

# Alternatives for using multivariate regression to adjust prospective payment rates

by Steven H. Sheingold

*Multivariate regression analysis has been used in structuring three of the adjustments to Medicare's prospective payment rates. Because the indirect-teaching adjustment, the disproportionate-share adjustment, and the adjustment for large cities are responsible for distributing approximately \$3 billion in payments each*

*year, the specification of regression models for these adjustments is of critical importance. In this article, the application of regression for adjusting Medicare's prospective rates is discussed, and the implications that differing specifications could have for these adjustments are demonstrated.*

## Introduction

The enactment and implementation of Medicare's prospective payment system in 1983 represented a fundamental and dramatic change in the way the Federal Government reimburses hospitals for inpatient services provided to beneficiaries. It also placed the Government in the position of administering prices—annually facing the task of both updating the average price level and adjusting the relative prices paid to different hospitals. To aid in this effort, analysts have attempted to develop various methodologies with which to process available data and provide usable information to policymakers.

Many issues have been examined using multivariate regression analysis, particularly the adjustment of relative rates. Through the prospective payment system (PPS), the Health Care Financing Administration (HCFA) pays national average prices for each of 473 diagnosis-related groups (DRGs), adjusted for a limited number of factors that are thought to reflect unavoidable differences in costs among hospitals. In many ways, regression provides a flexible and relatively easily understood tool to analyze the effects of these factors. Indeed, this technique has already been used considerably in setting the system's current adjustments for indirect teaching costs and the costs of serving a disproportionate share of low-income patients. Regression estimates have also been considered to adjust for number of beds and city size.

There are a number of important and perhaps overlooked issues to consider for the application of regression analysis to setting or adjusting prospective payment rates, however. The principles typically used to guide the choice of regression models for research and hypothesis testing may not be fully appropriate for specifying regressions for payment purposes. Policy goals and payment principles may require that regression estimates differ from the "unbiased estimate" sought in more academic research. Moreover, the same criteria may impose more restrictions on the estimating model than are typically applied. As a result, potential regression specifications for payment purposes and the resulting estimated coefficients may vary considerably. Because PPS payments amount to approximately \$40 billion annually, such differences in estimated coefficients can have significant implications. For example, a 0.1 difference in estimates of the indirect-teaching effect

could imply an annual shift of payments amounting to approximately \$250 million.

## Regression issues for payment

PPS has only three national average prices per DRG—one for hospitals in urban areas with more than 1 million population, one for other urban hospitals, and one for rural hospitals. However, payments per discharge differ among hospitals by local wage levels, teaching activity, and hospital share of low-income patients. The rationale for adjusting rates for these factors is that they lead to cost differences that are beyond hospitals' control and should therefore be recognized in payment rates. Indeed, regression analysis has been used extensively in setting the last two adjustments, in part because it provides information about the contributions of these factors to cost variation among hospitals. Moreover, regression can be used to provide such information in a way that is consistent with policy goals and prospective pricing principles. A complete list of variable names and definitions, as used in the regressions discussed in this article, appears in Table 1.

## The indirect-teaching adjustment

Since the inception of PPS, hospitals with approved teaching programs have received additional payments to compensate the hospitals for the indirect costs of medical education programs. These payments vary directly with the IRB for each hospital—a proxy measure for the scope of teaching activity. These additional payments amounted to somewhat less than \$2.1 billion in fiscal year 1989.

For the initial PPS proposals, HCFA used regression analysis to estimate the magnitude of indirect costs to be an increase in the average cost of a Medicare discharge of 5.795 percent for each 10-percent increase in the IRB. Congress legislated an initial indirect-teaching adjustment for PPS of 11.59 percent for each 0.1 percent increase in the IRB—double the estimate for indirect teaching costs. The rationale for this step was that, in addition to compensating these hospitals for indirect teaching costs, this adjustment would also serve as a partial correction for the system's inability to account for other factors that legitimately increase costs in teaching hospitals. In particular, there was concern about the ability of DRGs to account for severity of illness of patients requiring the specialized services often provided in teaching hospitals.

In response, many analysts began to examine methods

Reprint requests: Steven H. Sheingold, Ph.D., 370 L'Enfant Promenade, SW., Suite 900, Washington, D.C. 20024.

**Table 1**  
**Description of regression variables**

Variable	Description	Variable	Description
<b>Description and payment</b>		<b>Bed deciles—Urban hospitals—Cont.</b>	
IRB	1 plus the ratio of residents to beds	BED8	Number of beds greater than or equal to 310 and less than 387
CMI	1984 Medicare case-mix index (1981 case-mix index if 1981 cost data used)	BED9	Number of beds greater than or equal to 387 and less than 506
WI	Medicare wage index	BED10	Number of beds greater than or equal to 506
URBAN	1 if located in a metropolitan statistical area (MSA)	<b>Bed deciles—Rural hospitals</b>	
MPOP2	1 if located in an MSA with between 250,000 and 1 million in population	RBED2	Number of beds greater than or equal to 24 and less than 30
MPOP3	1 if located in an MSA with more than 1 million in population	RBED3	Number of beds greater than or equal to 30 and less than 37
CC	Central city within MPOP3	RBED4	Number of beds greater than or equal to 37 and less than 45
URB	Located in a rural county that contains a city or town of at least 20,000 in population	RBED5	Number of beds greater than or equal to 45 and less than 53
LOWIN	The sum of the percent of Medicare days attributable to Supplemental Security Income recipients and the percent of total days for which Medicaid is principal payer	RBED6	Number of beds greater than or equal to 53 and less than 66
URB>99	1 if urban hospital with more than 99 beds	RBED7	Number of beds greater than or equal to 66 and less than 85
DSHVAR5	Five variables restricted to produce the exact payment adjustment for each hospital specified under the current disproportionate-share adjustment	RBED8	Number of beds greater than or equal to 85 and less than 110
BEDS	Number of beds	RBED9	Number of beds greater than or equal to 110 and less than 159
<b>Bed deciles—Urban hospitals</b>		RBED10	Number of beds greater than or equal to 159
BED2	Number of beds greater than or equal to 54 and less than 90	<b>Dependent</b>	
BED3	Number of beds greater than or equal to 90 and less than 121	MOCC	Medicare cost per discharge, net of capital and direct medical education expenditures, hospital years beginning in fiscal year 1984
BED4	Number of beds greater than or equal to 121 and less than 159	STDCST	MOCC standardized by CMI and WI
BED5	Number of beds greater than or equal to 159 and less than 204	STDCSTO	STDCST net of estimated outlier payments per discharge
BED6	Number of beds greater than or equal to 204 and less than 251	FSTD	MOCC standardized for CMI, WI, indirect-teaching and disproportionate-share payments
BED7	Number of beds greater than or equal to 251 and less than 310	FSTD0	FSTD net of estimated outlier payments per discharge

NOTE: All continuous variables except LOWIN are measured in natural logarithms.

SOURCE: Sheingold, S.: Battelle Institute, Washington, D.C., 1990.

to determine the size of an indirect-teaching adjustment, looking at ways to compensate for teaching costs only or for other factors as well. One proposal was that the adjustment be structured to reflect indirect teaching costs and all other factors that affect costs but are not explicitly recognized in calculating PPS rates (Lave, 1984, 1985). For example, the adjustment could be used to provide partial compensation for severity-of-illness differences within DRGs, costs associated with being located in more populous metropolitan areas and central cities, costs associated with being larger (as measured by the number of beds) or with having more specialized facilities—all of which are correlated with teaching hospitals.

An interesting aspect of this proposal was that the size of this composite adjustment would be determined by multivariate regression analysis. In particular, an estimate of the indirect-teaching effect would be obtained from a payment model—that is, a regression including only variables used to calculate payment rates—rather than from the more fully specified regression model typically employed for such analyses. In this way, the influence of these excluded factors would be “loaded on” to the estimated indirect-teaching effect, to the extent that they were correlated with the size of teaching programs. Moreover, any errors in measuring the payment variables

themselves, such as case-mix compression, were also more likely to affect the teaching coefficient in this specification. Indeed, a similar methodology was used as a basis for reducing the adjustment to its current level of 7.7 percent. (Case-mix compression—the relative underpricing of DRGs with higher weights—was thought to adversely affect teaching hospitals because, on average, they tend to have larger case-mix indexes.)

The debate over the size of this adjustment continues as more recent cost data become available, the Federal budget deficit remains larger than desired, and equity questions arise concerning the current structure of the adjustment. It is likely that regression analysis will continue to play a role in the debate. As is demonstrated in this article, however, there are considerations other than the exclusion of nonpayment variables in specifying a payment regression for estimating this adjustment. These other considerations can have substantial implications for the estimated size of the adjustment.

### Disproportionate-share adjustment

In March 1986, Congress legislated an adjustment to PPS rates intended to compensate hospitals for the costs

of serving a disproportionately large share of low-income patients. It was hypothesized that such hospitals would experience higher costs for two reasons. First, it was thought that low-income patients were more severely ill within any given DRG than their higher income counterparts. Second, such hospitals may respond to the needs of low-income patients by adding specialized and costly staffing and facilities.

One of the most important questions relevant to the policy debate was whether hospitals' share of low-income patients significantly affected the average cost per Medicare discharge. Another of the important considerations was how "disproportionate share" should be defined. Estimates from multivariate regression analyses demonstrated that, after a threshold level, hospitals' low-income-patient share had a significant and positive impact on Medicare costs in urban hospitals with more than 99 beds. The estimated coefficients were used in a formula to adjust PPS rates and to define disproportionate-share hospitals.

Again, a number of questions arose as to how regression should be applied to provide estimates that were both statistically reliable and consistent with payment goals. For example, should the specification include only PPS payment variables, or should it include a number of nonpayment variables to provide a better estimate of the partial impact of low-income patients on costs? Should there be separate specifications for urban and rural hospitals? The impact of these issues on the estimates is examined in the following sections.

### Other adjustments

Both hospital size and the size of the metropolitan statistical area (MSA) in which a hospital is located have been discussed as possible adjustment factors to PPS rates. These are two of the more prominent factors whose impacts on costs are partially compensated through the indirect-teaching adjustment. Various regression specifications can provide useful information on the impact of these variables and on how much of these impacts may already be compensated for to some extent by other features of PPS.

### Applying regression to adjusting rates

The standard regression model can be represented as follows:

$$Y = XB + u \quad (1)$$

where  $Y$  is the dependent variable,  $X$  a vector of explanatory variables hypothesized to influence variations in  $Y$ , and  $u$  the disturbance or error term. The primary condition to obtain unbiased estimates of  $B$  is that  $E(Xu) = 0$ . That is, the disturbance term only reflects random influences on  $Y$ , or those with no systematic relationship to  $X$ . If the disturbance term reflects factors that are either unmeasured, incorrectly measured, or omitted from  $X$ , then  $B$  will be biased in proportion to the correlation between these factors and  $X$ . These problems are often referred to collectively as specification error.

The choice of functional form, variables, and any restrictions placed on the estimation is, in principle,

guided by a theory or model that leads to testable hypotheses about the relationships between variables. Much of the work in describing hospital costs has taken the approach of applying regression to a relationship known as a hospital cost function, which relates a measure of costs to a variety of factors. A commonly used variety of this relationship is known as a behavioral cost function and in general form is written:

$$C = f(I, O, S, T) \quad (2)$$

where  $I$  represents input prices,  $O$  represents variables describing output heterogeneity,  $S$  measures scope of services or inputs fixed in the short run, and  $T$  is a measure of teaching activity. Specifications of the behavioral cost function have sometimes been labeled "ad hoc," in that many of the variables used—particularly in  $O$ —are not strictly related to standard microeconomic theory. In contrast, a technical cost function is more stringently based on traditional theory of production—cost minimization, for example—and would only contain a contemporaneous output measure and input prices (Breyer, 1987). This last specification often leaves less flexibility in choosing a functional form for estimation.

The behavioral cost function recognizes that nonprofit hospitals may pursue objectives other than cost minimization, and hence, a number of other factors may affect costs (Evans, 1971; Lave and Lave, 1984). Such a cost function is generally implemented for the purposes of estimating the determinants of Medicare cost as:

$$MOCC = f(CMI, WI, LOWIN, LOC, BEDS, IRB) \quad (3)$$

Particularly when estimates are to be used for examining PPS, the dependent variable (MOCC) is the hospital's average cost for a Medicare discharge, net of capital and direct medical education costs. CMI represents Medicare's DRG case mix, WI the Medicare wage index, and LOWIN is a measure of the share of each hospital's patients that are low income. The number of hospital beds (BEDS) is often used as a proxy for scope of facilities and services, and LOC represents the type of area in which the hospital is located, such as size of MSA. These last two variables may reflect a number of factors that cause costs to vary, such as nonlabor input prices, tastes and preferences of consumers, and practice styles of physicians. Although teaching status may be measured in various ways, the IRB has been the measure of teaching intensity used by Medicare and is the basis for the current indirect-teaching adjustment. (Because regressions are often estimated in log-linear form and about 80 percent of hospitals would have an IRB of 0, the teaching variable is implemented as  $1 + IRB$ .) This basic formulation was employed by HCFA to determine the estimate of 5.795 for the indirect teaching costs. Other efforts have also included regional and ownership variables to further specify hospitals' objective functions.

### Specifying alternative payment models

A PPS payment regression would be a modification of this approach, specified to be consistent with PPS payment parameters, principles, and policy objectives. The specifications used for estimating the indirect-

teaching adjustment made a first step in that direction by omitting from equation (3) any variables not used in determining PPS prices, such as city size and number of beds (Sheingold, 1985; Anderson and Lave, 1986).

On further examination, however, there are a number of important issues that might be considered in specifying payment models:

- Should such specifications be restricted so as to be internally consistent with actual PPS payment parameters?
- Should the specifications account in some way for payments made through the system's provisions for outlier cases?
- Should payment models, in some cases, include variables not used in setting prices?
- Should separate regressions be used for urban and rural hospitals in estimating various rate adjustments?
- Should regressions be modified to account for a recent change in the way the standardized amounts are calculated (as discharges) rather than as hospital-weighted averages?

The first issue in the preceding list concerns whether estimated parameters of policy interest, such as the indirect-teaching adjustment, should be estimated using an unrestricted or restricted regression model. In the latter case, PPS parameters such as CMI or WI would be restricted to their actual system values. Although a 1-percent difference in CMI and WI would result in 1-percent and 0.75-percent differences in payments, respectively, estimated coefficients from an unrestricted model might deviate substantially from these values for a number of reasons. The consequences for the estimated policy parameter, and the distribution or level of payments should this parameter be implemented, could be considerable.

Although there are obvious statistical reasons for this potential difference, the issue of what type of regression model to use can also be considered within the context of the system's pricing mechanism. The basic PPS prices, or standardized amounts, are an average of each hospital's cost per Medicare case, adjusted or standardized for payment differences resulting from such factors as case-mix and wage index. These payment differences are set at the rate of 1 percent and 0.75 percent, respectively, for these factors. If the estimated coefficients for CMI and WI differ from these values, the estimated coefficients for the other variables will reflect their impact on a "standardized" cost that differs from the one that is used to calculate PPS rates.

Thus, a regression specified to be fully consistent with PPS pricing might restrict relevant parameters to their system values. One method of implementing this approach would be to estimate equation (3) with the appropriate restrictions placed on the coefficients of relevant variables. An equivalent approach, however, is to standardize the cost per case by these variables in the same manner used to calculate PPS prices, rather than including them on the right side of the equation in restricted form. (That is, each hospital's cost per case would be standardized by case mix and wage index as they are for calculating the system's standardized rates. In a log-linear regression, this method would be equivalent to restricting these coefficients to 1.0 and 0.75, respectively.) Such an approach would provide the same estimates for the nonrestricted parameters and has

additional appeal in keeping the dependent variable consistent with the basic system prices—the standardized amounts.

Another issue is how to include the effect that outlier payments have on the parameters of interest. Although other payment variables, such as teaching and case mix, are usually included in estimated cost functions—whether for payment purposes or not—outlier payments are a feature unique to PPS and are not usually included in cost analyses. Nonetheless, outlier payments are important to a payment specification because they tend to be correlated with variables of interest—such as teaching status and disproportionate share—and hence, will affect the estimated coefficients of these variables. Put another way, outlier payments currently serve a function somewhat similar to the indirect-teaching and disproportionate-share adjustments, accounting for severity of illness. Therefore, to exclude these payments could result in estimates that would effectively represent double payment.

Still another concern is whether to include only PPS payment variables in the regression. In some forms of the payment specification, it may be consistent with a policy goal to exclude nonpayment variables. To the extent that a policy variable is correlated with the excluded variables, its estimated coefficient will reflect their influence on costs through one type of specification error called "omitted variables bias." In other words, the ordinary least-squares requirement that  $E(Xu) = 0$  would be violated and therefore the estimated coefficients biased. For example, consider the case in which the true regression is:

$$Y = bX + cZ + u \quad (4)$$

but the following is estimated instead:

$$Y = bX + v \quad (5)$$

Rather than obtaining an unbiased estimate of  $b$ , the estimated coefficient for  $X$  would be  $b + cd$ , where  $d$  would be the coefficient of  $Z$ , if it were regressed on  $X$ . The bias therefore is the product of the impact the excluded variable would have on  $Y$ , multiplied by the regression coefficient of the excluded variable on the included variable.

This policy choice to include only the PPS payment variables was made for past estimates of the indirect teaching effect. Because teaching hospitals tend to be larger than other hospitals and are more likely to be located in large cities, the effects of these variables were loaded on to the estimated indirect-teaching effect in proportion to their correlation with the ratio of residents to beds.

In other instances, there may be interest in the unbiased effect of a particular variable on costs, although in a way that is consistent with the payment parameters. For example, the effects of low-income-patient share on costs might be estimated in a specification including number of beds and location variables, at the same time accounting for the system values of CMI, WI, and outlier payments.

An issue of recent importance is whether the variables should be weighted by each hospital's Medicare discharges, regardless of which specification is used. Through fiscal year 1987, the system's standardized

amounts have been based on a simple average of hospital costs. Regressions using the hospital as the unit of observation then provided estimates that could be used to measure variation around this hospital-weighted average. Beginning in fiscal year 1988, the standardized amounts were calculated as discharge-weighted averages of hospital costs. It can be argued, therefore, that weighted regression should be used to estimate adjustments to these new rates.

Finally, whichever decisions are reached with regard to the preceding issues, a question remains as to whether regressions should be estimated separately for urban and rural hospitals. On one hand, some argue for separate regressions to be consistent with calculating separate standardized amounts for urban and rural areas. Others argue, however, that, although the DRG weights are applied to separate urban and rural rates to determine payments, they are based on national averages.

The impact of these issues is examined in the following sections. The alternative regression specifications are applied to estimating the indirect-teaching effect, the impact of low-income patients on costs, and the impact of city size and number of beds.

## Data and variables

Most data used in the following sections were provided by HCFA. The basic cost data and other hospital characteristics were taken from Medicare cost reports for 1981 and for the first year of PPS (PPS1). The DRG case-mix index, the wage index, and disproportionate-share data for each hospital were provided separately. The case-mix indexes are based on Medicare inpatient bill files: Medicare Provider Analysis and Review (MEDPAR) file for 1981 and the patient billing (PATBILL) file for 1984. The wage indexes are based on a 1982 HCFA wage survey of hospitals. The sizes of MSAs are calculated from published population estimates for 1985, and central city location data were provided by the U.S. Bureau of the Census.

Five dependent variables are used for the analyses. All are based on hospital average costs for a Medicare discharge, net of capital and direct medical education expenditures. Which of these is used for a particular specification determines how many PPS payment variables are restricted, so that their estimated coefficients reflect their system payment values. The coefficients of all variables are unrestricted when MOCC is the dependent variable. STDCST is cost per discharge standardized for the case-mix and wage indexes, and STDCSTO is calculated net of estimated outlier payments per discharge as well. Using STDCST is equivalent to restricting the case-mix and wage index coefficients to 1 and 0.75, respectively, and using STDCSTO also accounts for any differences in costs compensated for by outlier payments. (In regressions using STDCST or STDCSTO, CMI and WI are not included as independent variables. This method is equivalent to using MOCC as the dependent variable [or MOCC net of outliers], including those variables on the right side and restricting their coefficients to 1 and 0.75.) Similarly, FSTD and FSTDO are fully standardized by the case-mix, wage index, indirect-teaching, and disproportionate-share payments, and hence, the coefficients of these variables

are restricted to their payment values. For the same reasons, when FSTD or FSTDO are used as dependent variables, IRB and the disproportionate-share variables are not included as explanatory variables.

In the following sections, results from both unrestricted regressions and restricted regressions are displayed. Both provide valuable information for examining PPS. Unrestricted regressions provide general information on hospital cost variation and specific information with which to contrast estimated effects of payment parameters (CMI, for example) with their assumed system values. In this way, potential problems with the system's current adjustments can be examined. On the other hand, the restricted regressions allow for particular current PPS rate adjustments or proposed adjustments to be evaluated relative to current prices. In other words, this method provides the impact of these variables after accounting for those cost variations already compensated for by the system.

## Estimates of indirect-teaching effect

In Table 2, estimated payment regressions based on the 1981 Medicare cost data that were originally used to estimate the indirect-teaching adjustment are presented. In the first three columns, the regressions include only the system's initial payment variables, CMI, WI, URBAN, and IRB. In the second through fourth columns, CMI and WI are used to standardize cost per case, rather than appearing as explanatory variables with restricted coefficient estimates. The coefficient for IRB in these equations will reflect not only indirect teaching costs, but the influence of bed size, city size, and low-income-patient share. To an extent, these three factors are correlated with teaching activity. In the third equation, outlier payments are accounted for by specifying the dependent variable to be standardized costs net of an estimate of each hospital's outlier payments per case. The fourth regression accounts for any overlap with the current disproportionate-share adjustment by adding five low-income-patient share variables and restricting their coefficients to produce the exact payment value of the adjustment to each hospital. For example, the impact for an urban hospital with more than 99 beds and 30-percent low-income share would involve 2 coefficients restricted so that the impact on costs would be:

$$0.025 + (0.3 - 0.15)0.5$$

The more typically employed specification, using MOCC, suggests an indirect-teaching adjustment of 8.3 percent for each 10-percent increase in IRB. The coefficients of 1.52 and 1.18 for CMI and WI, which are considerably larger than their payment values, will lead to a smaller teaching coefficient, to the extent that teaching hospitals tend to have larger case-mix and wage indexes relative to other hospitals. (If payments were determined at the rates suggested by these coefficients, the standardized costs of teaching hospitals would be lower relative to other hospitals.) Similarly, the estimated urban-rural differential of 10 percent is considerably smaller than the actual difference in the urban and rural standardized amounts (25 percent) that were in effect prior to 1987. When the dependent variable is

**Table 2**  
**Regression results using various measures of cost per case as a dependent variable**

Item	Cost per case	Standardized cost per case	Standardized cost per case net of outliers	Standardized cost per case net of outliers
<b>Independent variable</b>			Coefficient	
Constant	7.54	7.45	7.43	7.43
CMI	1.52	( <sup>1</sup> )	( <sup>1</sup> )	( <sup>1</sup> )
WI	1.18	( <sup>1</sup> )	( <sup>1</sup> )	( <sup>1</sup> )
URBAN	.099	.223	.176	.139
IRB	.829	1.09	.841	.646
DSHVAR5	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )
		Regression statistic		
F <sup>2</sup>	.646	.249	.167	.169

<sup>1</sup>Variable is used to standardize cost per case and therefore is not included in restricted form as an independent variable.

<sup>2</sup>Indicates variable is not included in the model.

<sup>3</sup>These coefficients are restricted to provide each hospital receiving the adjustment with their exact payment value.

NOTES: Variable names and definitions are shown in Table 1. All nonrestricted variables are statistically significant at the 1-percent level for a two-tailed test.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy; Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

standardized by CMI and WI, effectively restricting their coefficients to system values, the estimated teaching coefficient rises to 1.1—very close to the doubled teaching adjustment previously described—and the urban-rural differential is estimated to be very close to the actual difference of 25 percent. When the effect of outlier payments is included, the teaching coefficient falls to 8.4 and was the basis of the 8.7-percent adjustment initiated in 1986 for years in which there was no disproportionate-share adjustment. (This data set contained about 300 more hospitals than that originally used by HCFA to obtain the 5.795 estimate, resulting in slightly lower estimated coefficients. Therefore, a policy decision was made to raise this adjustment to 8.7, rather than 8.4 percent.) When the actual payment form of the disproportionate-share adjustment is included, the teaching effect drops to 6.5, because of the considerable overlap between the two effects. Of the 183 major teaching hospitals used in the estimation, 140 receive disproportionate-share payments. Together, all hospitals that receive indirect-teaching payments account for 65 percent of all disproportionate-share payments. A policy decision was made, however, to reduce the teaching adjustment to 8.1 percent to partially account for this overlap.

Which of these specifications is “correct” really depends both on what policy goals and payment principles are and why the coefficients for payment variables differ from their system values. For example, if the large coefficient for CMI results from pricing problems such as case-mix compression or severity of illness, then policymakers may wish to use a specification that restricts its coefficient to its payment value of 1.0—thereby causing the teaching effect to reflect the impact of these pricing problems. On the other hand, if the estimated CMI effect were thought to result from teaching hospitals systematically undercoding their diagnoses prior to PPS, then an unrestricted specification might be preferred, as this problem has been self-correcting.

Indeed, undercoding should be a minimal problem when cost and case-mix data from PPS years are used.

In Table 3, comparable specifications for estimating the indirect-teaching adjustment are presented, using data from hospitals’ first year under PPS. Separate estimates are presented with and without hospitals in New York, because of a special situation involving the system’s outlier payments. For a number of reasons, average length of stay in New York is considerably higher than in other States, yet New York’s average cost per case is relatively low. (In fiscal year 1984, average Medicare length of stay in New York hospitals was 13.6 days, compared with an average of 7.5 days for all other States. New York hospitals’ specific payment rates are, on average, 17 percent lower than national payment rates.)

As a result, New York’s hospitals, in comparison with those of other States, are estimated to receive outlier payments that are very large relative to their costs. For example, the ratio of estimated outlier payments to costs is 0.045 for all other States and 0.295 for New York. Hospitals in the District of Columbia have the next-largest ratio of 0.193. Because of this situation, the inclusion of New York hospitals can reduce the regression estimates for the teaching effect considerably. Both sets of results are presented here, because it is likely to be a policy decision as to whether the indirect-teaching adjustment for all hospitals will be reduced to reflect this situation in New York.

Again, there is considerable variation in the estimated teaching coefficient among the different specifications. As can be seen in the table, the estimated teaching effects with and without New York hospitals are similar when standardized cost (STDCST) is the dependent variable. When outliers are accounted for (STDCSTO is the dependent variable), the inclusion of New York hospitals reduces these estimates substantially, however. The estimated indirect-teaching effects without New York are 7.0 percent, if the disproportionate-share adjustment is not accounted for, and 4.9 percent, if the disproportionate-share variables are included. (The estimated effects were similar (6.9 and 4.8 percent), if, instead of excluding the New York hospitals, they were included and a dummy variable for these hospitals added to the regression.) With New York hospitals included, the corresponding estimates are 5.8 and 3.4 percent. In either case, the estimates are considerably lower than those obtained from the 1981 data and lower than the current law counterparts of 8.7 and 8.1 percent.

The estimates rise when the same regressions are based on discharge-weighted variables (Table 4). In these specifications, the indirect-teaching effect is estimated to be an increase in costs of 7.4 percent for each 10-percent increase in the IRB (5.5 percent if the disproportionate-share adjustment is included). The latter is in contrast to the 6.8-percent adjustment that would result under the policy that deducted 0.6 percentage points to account for disproportionate share. It is also of policy interest to use a more fully specified model to examine other factors that affect costs and to obtain estimates of the indirect-teaching effect itself—that is, to control for these other factors. Estimates of this partial effect are provided in Table 5. In columns 1 and 2, the indirect-teaching effect is estimated in payment-style specifications, but in columns 3 and 4, all coefficients are unrestricted. The

**Table 3**  
**Regression results using various measures of hospital payment as a dependent variable**

Independent variable	Dependent variable									
	Cost per case <sup>1</sup>		STDCST		STDCST without New York		STDCSTO		STDCSTO without New York	
	Without LOWIN	With LOWIN	Without DSHVARS	With DSHVARS	Without DSHVARS	With DSHVARS	Without DSHVARS	With DSHVARS	Without DSHVARS	With DSHVARS
	Coefficient									
Constant	7.62	7.61	7.58	7.56	7.62	7.60	7.55	7.53	7.50	7.49
IRB	0.886	0.719	1.020	0.785	0.996	0.766	0.578	0.343	0.701	0.487
CMI	1.291	1.239	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )
WI	0.979	0.949	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )	( <sup>2</sup> )
URBAN	0.105	0.060	0.169	0.137	0.175	0.143	0.119	0.086	0.139	0.107
LOWIN	( <sup>3</sup> )	0.149	( <sup>3</sup> )	( <sup>3</sup> )	( <sup>3</sup> )	( <sup>3</sup> )	( <sup>3</sup> )	( <sup>3</sup> )	( <sup>3</sup> )	( <sup>3</sup> )
	Regression statistic									
R <sup>2</sup>	0.646	.656	.222	.221	.232	.221	.091	.069	.144	.134
F-statistic	2,545	1,768	292	788	761	753	276	206	449	410
Number	5,562	5,562	5,562	5,562	5,324	5,324	5,562	5,562	5,324	5,324

<sup>1</sup>Unstandardized cost per Medicare discharge.

<sup>2</sup>Variable used to standardize Medicare cost per discharge.

<sup>3</sup>Variable not included.

NOTES: Variable names and definitions are shown in Table 1. All variables with unrestricted coefficients are statistically significant at the 1-percent level for a two-tailed test.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

estimated effect of IRB in the payment regressions might be interpreted as indirect teaching costs, net of any other payments in the system that might partially subsidize these costs, such as outlier payments or the disproportionate-share adjustment. Conversely, the IRB coefficients in columns 3 and 4 might be interpreted as estimates of the indirect teaching costs that reflect a less stringent accounting for its interaction with other aspects of the payment system.

Depending on the specification, the estimated indirect teaching costs can vary considerably. They range from a 1.7-to-3.1-percent increase in cost for each 10-percent increase in IRB in the payment regressions (for unweighted and weighted, respectively) to a 5.0-to-6.6-percent increase in the unrestricted specifications. It is interesting to note that the comparable unrestricted, unweighted specification using the 1981 data (not shown) produced an estimated indirect-teaching effect of 4.05 percent, which is the basis of the current teaching adjustment.

Finally, the use of a single regression for both urban and rural areas has implications for the 40 rural hospitals

that receive indirect-teaching payments. Although the estimated indirect-teaching effects do not differ substantially for urban hospitals if a separate regression is used, the rural indirect-teaching effect would be considerably larger than that found in the aggregate specification. Compared with the unweighted and weighted teaching effects of 7.0 and 7.4 percent for all hospitals (Table 4), the impact in rural areas would suggest teaching adjustments of 9.6 percent and 19.1 percent, respectively.

**Table 5**  
**Results of fully specified regressions using two alternative dependent variables**

Item	Dependent variable is STDCSTO		Dependent variable is MOCC	
	Unweighted	Weighted	Unweighted	Weighted
<b>Independent variable</b>	<b>Coefficient</b>			
Constant	7.37	7.42	7.36	7.42
IRB	0.17	0.31	0.50	0.66
CMI	( <sup>1</sup> )	( <sup>1</sup> )	0.91	0.83
WI	( <sup>1</sup> )	( <sup>1</sup> )	0.98	0.91
MPOPI	*0.02	*-0.01	*0.01	*-0.02
MPOP2	*0.02	*0.00	*0.00	*0.00
MPOP3	0.06	0.05	0.03	0.03
CC	0.07	0.05	0.08	0.05
BEDS	0.06	0.05	0.07	0.08
LOWIN	( <sup>1</sup> )	( <sup>1</sup> )	-0.09	-0.19
LOWIN X URB>99	( <sup>1</sup> )	( <sup>1</sup> )	0.27	0.41
DSHVARS	—	—	( <sup>1</sup> )	( <sup>1</sup> )
	Regression statistics			
R <sup>2</sup>	0.18	0.19	0.68	0.72
F-statistic	199.00	211.00	1,120	1,368

\*Not statistically significant at the 5-percent level for a two-tailed test.

<sup>1</sup>Variable not included.

NOTES: Variable names and definitions are shown in Table 1. Hospitals located in New York State are excluded.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

**Table 4**

**Weighted and unweighted teaching coefficients from payment regression excluding New York hospitals, using cost per discharge net of estimated outlier payments (STDCSTO) as a dependent variable**

Independent variable	Without DSHVARS		With DSHVARS	
	Unweighted	Weighted	Unweighted	Weighted
IRB	0.70	0.74	0.49	0.55
URBAN	0.14	0.11	0.11	0.08

NOTES: Variable names and definitions are shown in Table 1. Unweighted results are obtained from regressions presented in the last two columns of Table 3. In these regressions, each hospital's observation is weighted by the number of its Medicare discharges.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

**Table 6**  
**Estimated coefficients for a categorized teaching variable, from selected regression models of hospital payment**

Resident-to-bed ratio	Dependent variable is MOCC			Dependent variable is STDCSTO		
	Payment model <sup>1</sup>	Fully specified model <sup>2</sup>	Fully specified model <sup>3</sup>	Payment model <sup>1</sup>	Fully specified model <sup>2</sup>	Fully specified model <sup>3</sup>
	Coefficient					
0-0.1	0.06	0.03	0.03	0.05	0.02	0.02
0.1-0.2	0.10	0.05	0.06	0.07	0.02	0.02
0.2-0.3	0.19	0.15	0.14	0.14	0.08	0.07
0.3-0.4	0.24	0.19	0.14	0.20	0.15	0.10
0.4-0.5	0.29	0.31	0.22	0.24	0.25	0.17
0.5 or greater	0.32	0.31	0.22	0.24	0.25	0.17

<sup>1</sup>Regression includes URBAN, CMI, WI, and DSHVARS when MOCC is the dependent variable, URBAN and DSHVARS when STDCSTO is the dependent variable.

<sup>2</sup>Also includes MPOP1, MPOP2, MPOP3, CC, BEDS, LOWIN, and LOWIN X URB>99.

<sup>3</sup>Also includes a variable representing whether a hospital is university affiliated and a member of the Council of Teaching Hospitals. The coefficient of this variable is 0.11 when MOCC is the dependent variable and 0.09 when STDCSTO is the dependent variable.

NOTES: Variable names and definitions are shown in Table 1. All coefficients are statistically significant. In these regressions, each hospital's observation is weighted by the number of its Medicare discharges.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

### Are there threshold effects?

Another question that has arisen related to specification is whether the indirect teaching costs are best estimated as a log-linear relationship and indeed, whether an adjustment should be paid to all hospitals with teaching programs. Specifically, some have questioned whether there is some threshold size of a teaching program at which indirect costs are a significant factor or whether the relationship between size and costs is continuous. The implications are potentially important for PPS; an adjustment that began at higher levels of the IRB could reduce substantially the number of eligible hospitals and affect both the distribution and level of PPS payments. For example, one suggestion is to adjust rates only for hospitals whose IRBs are at least 0.15 (Welch, 1987).

This issue is examined by employing a series of dummy variables representing various ranges of the IRB. The impact of these variables is then estimated in the various regression specifications used previously. The results for selected specifications are presented in Table 6.

In general, when nonpayment variables are excluded from the equations (columns 1 and 4), there does not appear to be a threshold level of the IRB at which indirect-teaching effect becomes significant—even when outlier payments are accounted for. Rather, it appears that statistically significant cost effects exist throughout its range. Moreover, as demonstrated in Table 7, these estimated impacts are comparable to those obtained from a log-linear specification. For example, the coefficient for the 0.3-0.4 range of the IRB implies a cost impact of 22 percent, and the log-linear form evaluated at the midpoint of the range implies an impact of nearly 18 percent.

In addition, there does not appear to be a threshold for estimates of the indirect teaching costs themselves. In all of the fully specified regressions, the coefficients for all ranges of the IRB remain statistically significant, although smaller. (If unweighted regressions are used

instead, there is a threshold at about 0.2 of the IRB when STDCSTO is the dependent variable.) Adding a variable that controls for whether the hospital is university affiliated and a member of the Council of Teaching Hospitals (an alternative measure of teaching status) reduces the estimated impact for those hospitals with IRBs greater than 0.3, but has little effect on the estimates for smaller teaching programs.

### The impact of disproportionate share

The current disproportionate-share adjustment is also based on results from multivariate regression analyses. Unlike the indirect-teaching adjustment, however, the impact of low-income-patient share on Medicare costs was estimated for legislative purposes by using separate regressions for large urban hospitals (more than 99 beds), other urban hospitals, and rural hospitals. Although teaching programs are found mostly in larger urban

**Table 7**  
**Estimated effect of resident-to-bed ratio on indirect teaching costs under two alternative regression specifications**

Resident-to-bed ratio	Categorized IRB <sup>1</sup>	Log-linear <sup>2</sup>
		Percent
0-0.1	5.1	35.4
0.1-0.2	7.4	8.0
0.2-0.3	14.6	13.1
0.3-0.4	22.1	17.9
0.4-0.5	26.7	22.7
0.5 or greater <sup>4</sup>	27.4	29.5

<sup>1</sup>Calculated as e<sup>B</sup>.

<sup>2</sup>Calculated as IRB<sup>B</sup> where B = .55, from Table 4, column 4.

<sup>3</sup>Calculated using IRB = 1.1.

<sup>4</sup>Calculated using IRB = 1.6.

NOTE: Variable names and definitions are shown in Table 1.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.



hospitals, hospitals with large shares of low-income patients are spread across all three groups. In fact, the average share of low-income patients and the percent of hospitals with more than 40 percent low-income patients are higher for rural and small urban hospitals than for the larger urban group. (The average of LOWIN is 18.9, 19.2, and 14.5 percent, respectively, for rural, small urban, and larger urban hospitals. Moreover, the share of hospitals with more than 40 percent low-income patients is 10.2, 8.3, and 6.7, respectively.) For a number of reasons, however, it was hypothesized that the impact of these patients on costs would be greatest in the large urban hospitals (Sheingold, 1985). Therefore, separate regressions were used, and indeed, although statistically significant cost impacts were found for large urban hospitals, similar effects were not found for rural or small urban hospitals.

In Table 8, estimated regressions for urban hospitals with more than 99 beds are presented. The current formula for these hospitals provides an adjustment beginning at 2.5 percent for hospitals with 15 percent low-income share and increasing to a maximum adjustment of 15 percent for hospitals with 40 percent or greater low-income share. In general, all of the specifications tend to support this formula. The major exception is the group of hospitals with 45 percent or more low-income share for which the estimated unweighted effects are smaller than the next lower group. Moreover, the estimated weighted effect for this group is not significantly different from zero. Again, this would be a policy decision as to how these hospitals are paid. It is also apparent that the system's outlier payments subsidize some costs associated with low-income-patient share. The estimated effects uniformly decline when STDCSTO is the dependent variable relative to specification for which STDCST is used.

Another important question concerns how fully the regression for estimating low-income effects should be specified. In Table 9, the impact of including various nonpayment variables is explored for specifications using the same dependent variable—STDCSTO—and the various restrictions it imposes. Depending on the specification, the implied low-income effects on costs vary considerably. If city sizes and number of beds are included, there appears to be a threshold at which the cost impacts become significant at about a 15-percent low-income share (column 2, Table 9). If only payment variables are included, these effects begin at 30 percent low-income share instead, reflecting that a number of hospitals with large shares of low-income patients are located in the less costly smaller cities and have fewer beds. On the other hand, if regional and ownership variables are included with the city-size and bed-size variables, the threshold is at 10 percent, and the effects are generally of greater magnitude and have larger *t*-statistics (column 3, Table 9).

## Other potential rate adjustments

Both the size of the MSA in which a hospital is located and the number of beds have been discussed as possible adjusters to PPS rates. Although it is uncertain exactly what cost effects these variables represent, it is thought that MSA size may proxy for differences in nonlabor

**Table 8**

**Regression results for urban hospitals with more than 99 beds, specifying various levels of each hospital's share of low-income patients (LOWIN) as independent variables**

Item	Dependent variable is STDCST		Dependent variable is STDCSTO	
	Unweighted	Weighted	Unweighted	Weighted
<b>Independent variable</b>	<b>Coefficient</b>			
Constant	7.59	7.64	7.68	7.70
IRB	0.63	0.73	0.46	0.55
MPOP2	0.02	0.03	*0.01	*0.01
MPOP3	0.07	0.08	0.04	0.05
CC	0.07	0.05	0.08	0.07
LNBEDS	0.03	0.02	*0.01	*0.01
	<b>Coefficient and t-statistic</b>			
0.05-0.1	0.01 (0.77)	0.01 (0.73)	0.00 (0.32)	0.00 (0.60)
0.1-0.15	0.03 (2.37)	0.03 (2.43)	0.02 (1.56)	0.02 (1.61)
0.15-0.2	0.04 (3.57)	0.05 (3.61)	0.02 (1.47)	0.03 (2.03)
0.2-0.25	0.03 (1.47)	0.05 (2.73)	0.02 (0.27)	0.03 (1.26)
0.25-0.3	0.05 (2.07)	0.06 (3.01)	0.02 (1.02)	0.04 (1.65)
0.3-0.35	0.07 (2.43)	0.10 (3.41)	0.04 (1.42)	0.08 (2.53)
0.35-0.4	0.08 (2.66)	0.10 (2.63)	0.06 (1.69)	0.07 (1.85)
0.4-0.45	0.18 (6.34)	0.20 (5.76)	0.14 (4.54)	0.15 (4.21)
0.45 or greater	0.06 (2.87)	0.02 (0.66)	0.04 (1.60)	0.00 (0.03)
	<b>Regression statistics</b>			
Number	2,120	2,120	2,120	2,120
R <sup>2</sup>	.151	.162	.178	.191
F-statistic	19.2	20.1	19.8	22.8

\*Not statistically significant at the 1-percent, 5-percent, or 10-percent level for a two-tailed test.

\**t*-statistics are included in parentheses for the low-income-patient-share variables.

NOTES: Variable names and definitions are shown in Table 1. In these regressions, each hospital's observation is weighted by the number of its Medicare discharges.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

input prices, and that number of beds reflects more costly facilities and technology, and possibly severity of illness. Both MSA size and number of beds likely also reflect the effects of practice pattern differences among hospitals. The estimated impact of these variables is examined using separate urban and rural regressions and is presented in Tables 10 and 11.

The implications for a bed-size adjustment to PPS rates for urban hospitals differ, depending on the specification used. When outliers are not accounted for, there is a positive relationship between the number of beds and Medicare costs. Hospitals in the first 2 bed deciles (up to 120 beds) are considerably more costly than the smallest urban hospitals. Although generally increasing, the impact of number of beds on costs tends to be much smaller after this point. Such results indicate that a curvilinear adjustment for number of beds might be used, if it were consistent with policy goals. However, when FSTDO is the dependent variable—restricting the teaching and disproportionate-share adjustments to their

**Table 9**  
**Coefficients for low-income-patient-share variables from weighted regressions, using cost per discharge net of estimated outlier payments (STDCSTO) as a dependent variable**

Hospital's share (percent) of low-income patients	Alternative model specification		
	Payment variables only <sup>1</sup>	Fully specified <sup>2</sup>	Fully specified <sup>3</sup>
	Coefficient		
5-10	*-0.01	*0.00	*0.01
10-15	*0.00	*0.02	0.03
15-20	*0.01	0.03	0.04
20-25	*0.01	*0.02	0.03
25-30	*0.02	0.04	0.05
30-35	0.07	0.08	0.08
35-40	0.07	0.07	0.08
40-45	0.15	0.15	0.16
45 or more	*0.00	*0.00	*0.00

\*Not statistically significant.

<sup>1</sup>Includes IRB as an independent variable.

<sup>2</sup>Includes IRB, city size, and number of beds as independent variables.

<sup>3</sup>Includes IRB, city size, and number of beds as independent variables. Also, regions are represented by dummy variables for the four Census Regions. In addition, variables are added representing type of hospital ownership (public or investor).

NOTE: Variable names and definitions are shown in Table 1.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

payment values and accounting for outliers—the cost impact is smaller for the largest hospitals than for medium-size hospitals. These results suggest that the system's current adjustments partially compensate for costs associated with hospital size, and hence, a bed-size adjustment would provide double payment to the largest hospitals. In all specifications, hospitals in MSAs with more than 1 million population are more costly than those in smaller cities, with much of the effect concentrated in the central cities. Hence, there is some empirical support for paying higher rates to hospitals in the largest cities.

For rural areas—excluding designated referral centers—the largest hospitals are significantly more costly than their smaller counterparts, even if all payment parameters are accounted for. One implication is that the system's current adjustments better account for factors that cause large hospitals to be more costly than smaller ones in urban areas than they do for hospitals in rural areas. In addition, hospitals in urbanized rural (URB) counties are more costly than those in less populated counties to about the same extent that large MSAs are more costly than smaller MSAs.

## Considerations and caveats

Clearly, there are many considerations involved in adapting regression analysis to adjusting prospective payment rates. These considerations lead to a wide variety of specifications with varying implications for

**Table 10**  
**Regression results using various measures of urban hospital cost per case as a dependent variable**

Item	Dependent variable				
	MOCC	STDCST	STDCSTO	FSTD	FSTDO
<b>Independent variable</b>	Coefficient				
Constant	7.58	7.59	7.59	7.61	6.61
IRB	0.63	0.60	0.42	( <sup>1</sup> )	( <sup>1</sup> )
CMI	0.89	( <sup>1</sup> )	( <sup>1</sup> )	( <sup>1</sup> )	( <sup>1</sup> )
WI	0.81	( <sup>1</sup> )	( <sup>1</sup> )	( <sup>1</sup> )	( <sup>1</sup> )
MPOP2	*0.01	*0.01	*0.00	*0.01	0.00
MPOP3	0.04	*0.05	0.03	0.05	0.03
CC	*0.07	*0.09	0.09	0.07	0.06
LOW1	*0.01	*0.01	*0.01	( <sup>1</sup> )	( <sup>1</sup> )
LOW2	0.03	0.04	0.03	( <sup>1</sup> )	( <sup>1</sup> )
LOW3	0.04	0.04	0.02	( <sup>1</sup> )	( <sup>1</sup> )
LOW4	*0.02	*0.02	*0.00	( <sup>1</sup> )	( <sup>1</sup> )
LOW5	0.09	0.09	0.07	( <sup>1</sup> )	( <sup>1</sup> )
BED2	0.11	0.01	0.09	0.01	0.08
BED3	0.19	0.18	0.15	0.17	0.13
BED4	0.21	0.19	0.17	0.18	0.14
BED5	0.22	0.19	0.17	0.18	0.14
BED6	0.24	0.23	0.19	0.21	0.16
BED7	0.27	0.24	0.19	0.23	0.16
BED8	0.24	0.21	0.16	0.19	0.12
BED9	0.27	0.24	0.18	0.22	0.13
BED10	0.27	0.24	0.18	0.21	0.10
	Regression statistic				
R <sup>2</sup>	0.58	0.27	0.18	0.13	0.17

\*Not statistically significant at the 5-percent level for a two-tailed test.

<sup>1</sup>Variable not included.

NOTE: Variable names and definitions are shown in Table 1.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

**Table 11**  
**Regression results using various measures of rural hospital cost per case as a dependent variable**

Item	Dependent variable		
	MOCC	FSTD	FSTDO
<b>Independent variable</b>	Coefficient		
Constant	*7.63	7.58	7.57
CMI	1.12	( <sup>1</sup> )	( <sup>1</sup> )
WI	1.13	( <sup>1</sup> )	( <sup>1</sup> )
IRB	0.68	( <sup>1</sup> )	( <sup>1</sup> )
URB	0.05	0.06	0.06
LOW1	*0.00	( <sup>1</sup> )	( <sup>1</sup> )
LOW2	*0.02	( <sup>1</sup> )	( <sup>1</sup> )
LOW3	*0.02	( <sup>1</sup> )	( <sup>1</sup> )
LOW4	*0.00	( <sup>1</sup> )	( <sup>1</sup> )
LOW5	-0.03	( <sup>1</sup> )	( <sup>1</sup> )
RBED1	*-0.03	-0.04	-0.04
RBED2	*0.00	*-0.02	*-0.02
RBED3	0.04	*-0.03	0.40
RBED4	0.04	*-0.03	*0.02
RBED5	0.06	0.05	0.05
RBED6	0.07	0.06	0.06
RBED7	0.09	0.09	0.08
RBED8	0.12	0.11	0.10
RBED9	0.16	0.15	0.13
	Regression statistic		
R <sup>2</sup>	0.35	0.08	0.06

\*Not statistically significant at the 5-percent level for a two-tailed test.

<sup>1</sup>Variable not included.

NOTES: Variable names and definitions are shown in Table 1. Rural hospitals designated as referral centers under the prospective payment system are excluded.

SOURCE: Health Care Financing Administration, Bureau of Data Management and Strategy: Data from the Medicare Provider Analysis and Review file, Medicare Cost Reports, and the Provider-Specific file.

parameters of interest and therefore, for payment rates. The results presented in this article demonstrate the wide variety of estimates that might be obtained. For example, using PPS1 cost data, the estimated indirect-teaching effect consistent with current methods was 6.8 percent—rather than the 7.7 percent now used to determine payments. Other specifications, reflecting different policies and payment goals, resulted in estimates for the indirect-teaching effect as low as 1.7 percent, or as high as 19 percent for rural hospitals. Moreover, the threshold levels important to the system's disproportionate-share adjustment are substantially affected by the alternative regression specification. In some cases, the basic policy implications of the results may change depending on the specification—for example, whether a bed-size adjustment for urban hospitals is supported by the data.

One must bear in mind an important point concerning this analysis: Because we do not use purely statistical criteria for choosing a best estimate, none of these specifications is necessarily right or wrong for payment purposes. The preferability of any given specification depends on the goals and objectives set forth by policymakers, as well as on payment principles that are an integral part of the reimbursement system. For this analysis, it is more important to be able to explain the implications for the payment system of the various choices than it is to choose a right or wrong specification.

In this context, hospital simulation models can be used in conjunction with regression estimates to quantify the total and distributional impact on payments of the various estimates. For example, the results imply that a teaching adjustment of 6.8 percent, rather than the present 7.7 percent, would better reflect current policy, and an adjustment of 5.5 percent would better account for the double payment associated with the disproportionate-share adjustment. However, such changes could have substantial impact on particular hospitals. If the 5.5-percent adjustment were implemented, payments to teaching hospitals would be reduced by about \$550 million, with the largest impact being on central city hospitals. Thus, policy decisions might be made by balancing the effects estimated by regression and the payment impacts found for various hospitals using simulation models. For example, the regression effects might be implemented to improve the way payments are targeted for particular causes but could be phased in over a longer period of time to avoid large, sudden impacts on particular hospitals.

In a broader context, it is also important to consider what regressions can do and what they cannot do. They can be helpful in gauging the magnitude and efficacy of

particular adjustments to PPS's national payment rates by providing estimates of the relationship between variables with whatever payment restrictions are desired. Moreover, year-to-year changes in important parameters resulting from the release of new cost data can be readily observed. In this way, regressions can provide information for short-run adjustments and corrections to the system. Given the current scope and availability of data, it cannot provide us with the more detailed answers about hospital cost variation that are important for future payment policy, however. Although we can estimate that hospitals in larger cities are more costly than their smaller city counterparts, understanding how this difference is affected by such factors as severity of illness, practice patterns, and nonlabor input prices will still require considerable time and research.

## Acknowledgments

The author gratefully acknowledges the technical assistance of Jodi Korb of the Congressional Budget Office. The article benefited from the comments of James Bentley, Philip Cotterill, Julian Pettengill, Lisa Potetz, Bruce Steinwald, and three anonymous reviewers.

## References

- Anderson, G. F., and Lave, J.R.: Financing graduate medical education using multiple regression to set payment rates. *Inquiry* 23(2):191-199, Summer 1986.
- Breyer, F.: The specification of a hospital cost function—A comment on recent literature. *Journal of Health Economics* 6(2):147-157, 1987.
- Evans, R.C.: Behavioral cost functions for hospitals. *Canadian Journal of Economics* 5:198-214, 1971.
- Lave, J.R.: Hospital reimbursement under Medicare. *Milbank Memorial Fund Quarterly* 62(2):251-268, 1984.
- Lave, J.R.: The Medicare Adjustment for the Indirect Costs of Medical Education: Historical Development and Current Status. Washington, D.C. Association of American Medical Colleges, Jan. 1985.
- Lave, J.R., and Lave, L.B.: Hospital cost functions. *Annual Review of Public Health* 5:193-213, 1984.
- Sheingold, S.: The Indirect Teaching Adjustment to Medicare's Prospective Payment System. Washington, D.C. Congressional Budget Office, May 1985.
- Welch, W.P.: Do all teaching hospitals deserve an add-on payment under the prospective payment system? *Inquiry* 24(3):221-232, Fall 1987.