The Effects of Utilization Review on Hospital Use and Expenditures: A Covariance Analysis

Thomas M. Wickizer

Hospital utilization review (UR) has expanded rapidly in recent years and is now widely used by private payers as an approach to cost containment. This article reports estimates of the effects of UR on hospital utilization and medical expenditures based on a covariance estimation procedure. Claims data on 223 privately insured groups were analyzed covering a three-year period, 1984 through 1986. UR was associated with an approximate 12 percent decrease in admissions, a 14 percent decrease in hospital routine expenditures, and a 6 percent decrease in total medical expenditures. UR appears to reduce expenditures mainly by reducing admissions; hospital inpatient expenditures per admission were unaffected by the review activity. Analysis showed the effect of UR to have been greatest during the quarters immediately following implementation of the review activity. This finding underscores the need to analyze longitudinal data having sufficient time-series observations to obtain reliable estimates of long-term program impact. The analysis described here offers a computationally efficient alternative specification to the standard fixed-effects approach for analyzing pooled data, and is especially useful when the number of cross-section units is large.

In 1987, private health insurance payments totaled almost 158 billion dollars representing 32 cents of each health care dollar spent. As a result of accelerating insurance premiums, employers, in their role as payers of health care, sought to develop approaches for containing

The research reported here was supported by grants from the Robert Wood Johnson Foundation (#9894) and the Agency for Health Care Policy and Research (AHCPR) (HS05991).

Address correspondence and requests for reprints to Thomas M. Wickizer, Ph.D., Assistant Professor, Department of Health Services, University of Washington, Seattle, WA 98195. This article, submitted to *Health Services Research* on July 2, 1990, was revised and accepted for publication on May 16, 1991.

health care costs. Insurers have responded to employer pressure for cost containment by developing programs aimed at improving the efficiency of medical care resource consumption (Tell, Falik, and Fox 1984; Fox, Goldbeck, and Spies 1984; Herzlinger and Schwartz 1985; Payne 1987).

One of the more prominent cost-containment programs in use today is hospital utilization review (UR), which seeks to ensure that hospital care used is appropriate and necessary. UR has expanded rapidly in recent years (Gabel, Jajich-Toth, deLissovoy, et al. 1988; Ermann 1988), and 65 percent of private group insurance plans now use it as a cost-containment approach (Gabel et al. 1989). Research suggesting that 20 percent of hospital admissions and 35 percent of inpatient days may be inappropriate or unnecessary (Gertman and Restuccia 1981; Restuccia and Gertman 1984; Restuccia, Kemper, Payne, et al. 1986; Kemper 1988; Siu, Sonnenberg, Manning, et al. 1986) has led advocates to claim that UR can reduce costs (Donahue 1987; DiBlase 1986) and, at the same time, improve quality.

Evaluations of UR programs have produced a wide range of findings (Wickizer 1990). Early evaluations of professional standards review organizations (PSRO) prototype programs and other public UR programs produced findings suggesting that UR could reduce hospital inpatient rates by 10 to 15 percent, with attendant cost savings (Brian 1972, 1973; Westphal, Frazier, and Miller 1979; Flashner et al. 1973). However, later macro evaluations of the PSRO program found that UR had only a limited effect on reducing hospital days among Medicare patients (Dobson, Greer, Carlson, et al. 1978; Health Care Financing Administration 1979, 1980), with no net cost savings (Congressional Budget Office 1979, 1981). More recent assessments of private UR programs have produced some encouraging findings suggesting that UR can generate impressive cost savings (Richards 1984; Shahoda 1984; Ness 1985; O'Donnell 1987) but, like many earlier evaluations (Sayetta 1976), these assessments suffer from methodological problems that make their findings suspect (Wickizer 1990).

Analyses conducted by the present author and colleagues have shown that hospital inpatient UR can be effective, reducing hospital admissions by approximately 10 to 15 percent and inpatient hospital expenditures by 5 to 10 percent (Feldstein, Wickizer, and Wheeler 1988; Wickizer, Wheeler, and Feldstein 1989). This article reports estimates of UR effects based on a different statistical model that uses a covariance estimation approach. It evaluates outcome measures not previously examined and analyzes the effects of UR over time. The estimation approach used offers a computationally efficient method for analyzing pooled data.

UR PROGRAM OPERATION AND INCENTIVES

The UR program analyzed, established in 1983 by a private commercial insurance carrier, was made available to policyholders as a benefit plan option. Described in more detail elsewhere (Wickizer 1989), the UR program consists of two related review activities compulsory for all employees and dependents: preadmission authorization and concurrent review. The former focuses on ensuring the appropriateness of admission, the latter reviews the need for continued stay and certifies the patient's treatment plan.

Patients failing to comply with UR procedures are subject to financial penalties (physicians and hospitals are not subject to penalties). For example, employees who do not get authorization for admission as required by preadmission review may have their covered inpatient expenses reduced by some percentage (often 50 percent) or by a set dollar amount before normal benefits are applied. In effect, UR raises the relative price of hospital care judged inappropriate or unnecessary. Economic theory predicts that, as the price of a good or service increases, the quantity demanded decreases (more formally, $\partial Q/\partial P < 0$), and research has shown that the demand for medical care is sensitive to out-of-pocket price (Scheffler 1984; Newhouse and Phelps 1976; Manning, Newhouse, Duan, et al. 1987). Thus, one would expect UR to have a negative effect on the demand for inpatient care, *ceteris paribus*. By raising the price of inappropriate or unnecessary care, UR seeks to reduce the subsidy distortion caused by insurance (Pauly 1968; Dionne 1981; Dowd 1982) and thereby improve medical care consumption efficiency.

METHODS

DATA AND MEASURES

The unit of analysis for this study was the insured group. A group may represent an entire insured population covered under an employer's health insurance policy, or some subset of the employees and dependents covered under that policy, as with companies having offices or plants in different locations. The study population comprised groups that had active policies with the insurer as of the second quarter of 1985. All groups meeting the following criteria were selected for study: (1) claims data were available for a minimum of three quarters; (2) population data were available for employees and dependents; (3) the policy covered hospital and physician services; and (4) the group had a minimum of 150 insureds (employees and dependents).

The 223 groups that met these criteria were selected for study. Of these groups, 91 operated under UR for some or all of the period 1984–1986, and the remaining 132 did not have any cost-containment program in effect during this time. The study groups were well distributed in terms of geographic region and industry, and were representative of the under-65 general population in terms of hospital use. Eighteen percent of the groups were located in the Northeast, 35 percent in the North Central region, 26 percent in the South, and 21 percent in the West. The average group included approximately 1,500 insureds: 660 employees and 840 dependents.

Quarterly claims data for 1984–1986 on the 223 groups were pooled, making 1,844 observations available for analysis. The claims data provided information used to construct a set of utilization and expenditure measures, which formed the outcome variables for the analysis (see Table 1). The utilization variables included admissions per 1,000 insureds and length of stay. The expenditure measures, which represent total charges adjusted to reflect January 1985 prices, included: (1) hospital routine (room and board) expenditures per insured person; (2) hospital ancillary expenditures per insured person; (3) hospital inpatient expenditures per admission; (4) expenditures on physician hospital visits per insured person; and (5) total medical expenditures per insured person.

For the analysis of main effects, UR was measured as a binary variable, equal to 1 if a group operated under UR during a given quarter and 0 otherwise. Since most groups adopted preadmission authorization and concurrent review together, it was not possible to estimate the independent effects of these two review activities. Thus, the UR variable represents the aggregate effects of the two review procedures. It was expected that UR would have a negative effect on utilization and hence on expenditures.

For analysis performed to evaluate the effects of UR over time, the UR binary variable was respecified as a set of dummy variables where the subscript indicates the quarter of UR operation: $UR_1 = 1$ if quarter of UR implementation and 0 otherwise, $UR_{2,3} = 1$ if quarters 2 or 3 of UR operation and 0 otherwise, $UR_{4,5} = 1$ if quarters 4 or 5 of

UR operation and 0 otherwise, and so forth. (Unlike a quadratic model, this specification imposes no strong assumptions about the nature of UR's effect over time.) Anecdotal information suggests that the impact of UR may diminish with time. Specifying UR as a set of dummy variables representing different operational periods permits the analysis to examine this question empirically.

To control for the effects of external factors, a set of covariates representing employee demographic characteristics, health care market factors, and benefit plan features was included in the analysis (see Table 1). Six variables that measured the age-sex distribution of employees were intended to serve as indicators of age-associated morbidity and as proxy measures for factors related to age and sex that might affect people's use of medical care. Two variables served as proxy measures to control for differences in health status and population characteristics: expenditures on services related to childbirth/pregnancy and on services related to heart disease, both expressed as a percentage of total expenditures.

A small percentage of insureds in some groups who may be retired or employed but over 65 are eligible for Medicare benefits. (These persons are subject to all UR procedures.) To control for the effects of use among Medicare-eligible persons, the analysis included a variable representing payments made by Medicare on behalf of eligible insureds, expressed as a percentage of total medical expenditures. (The available data did not allow age to be broken down in a way that would directly capture use by Medicare eligibles: hence the need for the additional Medicare expenditure variable.)

Health care market factors have been found to be strong predictors of utilization (Wilson and Tedeschi 1984; McLaughlin, Merrill, and Freed 1984; Chiswick 1976). Although the insured groups analyzed represent only very small segments of the market, their utilization patterns may be influenced by broader market forces. Therefore, four variables measuring different characteristics of the health care market, defined as the metropolitan statistical area (MSA), were included in the analysis. These variables represent the penetration rate of health maintenance organization(s) (HMO), the number of practicing physicians per 1,000 population, the percentage of physicians who are general practitioners, and the hospital occupancy rate. It was expected that groups located in markets with higher HMO penetration rates would, *ceteris paribus,* have lower hospital use rates. Groups located in markets that had more physicians per population were expected to have higher hospital use rates.

Finally, to control for differences in benefit plan features, six

Table 1: Descriptive Information on Regression Variables; Pooled Quarterly Data on 223 Insured Groups, Years 1984–1986 (N = 1,844)

*Measure**	*Mean (Standard Deviation)*
Outcome Measures	
Admissions per 1,000 insured persons	26.27 (13.47)
Average length of stay	6.23 (3.79)
Hospital room and board expenditures per insured person	$41.84† ($36.22)
Hospital ancillary expenditures per insured person	$58.81 ($52.24)
Hospital expenditures per admission	$3,923.50 ($2,566.10)
Expenditures on physician hospital visits per insured person	$8.84 ($7.62)
Total medical expenditures per insured person	$216.89 ($117.24)
Covariates	
Percent male employees under 30	11.1% (6.0%)
Percent female employees under 30	9.1% (6.9%)
Percent male employees 30 to 50	36.5% (9.5%)
Percent female employees 30 to 50	17.7% (8.3%)
Percent male employees over 50	18.6% (9.9%)
Percent female employees over 50	6.7% (3.6%)
Expenditures on childbirth and pregnancy as a percent of total expenditures	7.8% (7.7%)
Expenditures on heart disease as a percent of total expenditures	(3.9%) (7.3%)
Medicare coordination of benefit savings as a percent of total expenditures	1.2% (4.2%)

Continued

Table 1: Continued

*Measure**	*Mean (Standard Deviation)*
HMO penetration rate (%)	11.9% (12.9%)
Practicing physicians per 1,000 population	2.1 (0.8)
Percent of general practice physicians	13.3% (10.1%)
Hospital occupancy	66.7% (8.9%)
Deductible level	$130.00 ($52.65)
Coinsurance rate‡	80.0% (3.8%)
Mental health inpatient coinsurance rate‡	82.0% (20.1%)
Percent expenditures covered for semiprivate hospital room	88.7% (26.7%)
Percent expenditures covered for office visits	86.3% (17.7%)
Percent expenditures covered for hospital outpatient services	89.8% (20.6%)
UR (1 = UR, 0 otherwise)	0.41 (0.49)

*All variables are measured on a per quarter basis.

†Expenditure data are adjusted to reflect January 1985 prices.

‡Represents portion of charges paid by insurance.

covariates were included in the analysis: three representing deductible level, coinsurance rate, and mental health coinsurance rate; and three measuring the percentage of expenditures covered for hospital semiprivate room, physician office visits, and hospital outpatient department services. In general, it was expected that groups with more comprehensive coverage for hospital care would, *ceteris paribus,* have higher utilization and expenditures.

EMPIRICAL MODEL

Many prior UR studies have used cross-sectional observational data to analyze the effects of UR. Such studies rely on between-group variance to estimate relationships, a reliance that imposes some important limitations on the analysis. It is rarely possible to collect data on all of the important determinants of utilization. Unmeasured factors, if correlated with explanatory variables, may lead to biased estimates. To the extent that omitted factors are group specific and constant over time, they can be summarized by a group-specific intercept. With cross-sectional data, inclusion of an intercept (a dummy term) for each group would be impossible, since there would be one "variable" for each observation. However, with longitudinal data, it is possible to include an intercept for each group.

Given the availability of longitudinal (pooled) data, there are several estimation approaches that one can use. One option is to estimate the model using ordinary least-squares (OLS) regression without including cross-section intercepts, which assumes no variance across groups to exist in the effects of omitted factors. This approach was used by Feldstein, Wickizer, and Wheeler (1988) and Wickizer, Wheeler, and Feldstein (1989). While the estimates generated by the OLS pooled regression model are consistent, the *t*-statistics may be incorrect (Pindyck and Rubinfeld 1981).

A less restrictive approach is to recognize that omitted factors are likely to exist and to have effects that do vary across groups. Here two estimation approaches are available: the covariance model, also known as the fixed-effects or least-squares dummy variable model (LSDV), and the error components (random-effects) model (Kmenta 1986). The rationale for using the former is the belief that omitted variables exert a systematic effect causing cross-section intercepts to shift, while the latter model assumes that omitted factors exert a random effect. However, the error components model forces one to assume that the cross-section component of the error term is uncorrelated with any included variable, an assumption that is often untenable (Solon 1984; Kmenta 1986, 634). The error components model was viewed as inappropriate for the present analysis because it required the strong assumption that omitted cross-section factors were uncorrelated with UR status.

The standard covariance (fixed-effects) model requires that dummy variables be included in the model for each cross-section unit. Since a large number of degrees of freedom were available, it would have been possible to use this approach for the analysis. However, the task of inverting a data matrix with over 250 variables would have added a very measurable computational cost to the analysis. For studies analyzing pooled data on individuals where the number of cases is often large, exceeding 1,000 to 2,000, the use of the fixed-effects approach may be problematic and prohibitively expensive.

There is an alternative approach to the standard fixed-effects model that yields the same results as this model but does not require inclusion of dummy variables (Brown 1980; Kmenta 1986, 630–33). This specification exploits the "within" component of the variance present in pooled data, that is, the variance over time around cross-sectional averages. The major advantage of the alternative approach is computational efficiency; the estimation procedure uses the same number of degrees of freedom as the standard fixed-effects approach.

The basic idea behind the model is to difference out the effects of cross-section omitted factors instead of including them in the model as parameters to be estimated. The general form of the model used to evaluate the effects of UR is shown below:

$$Y_{it} - \bar{Y}_i = a + B'(X_{it} - X_i) + C'T_t + D(UR_{it} - \overline{UR}_i) + (E_{it} - \bar{E}_i) \quad (1)$$

where i is the insured group; t is the quarter; $\underline{Y}$ is a vector of utilization or expenditure variables measured in logarithms; $\underline{X}$ is a vector of covariates; $\underline{T}$ is a vector of quarter dummy variables (entered in conventional form); UR is a binary variable representing utilization review; and E is a normally distributed error term, with mean zero and variance δ_i^2. The time means are calculated such that for any variable x_{it}, $\bar{X}_i = \Sigma_1^T x_{it}/T$.

As indicated, the model uses longitudinal variance to generate coefficient estimates. Of the 91 groups that operated under UR, 45 adopted UR during the study period, mainly during the latter part of 1984 and early 1985, and exhibit such variance. The estimates of UR effects are based on information provided by these groups. The remaining 46 groups had UR in effect throughout the study period, and thus do not exhibit longitudinal variance. These groups, however, do provide information for estimating the coefficients of other variables in the model.

Because aggregate data collected over time were being analyzed on groups differing in size, specification tests were performed to determine whether the model's error terms exhibited heteroscedasticity or serial correlation. The Goldfeld-Quandt (1965) test showed the error terms to be heteroscedastic, with the error variance inversely related to group size. A standard weighting procedure was performed on the data to give more weight in the analysis to larger groups, and the Goldfeld-

Quandt test was then replicated using the weighted data to verify that the correct weighting factor had been selected.

Since repeated measures were taken on groups over time, it could not be assumed that the model's error terms were independent, that is, $\mathrm{Cov}(E_{it}, E_{it-s}) = 0,\ s \neq t$ (Nickell 1982; MaCurdy 1982). However, conducting tests to determine serial correlation when pooled data are being analyzed presents special problems because the usual autocorrelation estimators are not consistent as $N \rightarrow \infty$ but T remains small. Solon (1984) developed a procedure for obtaining corrected estimates of the autocorrelation (ρk) between E_{it} and $E_{i,t-k}$ for fixed-effects covariance models. The autocorrelation estimates generated by this procedure for $k = 1, 2$, and 3 were all close to zero (range 0.003 to 0.125 in absolute value), indicating the model's error terms were not serially correlated. Therefore, the model described by Equation 1 should have the desirable properties needed to generate unbiased and efficient parameter estimates.

RESULTS

The results of the analysis are summarized in Tables 2 and 3. Table 2 presents findings regarding UR's main effects; Table 3 shows the results of analysis of UR's effects over time.

MAIN EFFECTS

As Table 2 shows,[1] UR had a statistically significant negative effect on admissions. On average, UR groups had approximately 11.5 percent fewer admissions after adopting UR ($p < .001$). This percentage decrease translates into an average annual reduction of approximately 16 admissions per 1,000 insureds. Although negatively signed, the UR coefficient for the length-of-stay equation failed to achieve statistical significance, possibly because of case-mix effects associated with the preadmission review component of UR. If preadmission review keeps persons with less serious illness out of the hospital, and if case mix is not controlled, then the estimated effect of UR on length of stay may be biased toward finding no effect (Health Care Financing Administration 1980).

Table 2: The Effects of Utilization Review on Hospital Use and Expenditures; Covariance Analysis of Pooled Quarterly Data on 223 Insured Groups, 1984–1986 (N = 1,844)

Measure†	*Admissions*‡	*Length of Stay*	*Hospital Routine Expenditures*	*Hospital Ancillary Expenditures*	*Hospital Expenditures per Admission*	*Hospital Physician Expenditures*	*Total Medical Expenditures*
Intercept	−0.014	0.000	−0.023	−0.031	−0.022	−0.252****	−0.047**
	(0.022)§	(0.019)	(0.033)	(0.033)	(0.019)	(0.043)	(0.019)
Percent females under 30††	−0.115**	0.011	−0.065	0.023	0.101**	0.020	−0.044
	(0.051)	(0.044)	(0.077)	(0.077)	(0.044)	(0.100)	(0.044)
Percent males 30 to 50	0.005	0.007	0.019	0.033	0.024*	0.059*	0.014
	(0.018)	(0.015)	(0.027)	(0.027)	(0.015)	(0.034)	(0.014)
Percent females 30 to 50	0.079**	−0.013	0.035	0.006	−0.054*	−0.044	0.044
	(0.036)	(0.031)	(0.054)	(0.055)	(0.031)	(0.070)	(0.031)
Percent males over 50	−0.009	0.004	0.003	0.018	0.021****	0.018	0.008
	(0.007)	(0.006)	(0.012)	(0.012)	(0.006)	(0.015)	(0.006)
Percent females over 50	−0.089	0.027	0.029	0.061	0.098	0.226	−0.021
	(0.083)	(0.070)	(0.123)	(0.124)	(0.071)	(0.161)	(0.071)
Percent childbirth/pregnancy expenditures	0.443**	−0.811****	−0.154	−0.644**	−0.940****	0.192***	−0.427***
	(0.180)	(0.152)	(0.268)	(0.271)	(0.153)	(0.070)	(0.154)
Percent heart disease expenditures	0.363**	0.170	0.591**	1.781****	0.847****	0.263****	0.766****
	(0.177)	(0.150)	(0.263)	(0.266)	(0.151)	(0.065)	(0.152)
Percent Medicare savings	0.321	−0.127	0.610	0.888**	0.385	0.402****	0.374
	(0.287)	(0.243)	(0.428)	(0.432)	(0.245)	(0.065)	(0.246)
HMO penetration rate	−0.017****	0.000	−0.022****	−0.009*	0.001	0.329****	−0.010****
	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.064)	(0.003)
Physicians per 1,000 population	0.258	−0.035	0.540**	0.639***	0.312**	0.338****	0.497****
	(0.164)	(0.139)	(0.245)	(0.247)	(0.140)	(0.060)	(0.141)
Percent general practice physicians	0.793**	0.346	1.308**	1.086**	0.334	0.829**	0.783**
	(0.355)	(0.301)	(0.530)	(0.535)	(0.303)	(0.342)	(0.305)

Continued

Table 2: Continued

Measure†	*Admissions*‡	*Length of Stay*	*Hospital Routine Expenditures*	*Hospital Ancillary Expenditures*	*Hospital Expenditures per Admission*	*Hospital Physician Expenditures*	*Total Medical Expenditures*
Occupancy	-0.011****	-0.002	-0.013***	-0.006	-0.000	0.309****	-0.015****
	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)	(0.061)	(0.003)
Percent coverage for semiprivate room	0.001	0.001	-0.004****	-0.003***	-0.004****	0.245****	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.063)	(0.001)
Deductible	0.028**	-0.019****	0.054***	0.021	0.006	0.249****	0.022**
	(0.012)	(0.001)	(0.018)	(0.018)	(0.011)	(0.064)	(0.011)
Coinsurance rate	0.492	0.478	1.086	0.014	0.403	0.241****	0.333
	(0.710)	(0.602)	(1.061)	(1.071)	(0.608)	(0.065)	(0.611)
Mental health coinsurance rate	-0.015	-0.003	-0.199****	-0.016	0.007	0.228****	-0.019
	(0.014)	(0.012)	(0.020)	(0.020)	(0.012)	(0.065)	(0.012)
Percent coverage for office visit	0.000	0.000	-0.001	0.000	0.001	0.204****	-0.002**
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.065)	(0.001)
Percent coverage for hospital outpatient department care	-0.001	0.000	-0.001	-0.001	-0.000	-0.623*	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.347)	(0.001)
Utilization review (UR)	-0.122****	-0.033	-0.149***	-0.104*	0.015	-0.098	-0.066**
	(0.045)	(0.039)	(0.067)	(0.068)	(0.039)	(0.088)	(0.039)

*$p < .10$; **$p < .05$; ***$p < .01$; ****$p < .001$.

†The t-statistics for UR are based on one-tailed tests, all other t-statistics are based on two-tailed tests. Table 2 does not show R^2s because this does not provide a meaningful measure of the model's overall explanatory power for mean-differenced equations.

‡The dependent variables are measured in logarithms, so the UR coefficients can be interpreted as approximate percentage differences.

§Standard errors, adjusted to reflect the loss of degrees of freedom resulting from the means-differencing procedure, are in parentheses.

††The reference variable is MUN30% (percent of males under 30).

Table 3: Abridged Results of Analyses of the Effects of Utilization Review over Time ($N = 1,844$)

Measure†	*Admissions*	*Length of Stay*	*Hospital Routine Expenditures*	*Hospital Ancillary Expenditures*	*Total Medical Expenditures*
UR_1	-0.242****	-0.085**	-0.323****	-0.277****	-0.222****
	(0.060)‡	(0.044)	(0.083)	(0.082)	(0.044)
$UR_{2,3}$	-0.091**	-0.045	-0.127**	-0.089	-0.054*
	(0.055)	(0.041)	(0.077)	(0.076)	(0.041)
$UR_{4,5}$	-0.053	-0.014	-0.043	0.018	0.019
	(0.058)	(0.043)	(0.081)	(0.080)	(0.043)
$UR_{6,7}$	-0.110**	0.039	-0.011	0.009	0.055
	(0.062)	(0.046)	(0.087)	(0.086)	(0.045)
$UR_{8,9}$	-0.087	0.016	-0.080	0.017	0.033
	(0.071)	(0.052)	(0.100)	(0.098)	(0.051)
UR_{10-12}	-0.100	-0.137***	-0.195**	-0.081	-0.060
	(0.083)	(0.061)	(0.117)	(0.116)	(0.061)

*$p < .10$; **$p < .05$; ***$p < .01$; ****$p < .001$.

†The t-statistics based on one-tailed tests.

‡Standard errors, adjusted to reflect the loss of degree of freedom resulting from means-differencing procedure, are in parentheses.

The effect of UR on expenditures is summarized in five equations shown in Table 2. Expenditures on routine and ancillary hospital services were, on average, 13.8 ($p < .05$) and 9.9 percent ($p < .10$) lower, respectively, for groups after they adopted UR. These percentage differences translate into a combined absolute difference of approximately $47 (January 1985 prices) per insured person per year.

The above estimates suggest that UR reduces hospital inpatient expenditures principally by reducing admissions through preadmission authorization. To explore whether or not review activities influenced the expenditures of patients admitted to the hospital, UR's effect on hospital inpatient expenditures per admission was estimated. As Table 2 shows, the estimated coefficient of the UR term in the expenditure-per-admission equation is not significantly different from zero. However, it is possible that this finding, like that for length of stay, reflects unmeasured case-mix effects associated with preadmission authorization.

To the degree that UR reduces hospital use, it should lower expenditures on physician inpatient hospital visits. The fourth expenditure equation in Table 2 summarizes information about the effect of UR on hospital physician expenditures. The estimated coefficient of the UR term is -0.098, indicating that UR was associated with an approximately 9 percent reduction in expenditures on hospital physician visits, on average. However, the estimate fails to achieve statistical significance because of the large standard error.

The last equation shown in Table 2 provides information on UR's impact on total aggregate medical expenditures. Adoption of UR was associated with an estimated 6.4 percent reduction in total medical expenditures ($p < .05$). When translated into a dollar figure, this percentage difference represents a reduced expenditure level of approximately $56 (1985 prices) per insured person per year.

The effects of the covariate factors were generally in the expected direction. The variables included in the model to control for differences in health status and population characteristics had a significant effect on both utilization and expenditures. Groups incurring a higher proportion of expenditures on services related to childbirth/pregnancy had, on average, significantly higher admission rates ($p < .01$), shorter average lengths of stay ($p < .001$), and lower hospital expenditures per admission ($p < .001$) as well as total expenditures ($p < .01$). Groups incurring a higher proportion of medical expenditures on services related to heart disease had significantly higher utilization as well as expenditures. Further, market characteristics appear to have exerted an important influence on the utilization and expenditure rates of the study groups. HMO penetration exerted a negative effect on both admissions and expenditures ($p < .001$), while physician supply had the opposite effect.

The variables representing benefit plan features present a mixed picture of results. The coefficients of most of the variables did not differ significantly from zero and those that did were small in magnitude. Contrary to expectations, the deductible coefficient for the admissions equation was positive. Groups with high admission rates might have been more likely to have raised deductibles. If so, the finding could reflect a potential problem of simultaneity.

EFFECTS OVER TIME

Table 3 presents abridged results of the analysis performed to evaluate the effects of UR over time. (Since the change in specification of the UR variable had little effect on the covariate estimates, these estimates are not presented in Table 3.) The table shows the estimated coefficients of the UR variables for five equations representing admissions, length of stay, hospital routine expenditures, hospital ancillary expenditures, and total medical expenditures. The coefficients shown in the table provide information indicating the approximate percentage changes in utilization and expenditures that are associated with different periods of UR operation relative to the pre-UR period.

As shown, UR appears to lead to an initial reduction in admissions and expenditures of approximately 20 to 30 percent, as indicated by the estimated coefficients of the UR_1 term. In three equations, the coefficients of $UR_{2,3}$ term are also statistically significant, although smaller in absolute value. The estimated coefficients representing year three of operation (UR_{10-12}) are also negative and are reasonably large in magnitude, although only two of these coefficients achieve statistical significance. While there is some evidence of a "decay effect," the pattern is not well defined; that is, the UR coefficients do not decrease in magnitude monotonically over time.

The large reduction in hospital use observed during the initial implementation quarter may reflect patient uncertainty about UR and its authorization procedures. This uncertainty may have led to a decision by patients to delay seeking care for elective surgical procedures. Wickizer (1991) found UR to have its greatest impact in the surgical area. It is also possible that over time providers "learned" to adjust to the review procedures to minimize the impact of UR; for example, physicians may have learned to code diagnostic information to reduce a patient's likelihood of denial of admission or continued-stay authorization.

CONCLUSION

The results of this analysis suggest that compulsory hospital inpatient UR can reduce hospital use and expenditures. It was found that UR reduced admissions by approximately 12 percent, hospital routine expenditures by 14 percent, hospital ancillary expenditures by 10 percent, and total medical expenditures by 6 percent, even after controlling for a large number of external factors. UR appears to reduce expenditures mainly by reducing admissions through preadmission authorization. These findings are consistent with findings reported earlier (Feldstein, Wickizer, and Wheeler 1988; Wickizer, Wheeler, and Feldstein 1989), and provide further evidence of a UR effect. Because the estimation procedure used here relies on longitudinal variation to estimate UR effects (in effect using groups as their own controls), it is less susceptible to problems of bias that might arise from unmeasured factors. The consistency of these findings with those of earlier analyses is encouraging.

One important aspect of this study was analysis of the effects of

UR over time. It was possible to undertake this analysis because longitudinal data were available. The effects of UR were found to be greatest during the initial quarters of review activity. The observed pattern of effects could reflect the decision of some patients to delay seeking care, a change in physician behavior, or both of these, or some other factor(s).

These findings underscore the need to use longitudinal data with sufficient time-series observations to generate reliable estimates of program effect. Analysis based on less than two years of follow-up data is unlikely to provide reliable estimates of long-term program effect. This could be the reason why some earlier studies (Brian 1972, 1973; Flashner et al. 1973; Ness 1985) that used a one-year follow-up period to analyze UR found large program effects.

The covariance model used here does not eliminate the possibility of selection bias. As part of earlier analyses, however, Wickizer, Wheeler, and Feldstein (1989) tested for selection using the Heckman (1979) procedure and found no evidence of selection bias. Furthermore, if anything, selection is likely to lead to an underestimate of UR effect, since groups adopting UR would be more likely to have higher, not lower, expenditure growth rates.

The approach used for this analysis provides a computationally efficient method of estimating covariance models, and is especially useful when the number of cross-section units is large. The analysis showed UR to have important effects, even after controlling for group-specific differences. These findings lend further support to the accumulating evidence that UR can reduce hospital use and help control expenditures. However, the data analyzed for this study are based on just one insurance carrier's UR program. Therefore, caution should be used in generalizing the findings. Other UR programs using different review criteria and incentives may have different outcomes.

This analysis did not address important questions concerning the effects of UR on the quality of care or on patient health status. These questions should be addressed by future research.

ACKNOWLEDGMENTS

The author is grateful to CNA Insurance Companies of Chicago for providing claims data used for this study. The author wishes to thank Catherine McLaughlin, Jack Wheeler, Gary Solon, and Paul Feldstein for their assistance. Appreciation is also expressed to an anonymous reviewer for helpful comments, and to Laura Larsson for valuable bibliographic support.

NOTES

1. The dependent variables are measured in logarithms, so the estimated coefficients shown in Tables 2 and 3 represent approximate percentage changes. For example, the estimated UR coefficient for the admissions equation (-0.122) represents a change in admissions of 11.5 percent. The percentage change figures presented in the text have been calculated based on the coefficient estimates shown in Tables 2 and 3. The coefficients of the quarter dummy terms are omitted for economy of presentation.

REFERENCES

Brian, E. W. "Foundation for Medical Care Control of Hospital Utilization: CHAP—a PSRO Prototype." *New England Journal of Medicine* 288 (26 April 1973): 878-82.

———. "Government Control of Hospital Utilization: A California Experience." *New England Journal of Medicine* 286 (22 June 1972): 1340-44.

Brown, C. "Equalizing Differences in the Labor Market." *Quarterly Journal of Economics* 90 (February 1980): 113-34.

Chiswick, B. R. "Hospital Utilization: An Analysis of SMSA Differences in Occupancy Rates, Admission Rates, and Bed Rates." *Explorations in Economic Research* 3, no. 3 (Summer 1976): 326-78.

Congressional Budget Office. *The Effects of PSROs on Health Care Costs: Current Findings and Future Evaluations.* Washington, DC: Government Printing Office, 1979.

———. *The Impact of PSROs on Healthcare Costs: Update of CBO's 1979 Evaluation.* Washington, DC: Government Printing Office, 1981.

DiBlase, D. "Effective Review Can Trim Health Expenses." *Business Insurance* 20, no. 41 (1986): 35-38.

Dionne, G. "Moral Hazard and Search Activity." *The Journal of Risk and Insurance* 48, no. 3 (September 1981): 422-34.

Dobson, A., J. G. Greer, R. H. Carlson, F. A. Davis, L. E. Kucken, B. J. Steinhardt, T. P. Ferry, and G. S. Adler. "PSROs: Their Current Status and Their Impact to Date." *Inquiry* 15, no. 2 (June 1978): 113-28.

Donahue, R. "Unnecessary Health Care Wastes $100 Billion." *Medical Benefits* 4, no. 6 (1987): 38-42.

Dowd, B. E. "The Logic of Moral Hazard: A Game Theoretic Illustration." *The Journal of Risk and Insurance* 49, no. 3 (September 1982): 443-47 .

Ermann, D. "Hospital Utilization Review: Past Experience, Future Directions." *Journal of Health Politics, Policy and Law* 13, no. 4 (1988): 683-96.

Feldstein, P. J., T. M. Wickizer, and J. R. Wheeler. "Private Cost Containment: The Effects of Utilization Review Programs on Health Care Use and Expenditure." *New England Journal of Medicine* 318 (19 May 1988): 1310–15.

Flashner, B. A., S. Reed, R. W. Coburn, and P. R. Fine. "Professional Standards Review Organization." *Journal of the American Medical Association* 223, no. 13 (1973): 1473–84.

Fox, P. D., W. B. Goldbeck, and J. J. Spies. *Health Care Cost Management: Private Sector Initiatives.* Ann Arbor, MI: Health Administration Press, 1984.

Gabel, J., C. Jajich-Toth, G. deLissovoy, T. Rice, and H. Cohen. "The Changing World of Group Health Insurance." *Health Affairs* 7, no. 3 (1988): 48–65.

Gabel, J., S. DiCarlo, S. Fink, and G. deLissovoy. "Employer-sponsored Health Insurance in America." *Health Affairs* 8, no. 3 (1989): 116–28.

Gertman, P. M., and J. D. Restuccia. "The Appropriateness Evaluation Protocol: A Technique for Assessing Unnecessary Days of Hospital Care." *Medical Care* 19, no. 8 (1981): 855–71.

Goldfeld, S. M., and R. E. Quandt. "Some Tests for Homoscedasticity." *Journal of the American Statistical Association* 60 (June 1965): 539–45.

Health Care Financing Administration. *Professional Standards Review Organization 1978 Program Evaluation.* Washington DC: Department of Health, Education, and Welfare, 1979.

———. *Professional Standards Review Organization 1979 Program Evaluation.* Washington DC: Department of Health, Education, and Welfare, 1980.

Heckman, J. "Sample Selection Bias as Specification Error." *Econometrica* 47, no. 1 (1979): 153–62.

Herzlinger, R., and J. Schwartz. "How Companies Tackle Health Care Costs: Part I." *Harvard Business Review* 63, no. 4 (1985): 68–80.

Kemper, K. J. "Medically Inappropriate Hospital Use in a Pediatric Population." *New England Journal of Medicine* 318, no. 16 (1988): 1033–37.

Kmenta, J. *Elements of Econometrics.* New York: Macmillan, 1986.

MaCurdy, T. E. "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis." *Journal of Econometrics* 18 (January 1982): 89–94.

Manning, W. G., J. P. Newhouse, N. Duan, E. B. Keeler, A. Leibowitz, and M. S. Marquis. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *American Economic Review* 77, no. 3 (June 1987): 251–77.

McLaughlin, C. G., J. C. Merrill, and A. J. Freed. "The Impact of HMO Growth on Hospital Costs and Utilization." *Advances in Health Economics and Health Services Research* 5 (1984): 57–76.

Ness, C. O. "Utilization Review, in Conjunction with Redesign, Saves Employer 27% Over Projected 1984 Costs." *Employee Benefit Plan Review* 40, no. 3 (1985): 104–108.

Newhouse, J. P., and C. E. Phelps. *New Estimates of Price and Income Elasticities of Medical Care Services.* Santa Monica, CA: The RAND Corporation, 1976.

Nickell, S. "The Determinants of Occupational Success in Britain." *Review of Economic Studies* 49 (January 1982): 43–52.

O'Donnell, P. S. "Managing Health Costs under a Fee-for-Service Plan." *Business and Health* 4 (March 1987): 38–40.
Pauly, M. V. "The Economics of Moral Hazard: Comment." *The American Economic Review* 58, no. 3 (June 1968): 530–38.
Payne, S. M. "Identifying and Managing Inappropriate Hospital Utilization: A Policy Synthesis." *Health Services Research* 22, no. 5 (1987): 709–69.
Pindyck, R. S., and D. Rubinfeld. *Econometric Models and Econometric Forecasts.* New York: McGraw-Hill, 1981.
Restuccia, J. D., and P. Gertman. "A Comparative Analysis of Appropriateness of Hospital Care." *Health Affairs* 3, no. 2 (1984): 130–38.
Restuccia, J. D., B. E. Kreger, S. M. Payne, P. M. Gertman, S. J. Dayno, and G. M. Lenhart. "Factors Affecting Appropriateness of Hospital Care in Massachusetts." *Health Care Financing Review* 8, no. 1 (1986): 47–54.
Richards, G. "Business Spurs UR Growth." *Hospitals* 58, no. 5 (1984): 96–100.
Sayetta, R. B. "Critique of an Earlier Study of the Sacremento Medical Care Foundation's Certified Hospital Admission Program (CHAP)." *Medical Care* 14, no. 1 (1976): 80–90 .
Shahoda, T. "Preadmission Review Cuts Hospital Use." *Hospitals* 58, no. 15 (1984): 54–56.
Scheffler, R. M. "The United Mine Workers' Health Plan: An Analysis of the Cost-Sharing Program." *Medical Care* 22, no. 3 (1984): 247–60.
Siu, A. L., E. A. Sonnenberg, W. G. Manning, G. A. Goldberg, E. S. Bloomfield, J. P. Newhouse, and R. H. Brook. "Inappropriate Use of Hospitals in a Randomized Trial of Health Insurance Plans." *New England Journal of Medicine* 315, no. 20 (1986): 1259–66.
Solon, G. "Estimating Autocorrelations in Fixed-Effects Models." Technical Working Paper no. 32. Cambridge, MA: National Bureau of Economic Research, 1984.
Tell, J. T., M. Falik, and P. D. Fox. "Private-Sector Health Care Initiatives: A Comparative Perspective from Four Communities." *Milbank Memorial Fund Quarterly* 62, no. 3 (1984): 357–77.
Westphal, M., E. Frazier, and C. Miller. "Changes in Average Length of Stay and Average Charges Generated following Institution of PSRO Review." *Health Services Research* 14, no. 4 (1979): 253–65.
Wickizer, T. M. "The Effect of Hospital Utilization Review in Selected Diagnostic Areas: An Exploratory Study." *American Journal of Public Health* 4 (April 1991): 482–84.
———. "The Effects of Utilization Review on Hospital Use and Expenditures: A Review of the Literature and an Update on Recent Findings." *Medical Care Review* 47, no. 3 (1990): 327–63.
———. *The Impact of Private Utilization Review on Hospital Use and Medical Care Expenditures among Insured Groups.* Ph.D diss. The University of Michigan, 1989.
Wickizer, T. M., J. R. Wheeler, and P. J. Feldstein. "Does Utilization Review Reduce Unnecessary Care and Contain Costs?" *Medical Care* 27, no. 6 (1989): 632–46.
Wilson, P., and P. Tedeschi. "Community Correlates of Hospital Use." *Health Services Research* 19, no. 3 (August 1984): 333–55.