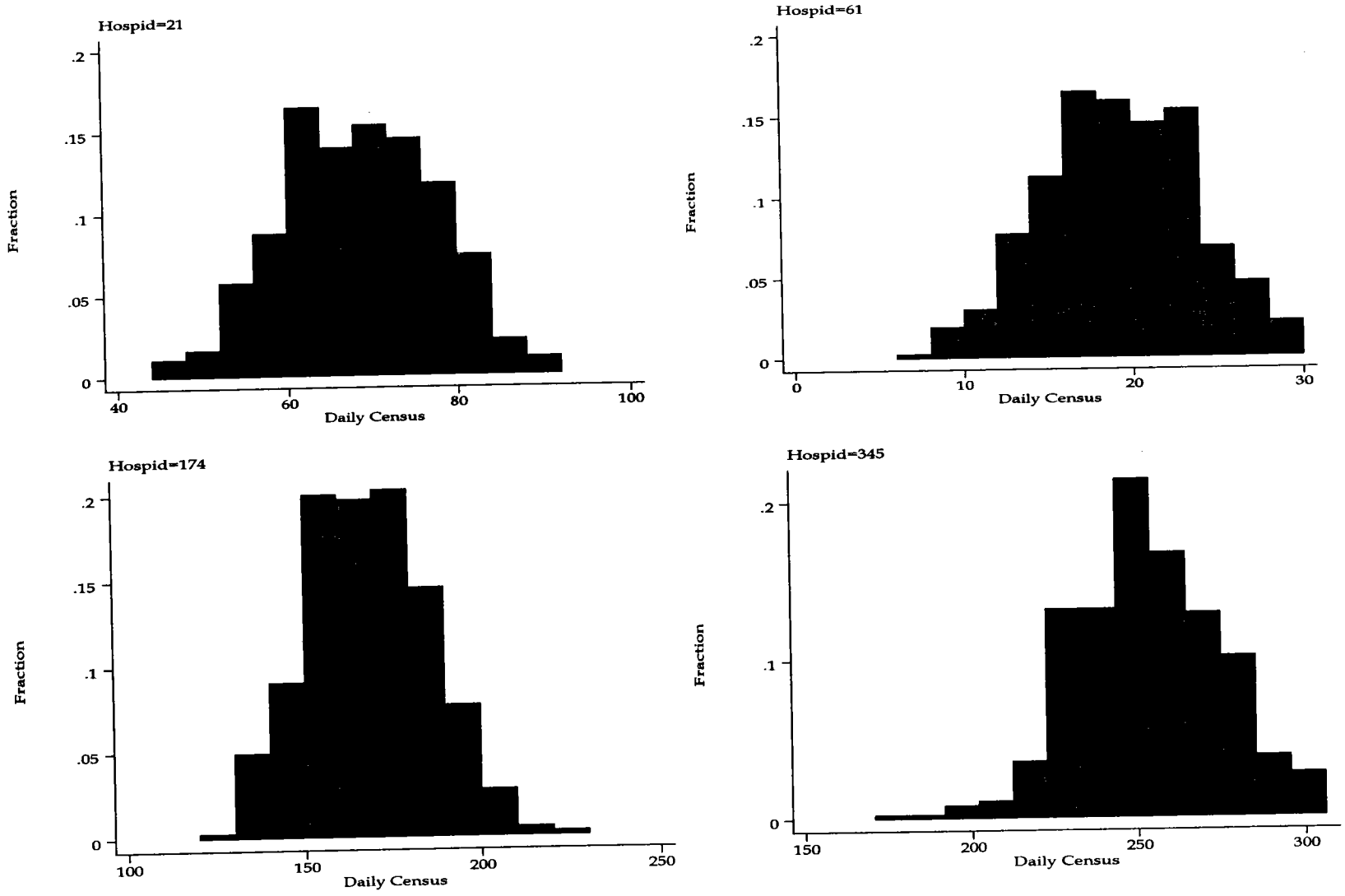
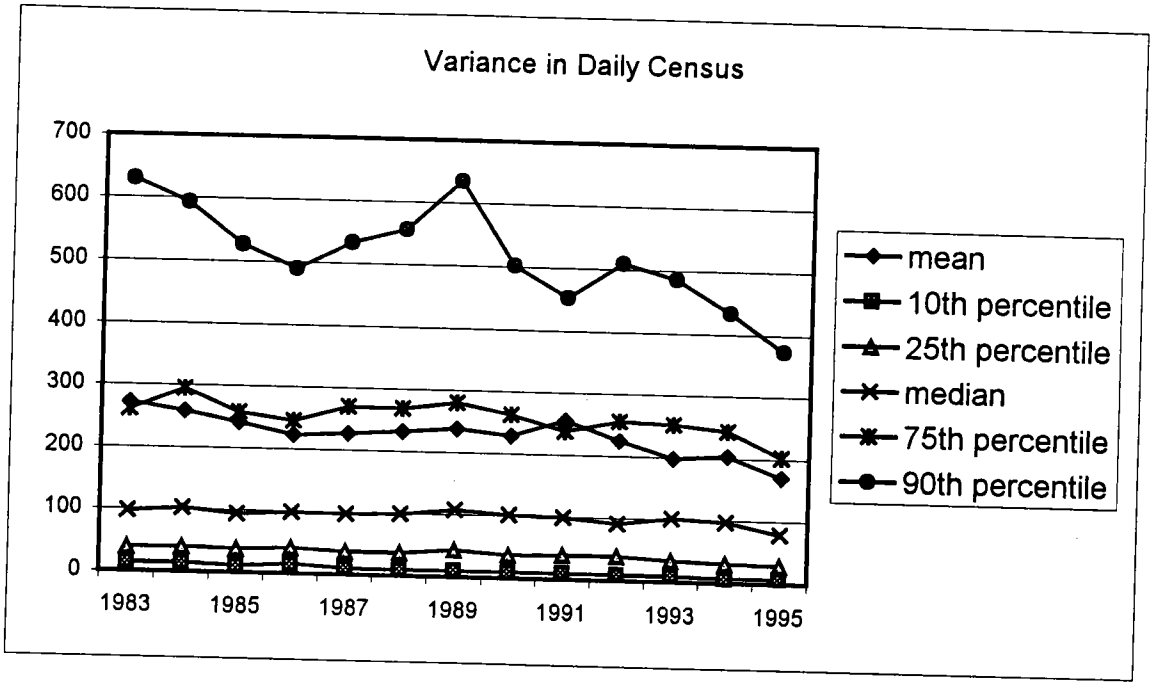


Figure 5: Variance and Coefficient of Variation, California Hospitals



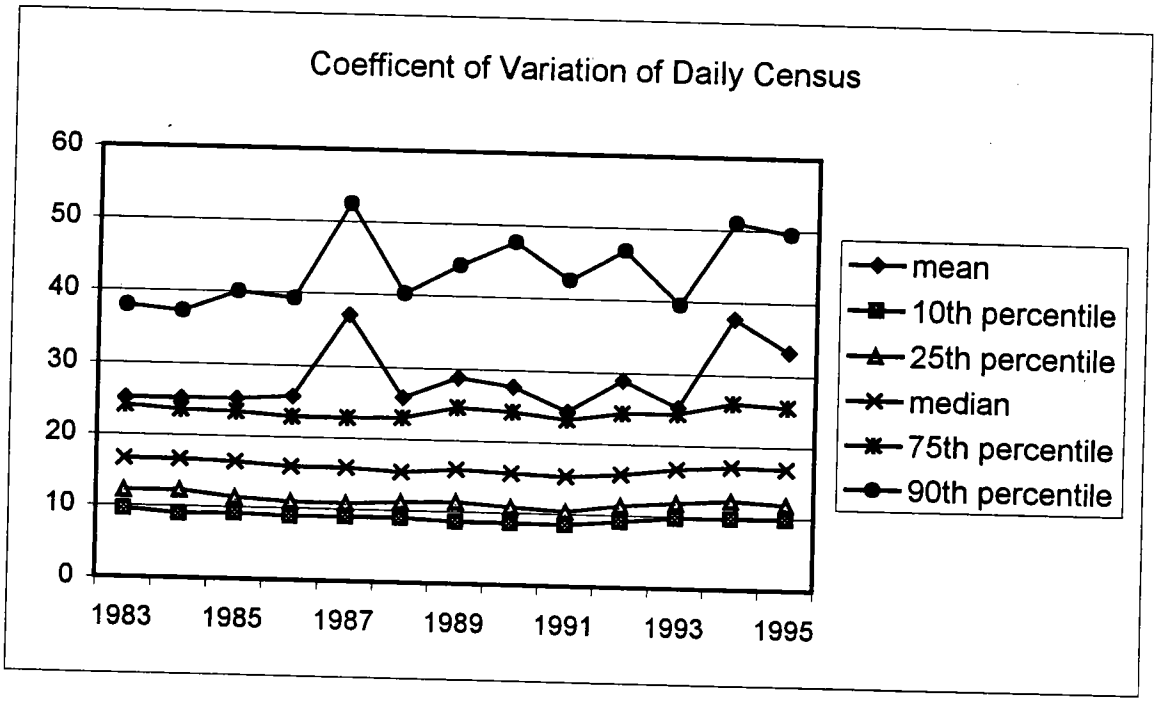
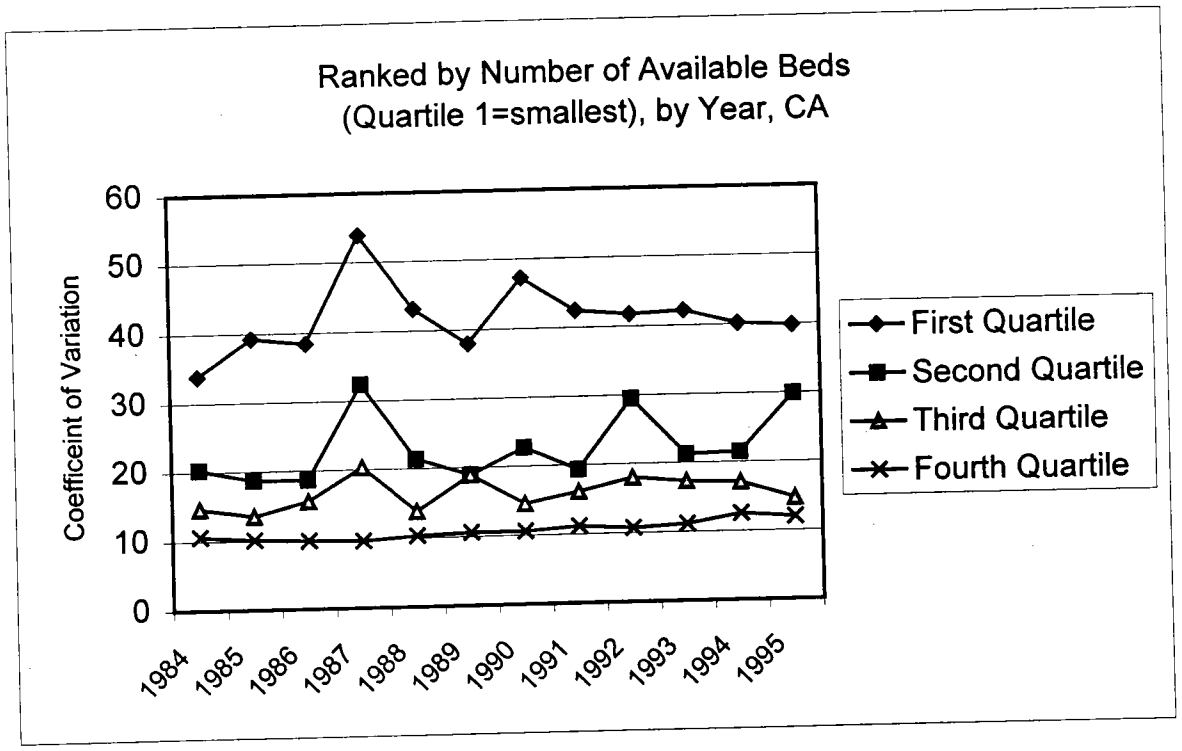
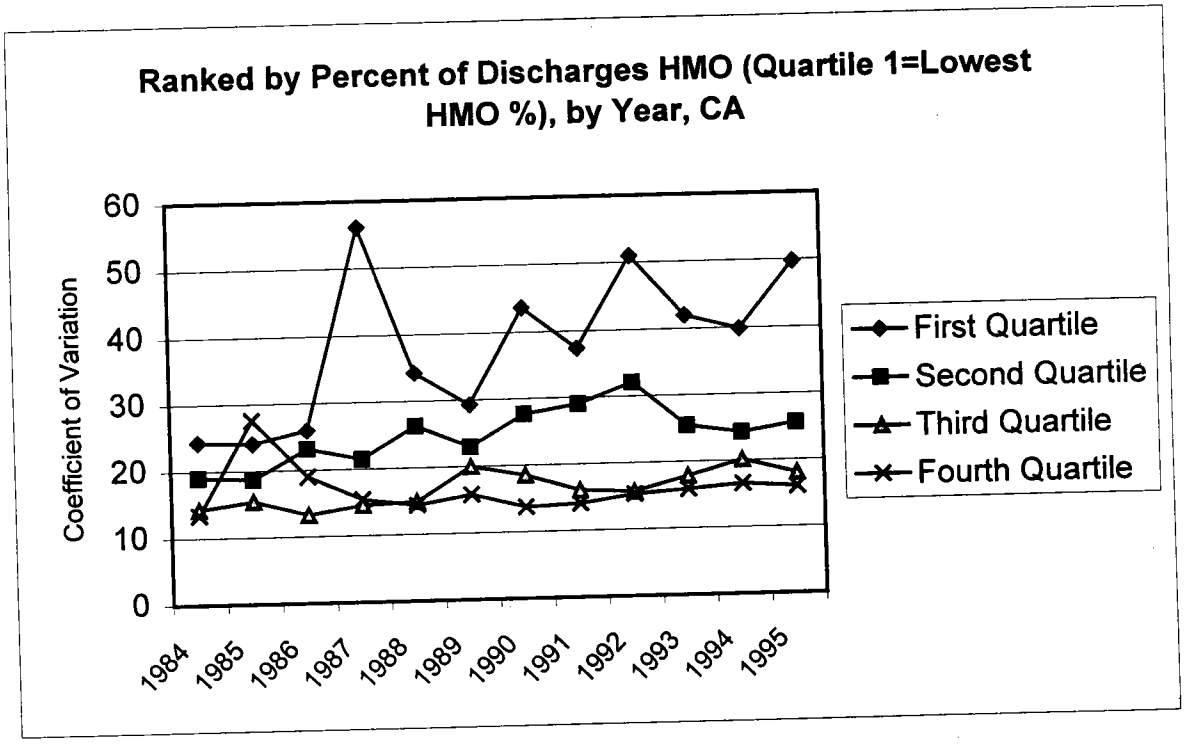


Figure 6: Coefficients of Variaion by Bed Size and HMO Quartiles





Note: N=5357 hospital-year observations