A MEASURE OF PRODUCTIVE EFFICIENCY WITH APPLICATION IN INCENTIVE REIMBURSEMENT FOR HOSPITAL CARE

Dissertation for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

LYCURGUS LYCURGUS LIAROPOULOS

1973

This is to certify that the

thesis entitled

A MEASURE OF PRODUCTIVE EFFICIENCY WITH APPLICATION IN INCENTIVE REIMBURSEMENT FOR HOSPITAL CARE

presented by

Lycurgus Lycurgus Liaropoulos

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Economics

Mitchell Stengel
Major professor

Date July 24, 1973

O-7639

A MEA
APPLI

The p

aspects of the

services and t

produced. Th

to the notion t

efficiently, th

compromising

The

to the various

which make p

pendent on th

operates. U

obviously ess

centive reim

of this thesis

ABSTRACT

A MEASURE OF PRODUCTIVE EFFICIENCY WITH APPLICATION IN INCENTIVE REIMBURSEMENT FOR HOSPITAL CARE

by Lycurgus Lycurgus Liaropoulos

The purpose of this dissertation is to investigate certain aspects of the relationship between the cost of producing hospital services and the level of efficiency with which such services are produced. The skyrocketing cost of hospital care has given rise to the notion that, if hospitals are induced to operate more efficiently, the cost of hospitalization can be contained without compromising the quality of care.

The relationship between costs and efficiency is central to the various recent incentive reimbursement proposals and plans which make payments to hospitals by the various third parties dependent on the level of efficiency with which each institution operates. Unfortunately, a satisfactory measure of efficiency, obviously essential to the equitable and effective application of incentive reimbursement, has not yet been developed. The objective of this thesis, therefore, is to construct and test a measure of

hospital cost

than the ofte

The

constructed

and it incorpo

are adjusted

pital output in

pediatrics, ge

average cost

through regres

comparing a h

responding san

hospital's case

The se

pital length of s

care produced.

positive but dec

a given hospital

hospital costs

given institutio

The r

per case are t

hospital costs which is also a more accurate measure of efficiency than the often used average cost per case or per patient day.

The proposed measure of costs, or "costliness index", is constructed for a sample of 94 Michigan short-term general hospitals, and it incorporates two types of adjustments. First, hospital costs are adjusted for differences in patient-mix by disaggregating hospital output into six types of cases: medical-surgical, obstetrics, pediatrics, geriatrics, psychiatric cases, and outpatient care. Average cost weights for each of the six case-types are derived through regression analysis, and an index number is developed comparing a hospital's costs for specific case-types with the corresponding sample average and weighting by the composition of the hospital's casemix.

The second adjustment assumes that differences in hospital length of stay imply differences in the actual amount of patient care produced. A logarithmic transformation is used to assign positive but decreasing weights to each additional patient day within a given hospital stay. This transformation, therefore, adjusts hospital costs for the actual amount of patient care produced by a given institution relative to the sample average.

The resulting costliness index and the actual average cost per case are then shown to have radically different reimbursement mplications.

hypothetical in

financial rewar

or relative ave

efficiency. Th

with regard to

are adjusted fo

and teaching pr

A fina

costliness and

constructed fro

hospital produc

closer relations

The theoretical

as the empiric

measure of hos

ment.

implications. For more than a quarter of the hospitals studied, a hypothetical incentive reimbursement plan provides the opposite financial rewards and penalties depending on whether costliness or relative average cost per case are used as measures of hospital efficiency. The distinction between costliness and average cost with regard to reimbursement holds even when the two measures are adjusted for factors such as location, facilities and services, and teaching programs.

A final step is to test the actual relationship between costliness and efficiency. This is done via a productivity index constructed from the residuals of an estimated Cobb-Douglas hospital production function. The productivity index shows a closer relationship to costliness than to average cost per case. The theoretical properties of costliness index, therefore, as well as the empirical findings, suggest the use of costliness as a measure of hospital costs for the purposes of incentive reimbursement.

A MEAS
APPLIC

in ;

A MEASURE OF PRODUCTIVE EFFICIENCY WITH APPLICATION IN INCENTIVE REIMBURSEMENT FOR HOSPITAL CARE

by

Lycurgus Lycurgus Liaropoulos

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

1973

TO
YOLANDA

Knowin
which I receive
nevertheless. I
following:

To Dr.
ture into the fie
advice and direc
Ginsburg who o
of interest;

To the
Department of P
unpublished data

To Pro
couragement and
dissertation;

To my
provide criticis
various crucial

I feel
Miss Sue Peter
over the typing

The fi
some technical
the moral supp
aration poss
and, in particu
Bacopoulos.

ACKNOWLEDGEMENTS

Knowing that I cannot fully account for all the assistance which I received throughout the preparation of this thesis, I would, nevertheless, like to express my gratitude and appreciation to the following:

To Dr. Mitchell Stengel who, with his willingness to venture into the field of Medical Economics, provided me with sound advice and direction throughout the study. Also, to Dr. Paul Ginsburg who offered valuable assistance on many important points of interest;

To the American Hospital Association and the Michigan Department of Public Health, which provided me with necessary unpublished data:

To Professors Jan Kmenta and C. P. Larrowe whose encouragement and support were essential at different stages of this dissertation;

To my good friend Harold Reinholds whose willingness to provide criticisms and to suffer through tedious discussions of various crucial concepts was above and beyond the call of friendship.

I feel personally responsible for the nightmares which Miss Sue Peterson must undoubtedly have suffered as she laboured over the typing of the manuscript.

The final thanks is for a group of friends who provided some technical and conceptual assistance, but, more importantly, the moral support and sense of direction which made this dissertation possible. I wish to thank my friends in East Lansing, and, in particular, Harold Reinholds, Evan Jones, and Alex Bacopoulos.

TABLE OF CONTENTS

Chapter

Appendix

Table

LIST OF TABLES

Table		Page
13	Actual Reimbursement Amounts to High Cost-Low Costliness Hospitals	108
14	Distribution of Cr and C* by Location	113
15	Number of Hospitals Where Regional Average Cost and C* Move in Opposite Directions	115
16	Classification of Hospitals by Blue Shield Prevailing Areas	117
17	Total Reimbursement Amounts Based on Regional Average Cost and Costliness Estimates (in millions)	118
18	Cost and Productivity Performance According to Cr, C*, Pr and P* for 22 Hospitals	150
19	Correlations Among Number of Cases in Each Case-type and Size of Hospital	174
20	Correlations Among Proportions of Cases in Each Case-type.	176
21	Average Cost Per Day and Per Visit by Case-type	181
22	Results From Regression Number One . . .	191
23	Results From Regression Number Two . . .	192

Figure

LIST OF FIGURES

During

the most explos

charges, in fac

Consumer Pric

155.5 percent a

prices and a 3

commodities c

service charge

as all other pr

cern among th

of the annual

Cross, and pr

The

represents on

spiraling cos

the increase

the improved

hospital as th

CHAPTER I

INTRODUCTION

During the past ten years hospital prices have been among the most explosive of all consumer prices. Hospital daily service charges, in fact, represent the fastest rising component of the Consumer Price Index. From 1960 to 1970, these charges increased 155.5 percent as compared to a 52.5 percent increase in all medical prices and a 31.2 percent increase in the prices of all consumer commodities combined. During the last half of the period, daily service charges increased by 87.8 percent, or four times as fast as all other prices. Such dramatic increases have caused great concern among the various third parties responsible for over 85 percent of the annual payments to hospitals, namely, government, Blue Cross, and private insurance companies.

The concept of incentive reimbursement for hospital care represents one of the recent attempts to deal with the continuously spiraling cost of hospitalization. It is undoubtedly true that part of the increase in hospital costs is due to necessary wage increases, the improved quality of hospital care, and the expanding role of the hospital as the central provider of medical care with an ever increasing conc

required to prov

this time, they

proponents of in

crease in prices

the hospital to k

culprit, as they

hospitals are pa

of charges whic

The m

ments to hospit

efficiency with

financial rewar

pitals and by pe

tive reimburse

can be modera

bursement is a

to bear on an i

determined on

considerations

If ho

level of effici

important.

increasing concern that costs have already gone beyond the levels required to provide high quality care, and that, from indications at this time, they may climb even faster in the immediate future. The proponents of incentive reimbursement maintain that the rapid increase in prices is partly due to a lack of economic incentives for the hospital to keep operating and fixed costs down. The ultimate culprit, as they see it, is the financing mechanism through which hospitals are paid either on the basis of full costs, or on the basis of charges which they are at relative liberty to set.

The main idea behind incentive reimbursement is that payments to hospitals should be at least partly related to the degree of efficiency with which a given institution operates. By offering financial rewards in the form of higher payments to efficient hospitals and by penalizing the inefficient ones, the advocates of incentive reimbursement argue that increases in the cost of hospital care can be moderated. From an economic point of view incentive reimbursement is an attempt to bring economic variables and incentives to bear on an industry where output and costs have traditionally been determined only on the basis of medical, ethical, and professional considerations.

If hospitals are to be evaluated and paid according to their level of efficiency, then the question of measuring efficiency becomes important. As we will see in Chapter II, most incentive reimbursement schemes rate h

per case or per

composed of fir

unit cost is dir

case in the hos

ence is that th

cept because h

services which

quirements.

meaningless u

certain adjus

The

average cost

by the variou

mate measur

portant objec

measure of h

ship to effic

adjusts hosp

length of sta

of output he

becomes a

schemes rate hospitals according to some estimate of average cost per case or per patient day. Although in a profit-oriented industry composed of firms manufacturing a homogeneous product average unit cost is directly related to productive efficiency, this is not the case in the hospital industry. The basic reason behind this difference is that the unit of hospital output is not an easily defined concept because hospitals produce varying amounts of a wide mix of services which are not directly comparable in terms of input requirements. As a result, cost comparisons among hospitals are meaningless unless the statistical or actuarial techniques used make certain adjustments for the basic heterogeneity of hospital output.

The first objective of this thesis is to show that the use of average cost per case or per patient day, although often suggested by the various incentive reimbursement proposals, is an inappropriate measure of hospital efficiency. The second and more important objective is to propose, estimate, and evaluate an alternative measure of hospital cost performance which bears a closer relationship to efficiency of operation. This proposed "costliness index" adjusts hospital average costs for differences in casemix and in length of stay for various types of cases. Two of the basic sources of output heterogeneity are, therefore, removed, and cost performance becomes a better indicator of hospital efficiency.

The se

theory of hospi

bursement mec

and a variant i

hospitals try to

vices subject t

the physician s

operating cos

creased reven

hospital reim

and payment c

objectives wi

examines the

under which

costs but rat

institution o

reimburseme

improvemen

only be achi

part of the v

Ch

measuring

Chapter III

The second chapter of the thesis deals with the economic theory of hospital behavior and the implications of alternative reimbursement mechanisms. Various suggested theories are reviewed, and a variant is developed based on the proposition that nonprofit hospitals try to increase the quality and extensiveness of their services subject to meeting a largely exogenous demand originating with the physician staff. This implies a tendency to increase fixed and operating costs continuously, and, therefore, a constant need for increased revenues. After a review of the major current forms of hospital reimbursement it is shown that both full cost reimbursement and payment on the basis of charges allow the hospital to attain its objectives with little regard to cost and efficiency. The thesis then examines the various incentive reimbursement plans and proposals under which payments are no longer in direct proportion to hospital costs but rather depend on the degree of efficiency with which an institution operates. The analysis ultimately shows that under such a reimbursement method efficiency incentives do exist, and that quality improvements and increases in the scope of services in a hospital can only be achieved with a greater amount of cost consciousness on the part of the various decision makers.

Chapters III through VI deal with the actual question of measuring efficiency for the purposes of incentive reimbursement. Chapter III deals with the question of output heterogeneity and defines output as the to

during the year.

of stay differen

are rejected as

oped which as

and to the amou

stay. The res

patient care" a

Chapter VI.

Part

to show that av

factors which

location of the

wage different

Hospital aver

ences in case

costliness in

mance reflec

that a hospita

casemix or b

reimbursem

efficient than

output as the total amount of inpatient and outpatient care produced during the year. Because of the importance of casemix and length of stay differences among hospitals, the case and the patient day are rejected as measures of patient care. A new measure is developed which assigns different weights to patients of different types and to the amount of output produced during each day of a patient's stay. The resulting new measure of output is called "adjusted patient care" and is used in an analysis of hospital productivity in Chapter VI.

Part of the analysis in Chapter III is used in Chapter IV to show that average cost per case must be adjusted for a variety of factors which affect costs but are not related to efficiency. The location of the hospital is used as a surrogate for factors such as wage differentials, facilities and services, and teaching programs. Hospital average cost per case is then adjusted directly for differences in casemix and length of stay. The resulting index number, or costliness index, is suggested as a measure of hospital cost performance reflecting efficiency differences among hospitals. It is shown that a hospital showing high average costs because of an "expensive" casemix or because of long average stays would not be penalized by a reimbursement plan based on costliness unless it was actually less efficient than the average hospital in a certain population of institutions.

Chapte

ndex and the in

cost as the bas

94 Michigan sh

forty percent o

results show th

costliness and

bursement amo

show high ave

tensive mix of

hospitals, on

show that this

various instit

short lengths

costliness as

payments to

could often re

quality, spec

Cha

efficiency. I

productivity

dexes are es

examined for

Chapter V presents an empirical analysis of the costliness index and the implications of using costliness as opposed to average cost as the basis for reimbursement. The sample examined includes 94 Michigan short term general nonprofit hospitals which represent forty percent of the industry, and the data are for the year 1969. The results show that for roughly a quarter of the hospitals examined, costliness and relative cost would result in radically different reimbursement amounts. It is observed that urban hospitals tended to show high average costs but lower costliness indicating a more extensive mix of services, but also a higher degree of efficiency. Rural hospitals, on the other hand, have low average costs, but the results show that this is not due to a high degree of efficiency but rather to various institutional factors as well as inexpensive casemixes and/or short lengths of stay. The most important conclusion is that the use of costliness as the reimbursement standard would actually relate hospital payments to efficiency. On the other hand, the use of average costs could often result in rewards for inefficiency or penalties for high quality, specialized, and therefore, expensive hospital care.

Chapter VI examines the formal relationship between costs and efficiency. Productive efficiency is seen to include two elements, namely, productivity and input efficiency. The two corresponding efficiency indexes are estimated from a Cobb-Douglas production function and examined for their relationship to costliness and average relative costs.

Analysis of the
costliness is s
The results fro
this is explaine
and costs are n
combinations.

Chapte
conclusions, a

Analysis of the productivity index gives additional indications that costliness is superior to average costs as a measure of efficiency. The results from the input efficiency index are inconclusive, but this is explained by the fact that the decisions affecting productivity and costs are made by agents other than those determining input combinations.

Chapter VII contains a summary of the thesis, the major conclusions, and certain implications and policy recommendations.

HO

IN

Durin

amounted to $

direct paymen

cent was reim

insurance, go

in particular,

creased by th

and Medicaid

for 50 percen

This

distinct featu

on the produ

enactment o

dicted that t

to prove hig

number of i

consumers

CHAPTER II

HOSPITAL REIMBURSEMENT, EFFICIENCY INCENTIVES, AND ECONOMIC BEHAVIOR

During 1971 total national expenditures for hospital care amounted to $29.6 billion. Of this only 13 percent represented direct payments by patients to hospitals while the remaining 87 percent was reimbursement by various "third parties" such as private insurance, government, and Blue Cross.[1] The role that government, in particular, plays in the financing of hospital care was greatly increased by the enactment of the 1965 Social Security Act (Medicare and Medicaid), to the point where public revenues alone now account for 50 percent of all expenditures for hospital care.[2]

This virtual separation of consumption from payment is a distinct feature of the hospital industry, and it has had serious effects on the production and cost of hospital services. At the time of the enactment of the 1965 Social Security Act, a number of writers[3] predicted that the current methods of financing hospital care were likely to prove highly inflationary for three major reasons. First, as the number of individuals covered by some form of insurance increased, consumers would have less reason to be concerned with the direct cost

of hospitalizati
increase, and
expenditure is
given that phys
care, which is
would increas
ance. Finally
virtually assu
would face fe
Moreover, si
there were no
level of quali

The
come true, m
other factors
care and risi
generally agr
contributory
party paymer
for more and
pitals as to
higher costs
increased de

of hospitalization. As a result consumption of hospital care would increase, and incentives to overspend might exist since higher expenditure is usually associated with higher quality of care.[4] Second, given that physicians are trained to provide the highest quality medical care, which is often also the most expensive, their tendency to do so would increase since the cost to the patient would be of lesser importance. Finally, with an increasing percentage of their revenue being virtually assured by third party payments, hospital administrators would face fewer pressures to achieve reductions in operating expenses. Moreover, since reimbursement was often based on incurred costs there were no penalties for higher costs, which presumably raise the level of quality in an institution.

The prediction of "rapidly rising hospital costs" has definitely come true, making the expression a painful household word. Although other factors such as increasing demand for more and higher quality care and rising labor costs should not be underestimated, it is also generally agreed that the financing mechanism has been an important contributory factor. One writer in particular, sees the growth of third party payments as responsible for a vicious cycle of increased demand for more and more expensive care, which gives false signals to hospitals as to "necessary" expansion or quality improvements, leading to higher costs, more comprehensive insurance policies and again increased demand.[5] Since third party payments are here to stay and most likely will

health insuranc

of reimbursem

obvious that pa

provide the va

economic effic

this realizatio

ment to the fo

we deal with t

methods of ho

efficiency and

Unti

was on the ba

for services

directly, or

hospital. Th

the second,

clusively us

main form o

Th

in 1971 rang

most likely will increase with the enactment of some form of national health insurance, attention has been focused increasingly on the methods of reimbursement used by the various third parties. It is becoming obvious that payments to hospitals should be made in such a way as to provide the various institutions with sufficient incentives to promote economic efficiency without compromising the quality of care. It is this realization that has brought the concept of incentive reimbursement to the forefront of recent research in hospital economics. Before we deal with this, however, we must first discuss the major current methods of hospital reimbursement and their implications for hospital efficiency and cost behavior.

Current Methods of Reimbursement

Until fairly recently the most common type of hospital payment was on the basis of charges. Under this system the hospital sets charges for services rendered and the third party either pays the hospital directly, or payments are made to the patient who, in turn, pays the hospital. The first form is still used by a few Blue Cross plans while the second, which is also called "indemnity payment," is almost exclusively used by commercial insurance carriers and represents the main form of charge reimbursement.

The typical payment for room and board by insurance companies in 1971 ranged from $40 to $50 per day[6] with any charges in excess of

this amount pa
few contracts
accommodatio
plans is 70 to
30 days to two
made for othe
operating and
age varies wid
paying only fo
up to a limit,
point.

What
insurance the
One result of
of hospitaliza
ubles, coins
cost of insur
price mainly
tion of hospi
price of hos
sumer's den
services tha
dictate. Th
ment, but it

this amount paid by the patient or other supplementary insurance. Very few contracts pay the full charge for allowable (usually a two-bed room) accommodations. The maximum hospitalization covered under most plans is 70 to 120 days although some contracts exist with periods from 30 days to two years. Besides the daily charge, payments are usually made for other hospital services such as x-rays, drugs and dressings, operating and delivery room use, anesthetics, tests and others. Coverage varies widely with regard to these services with some contracts paying only for specific services while others cover all service charges up to a limit, with coinsurance clauses[7] becoming effective after that point.

Whatever the exact nature of hospital payments by private insurance the main point is that they are made on the basis of charges. One result of this method of reimbursement is that the effective price of hospitalization to the consumer is reduced to the amounts of deductibles, coinsurance, if any, and charges for uncovered services. The cost of insurance premiums to the consumer is not part of the effective price mainly because it is essentially unrelated to the actual consumption of hospital services[8] at the time of need. The reduction of the price of hospital care, often to near zero, certainly affects the consumer's demand by allowing him to purchase more and higher quality services than his income or even possibly his medical needs would dictate. This is obviously an inflationary aspect of charge reimbursement, but it is shared by all other forms of medical insurance regardless of

What distingui

surance paym

The s

this chapter w

does not seek

zation of the

as perceived

traditionally

and philanthr

revenue from

determine an

accordingly,

output. This

ditions which

patient's bil

is less conc

the bill hims

fairly inelas

crease in pr

larger perc

ticity of den

regardless of the method of payment used or the third party involved. What distinguishes charge reimbursement from other types of insurance payments is the set of financial incentives it affords hospitals.

The simple economic model constructed in a later section of this chapter will show that, although the typical nonprofit hospital does not seek to maximize profits, its objectives include the maximization of the quality, quantity, and scope of services. These needs as perceived by the hospital require a certain target revenue which traditionally hospitals have collected from patients or third parties and philanthropy.[9] If a hospital receives a substantial part of its revenue from charge reimbursement, it is in a good position to determine and reach its target revenue either by setting charges accordingly, or, less importantly, by manipulating the volume of output. This situation arises because of the special demand conditions which the hospital faces:[10] Since a substantial part of a patient's bill is usually paid by some form of insurance, the patient is less concerned with the actual hospital price than if he had to pay the bill himself. From the hospital's viewpoint, this results in a fairly inelastic demand, which means that a certain percentage increase in prices (or charges) will increase hospital revenue by a larger percentage. Other factors responsible for the low elasticity of demand are the fact that hospital care is a "need" rather than a "want"

hospital's nor

on its staff an

In vi

the basis of c

have few ince

necessitated

on to third pa

marginal cos

certain quan

surance that

cost of the u

mechanism c

cation of res

the economy

Wi

last twenty

and Medica

the basis of

assessment

vision of se

by the vario

on a cost b

while only

than a "want", especially in emergency cases, and that demand for a hospital's nonemergency services largely depends on the physicians on its staff and not on consumer discretion.

In view of the low elasticity of demand, reimbursement on the basis of charges has two major implications. First, hospitals have few incentives for economic efficiency since cost increases necessitated by possible inefficiencies can be, at least in part, passed on to third parties simply by charge increase. Second, the lack of marginal cost pricing by the hospital means that even if it provides a certain quantity and quality of services efficiently, there is no assurance that the cost of a unit of services to the hospital equals the cost of the use of resources to society. As a result, the price mechanism does not work as a signaling device for the correct allocation of resources between the hospital industry and other sectors of the economy.

With the development of the various Blue Cross plans in the last twenty years, and especially after the introduction of Medicare and Medicaid in 1965, hospital reimbursement shifted to payment on the basis of costs. This form of payment is based upon third party assessments of the actual costs incurred by an institution in the provision of services to subscribers. Of the $24.8 billion paid to hospitals by the various third parties during 1971, $19.8 billion was reimbursed on a cost basis ($14.8 billion by government and $5.0 by Blue Cross), while only $5.0 billion was reimbursed on a charge basis.

In th

by insurance

is identical to

does not ente

profit nature.

any point of

independent c

Here again,

formance de

of the admin

expansion, a

financed ini

and recover

Th

if ever, enc

fining allow

interest on

limit the ex

are only eff

Moreover,

creases are

It is obviou

In the pure case where all patients in a hospital are covered by insurance plans which pay on the basis of full costs, the situation is identical to that of charge reimbursement. Since economic profit does not enter the cost functions of hospitals because of their non-profit nature, average cost becomes the effective price. Since at any point of production all costs are met, the hospital's output is independent of its cost structure and of the demand for its services. Here again, efficiency is of secondary importance, and cost performance depends upon noneconomic criteria such as the priorities of the administrators and medical staff. Salaries, bonuses, capital expansion, and the addition of new facilities and services can be financed initially from the private sector, recorded as new costs, and recovered through the cost reimbursement mechanism.

This type of full cost reimbursement, however, is rarely, if ever, encountered. Third parties often impose stipulations defining allowable costs, and placing limitations on depreciation, interest on loans, and permissible cost increases. These constraints limit the extent to which hospitals can increase costs at will, but they are only effective for cost increases exceeding the allowable limits. Moreover, there are considerable ambiguities as to what cost increases are necessary for quality improvements or facility expansion. It is obviously true that the ability of third parties to monitor cost

estimates (or
Fairly often,
or third part
hospital cost

The
of cost utiliz
reimbursabl
use differen
portant diffe
Blue Cross
wise allowa
cost conditi
labor costs
hospital ind
Medicare a
the plus fac
that the pol
expansion c
incentive to

W
nature of c
they have h
The fact th

estimates (or bills) submitted by hospitals is less than perfect.[11] Fairly often, either accounting methods vary among institutions, or third parties lack sufficient resources for a thorough audit of hospital cost reports.

The various third parties which reimburse on the basis of cost utilize a variety of specific formulas in determining the reimbursable amounts. Medicare, Medicaid, and Blue Cross also use different limitations and ceilings on their payments. One important difference is the use of a community or "plus" factor by most Blue Cross plans. This is a payment allowance in addition to otherwise allowable costs in recognition of unaccounted costs or of special cost conditions prevailing in certain communities, such as higher labor costs. This cost-plus factor has also been defended by the hospital industry as a necessary growth factor. Initially both Medicare and Medicaid made a similar allowance. In 1969, however, the plus factor was dropped from both programs because it was found that the policy encouraged duplication, overlapping, and unnecessary expansion of facilities and services and created an unhealthy economic incentive to maximize operating costs.

Whatever the specific forms of cost reimbursement or the nature of ceilings and limitations, the question still remains whether they have been effective in containing costs and promoting efficiency. The fact that hospital costs have been rising more than three times as fast as the

that they have

estimate emp

costs. Pauly

tinguish betw

status in an a

from four sta

point in time

bursement ef

the model by

areas of the

measure of c

coverage. S

rejection of

siveness of c

of the two st

question they

cost inflation

bursement.

significant d

The nonprof

be more or

ment a hosp

as fast as the consumer price index suggests perhaps tenuously, that they have not. There have also been at least two attempts to estimate empirically the impact of cost reimbursement on hospital costs. Pauly and Drake[12] use a simple dummy variable to distinguish between cost reimbursement and charge reimbursement status in an average cost regression based on a sample of hospitals from four states (Illinois, Indiana, Michigan, and Wisconsin) at one point in time. They conclude that there is no significant cost reimbursement effect on costs per patient day. K. Davis[13] elaborates the model by including data before and after Medicare, from all areas of the country, and by incorporating a more comprehensive measure of cost reimbursement varying with the extensiveness of coverage. She also concludes that "the empirical results lead to the rejection of the hypothesis that hospital costs increase with the extensiveness of cost reimbursement within the range observed."[14] Neither of the two studies tests the relevant hypothesis, however, since the question they ask is not whether cost reimbursement in itself leads to cost inflation, but whether it is more inflationary than charge reimbursement. Based on the previous discussion we would expect no significant differences in cost behavior under those two systems. The nonprofit nature of the hospital implies that average revenues will be more or less in line with average costs. Under charge reimbursement a hospital would normally attempt to have some "profit" for

improvements
improvements
able costs fo
that cost rei
to curb hosp
hospitals wi
interest in i
such a reali

The
officially in
the Secretar
hospital pay
Child Healt
ing reimbur
while provid
lower progr
to meet the
cost increa
hospitals s
gains in ef

improvements, while under full cost reimbursement the cost of these improvements, whenever they are made, will be added to reimbursable costs for the next period. As a result, it has become apparent that cost reimbursement, just like charge reimbursement, has failed to curb hospital cost inflation mainly because it has failed to provide hospitals with an adequate set of efficiency incentives. The current interest in incentive reimbursement is the natural consequence of such a realization.

Incentive Reimbursement

The need for incentive reimbursement was recognized officially in the 1967 Social Security Amendment, which authorized the Secretary of HEW to experiment with alternative methods of hospital payments under the Medicare, Medicaid, and Maternal and Child Health Programs. The provision reflected interest in developing reimbursement methods which would support high quality services while providing incentives for efficiency and economy and leading to lower program costs.[15] The purpose of incentive reimbursement is to meet the financial needs of hospitals in such a way as to slow down cost increases without a deterioration in the quality of care. Ideally, hospitals should be reimbursed so that those institutions which show gains in efficiency while maintaining quality are rewarded and those

which appear
levels of qua

The
which can be
prospective
the concept o
hospital ope
hospital dep
savings, and
from sugges
year depend
target budge
to the needs
savings tend
ments, per
the most in
ous medica
and elimina
ductivity is
and, second
dependent p
eventually
geting and

which appear less efficient or show lower costs because of lower levels of quality are penalized.

There is a variety of incentive reimbursement proposals which can be differentiated according to whether payment is on a prospective or retrospective basis.[16] In the first category belongs the concept of prospective budgeting. This involves a survey of hospital operations in order to assess the particular needs of each hospital department. The survey determines sources of potential savings, and a target budget is prepared based on savings expected from suggested cost reductions. Incentive payments at the end of the year depend on the extent to which each hospital has stayed within the target budget. The advantage of this approach is that it can be tailored to the needs of each individual institution. The disadvantage is that savings tend to be small[17] and concentrated on the nonmedical departments, perhaps because this is where the hospital administrator has the most influence. The more substantial savings possible in the various medical departments from improvements in utilization patterns and elimination of inefficiency are hard to achieve, first because productivity is difficult to measure in order to set the necessary targets and, second, because such savings often depend on cooperation by independent physicians. Certain experiments currently under way may eventually provide more evidence on the efficacy of prospective budgeting and industrial engineering techniques. Among these are a plan by the Connec

Blue Cross o

Ano

recently reco

It calls for a

factors other

burse hospit

second sugg

individual ho

reimbursem

fall below th

ed methods

could involv

the rate of i

capital fund

Mo

of a retrosp

upon some e

and predete

pare hospit

group of si

form of rev

by the Connecticut Hospital Association, and two experiments by Blue Cross of Southern California and Western Pennsylvania.[18]

Another form of prospective reimbursement has been recently recommended by the American Hospital Association.[19] It calls for a formula to be negotiated in advance, which depends on factors other than incurred costs. One method would be to reimburse hospitals with a fixed amount per patient day or per case. A second suggestion is to set target rates of cost increase, either for individual hospitals or for groups of similar institutions. Incentive reimbursement would then take the form of rewards if actual costs fall below the target or of penalties for cost overruns. The suggested methods of rewards and penalties vary. Penalties, for example, could involve reimbursement of less than actual incurred costs if the rate of increase exceeds the target, or a smaller allocation of capital funds in the future.

Most of the current incentive reimbursement proposals are of a retrospective nature, which would make payments dependent upon some evaluation of incurred costs rather than upon some desired and predetermined standard of performance. Such plans would compare hospital cost performance with the average performance in a group of similar hospitals and would make incentive payments in the form of rewards or penalties. Since the groups of hospitals are usually com

is also call

On

adopted in 1

member hos

on location

teaching pr

Reimbursen

costs relati

The unit of

tive reimbu

More speci

ceiling (whi

hospital re

appeals me

sidered un

scheme is

with lower

A

proposal is

pitals[21] in

adjustmen

reimburse

usually composed of hospitals in different geographical regions, this is also called Regional Average Cost Reimbursement.

One such specific incentive reimbursement plan is the one adopted in 1966 by Blue Cross of Western Pennsylvania and its member hospitals.[20] Nine groups of institutions are formed based on location (metropolitan, urban, or rural) and the nature of their teaching programs (advanced teaching, teaching, and nonteaching). Reimbursement at the end of each period is on the basis of actual costs relative to the mean cost of the other hospitals in the group. The unit of measurement is average cost per patient day, and incentive reimbursement takes the form of penalties for excessive costs. More specifically, if a hospital has costs in excess of the group ceiling (which is set at 10 percent above the group mean), the hospital receives only the ceiling rate. The plan provides for an appeals mechanism to handle cases where reimbursement is considered unfair by the hospital. One drawback of this particular scheme is that it provides no positive incentive payments to hospitals with lower than average costs.

A considerably more sophisticated incentive reimbursement proposal is made in a recent study of 93 Western Pennsylvania hospitals[21] in which the authors recognize the need for considerable adjustments before meaningful cost comparisons among hospitals for reimbursement purposes can be made. Using multiple regression techniques th

size, non-ro

grams, case

Since a satis

assume that

cation which

in their sam

total hospita

by a third p

year. Unde

accepting th

hearing. C

tailed cost i

Th

predictive

and maxim

than the pr

exceed 110

percent of

actual cost

total costs

and penalty

between ac

techniques they estimate the influence on hospital costs of: location, size, non-routine or extraordinary inpatient services, teaching programs, casemix, quality of medical staff, and outpatient activity. Since a satisfactory measure of casemix is not available, the authors assume that it is correlated with an index of medical staff sophistication which they construct from a questionnaire sent to the hospitals in their sample. The six variable are used in a predictive model with total hospital cost as the dependent variable. This model can be used by a third party to determine a hospital's predictive cost for a given year. Under the proposed plan the hospital has the choice of either accepting the predictive cost or requesting a formal budget review hearing. Out of such a hearing, and on the basis of much more detailed cost information a prospective rate can be set.

The proposed plan provides for certain intervals about the predictive cost which can be used to establish incentives or penalties and maximum reimbursement. If the hospital's actual costs are less than the predicted or negotiated rates, the reimbursement (never to exceed 110 percent of actual costs) will be the actual cost plus a percent of the difference between predicted and actual costs. If actual costs are higher than predicted, the hospital will receive total costs minus a percent of the difference. The percentage reward and penalty factors vary according to the magnitude of the difference between actual and predicted costs.[22]

A f

total costs w
signed to co
model. The
sample and
necessary f

Th
payment for
costs. An
on the aver
reimburser
tal's actua
in costs for
takes the b
the averag

A
of the prob
different h
pared with
of serious
stantially
reimburs
year does

A five percent "slack" between 97.5 and 102.5 percent of total costs where hospitals receive full cost reimbursements is designed to correct for the standard error of estimate in the predictive model. The authors compare predicted and actual costs in their sample and conclude that formal budget review would have been unnecessary for up to 60 percent of the hospitals.

The incentive reimbursement plans mentioned above involve payment formulas based on the absolute amounts of coverage or total costs. An alternative formula, proposed by Saul Waldman,[23] is based on the average increase in costs between two periods. Under this plan reimbursement depends on two major factors: (1) the individual hospital's actual costs in a base period and (2) the average rate of increase in costs for a control group of hospitals. A simple form of this plan takes the base costs of a hospital and allows for an increase equal to the average rate of increase for the control hospitals.

A major advantage of this approach is that it avoids certain of the problems involved in comparing average or total costs in different hospitals. Since an institution's costs in any period are compared with the same hospital's costs during the base period, the risk of serious inequities with possible repercussions on quality is substantially diminished. A high-cost institution could, theoretically, be reimbursed its full costs provided its rate of increase over the previous year does not exceed the average rate for the control group. A disadvantag

seek to hav

and efficie

currently e

such avera

W

ate the sha

depend on t

policies wh

set of objec

will deal in

Th

altering the

pital to inc

of producti

plans with

price whic

shows tha

can affect

ciency of

however,

disadvantage of such a plan, on the other hand, is that it does not seek to have any impact on the original or base period level of costs and efficiency. If, as we have reasons to believe, inefficiencies do currently exist in the hospital industry, they will not be affected by such average cost increase incentive reimbursement plans.

Whether incentive reimbursement can actually help moderate the sharp increases in hospital costs by increasing efficiency will depend on the way in which hospitals react to changes in economic policies which affect their revenues. This, in turn, depends on the set of objectives which determine hospital behavior, with which we will deal in the next section.

Incentive Reimbursement and the Economic Behavior of Hospitals

The basic idea behind incentive reimbursement is that, by altering the payment formula, incentives will be created for the hospital to increase economic efficiency and, thus, to lower the unit cost of production. Most of the retrospective incentive reimbursement plans with which this thesis is concerned would affect the effective price which hospitals receive for their services. Economic theory shows that changes in the price that a firm can charge for its product can affect the volume of output, the quality of the product, the efficiency of operation, and, therefore, the cost of production. Hospitals, however, do not operate like the standard firms of economic theory

in that they
can be sure,
have the des
nature of th
structure of
set of objec
quantity and

Be
sidered ina
on differen
A brief out
develop a t
of the alter

T
recover co
nates from
no interes
Associati
recommen
plant expa
pace with
possibilit
hospital r

in that they do not follow policies of profit maximization. Before we can be sure, therefore, that changes in the payment mechanism will have the desired efficiency incentives, and in order to determine the nature of these changes, we should examine the organizational structure of the hospital, its decision-making mechanism, and its set of objectives, both economic and those associated with the quantity and quality of the services it provides.

Because the traditional model of profit maximization is considered inapplicable to hospital economics, a variety of models based on different behavioral hypotheses have recently been suggested.[24] A brief outline of the major models follows, after which we will develop a theory of hospital economic behavior based on a synthesis of the alternative hypotheses.

The most prevalent view is that hospitals simply attempt to recover costs by setting price equal to average cost.[25] This originates from the belief that hospitals exist to serve the public and have no interest in profits. Guidelines set forth by the American Hospital Association emphasize the recover-of-costs theme. The AHA also recommends that prices should also "cover the funds necessary for plant expansion due to improvement of services required to keep pace with technological and scientific advances."[26] It is precisely the possibility for such a markup that has important implications for hospital reimbursement. If competitive pressures are not important, there is not

additions to

recovery-of

is at a level

costs in the

ments simp

payments a

reduce out

efficiency.

A

zation.[27]

"maximize

possible s

constraint

major imp

price as p

Some evid

set room c

services s

competitio

ally for p

multiple a

there is nothing to guarantee cost minimization or that expansion and additions to services will be always economically justifiable. The recovery-of-costs hypothesis therefore implies that if reimbursement is at a level higher than costs, hospitals will increase their total costs in the next period by spending for expansions or quality improvements simply because the funds are available. If, on the other hand, payments are below average costs, the hospital may be forced to reduce output, lower the quality of care, or increase its level of efficiency.

A different behavioral hypothesis is that of output maximization.[27] It is based on the assumption that hospitals seek to "maximize the welfare of society by serving as many patients as possible subject to certain constraints,"[28] one of which is a budgetary constraint determining the maximum size of the allowable deficit. The major implication of this model is that hospitals will charge as low a price as possible in order to increase the amount of output sold.[29] Some evidence of such behavior may be the fact that hospitals typically set room charges lower relative to costs than charges for ancillary services such as x-ray and laboratory tests.[30] Since some amount of competition among hospitals at the admission stage exists,[31] especially for patients with indemnity coverage whose physicians hold multiple appointments, the demand for hospital routine care is more elastic than

literally in

therefore,

A

an examina

bursement.

creased, t

least its re

sold. Afte

were reim

windfall re

much faste

A

third beha

zation.[33]

hospital to

period for

A variant

used in th

A

maximiza

function f

administr

elastic than that for ancillary services, where the patient is quite literally in a "captive" market. A room rate set lower than cost, therefore, may be an attempt by the hospital to maximize its output.

A rough test of this model could be conducted by means of an examination of recent hospital experience with changing reimbursement. Whenever the supply of funds to the hospital is increased, the model would predict that the institution would reduce at least its room charges in order to increase the quantity of services sold. After the enactment of Medicare, however, when hospitals were reimbursed at a cost plus two percent basis, and thus received windfall revenue, the opposite happened, with prices increasing at a much faster pace than before.[32]

A phenomenon such as the above would be consistent with a third behavioral hypothesis, namely, that of quality-quantity maximization.[33] It implies that during any period incentives exist for the hospital to accumulate a certain surplus which can be used in the next period for quality improvements and additions to plant and services. A variant of this hypothesis will be adopted in the theoretical model used in this thesis.

A generalized version of the quantity or quantity-quality maximization hypothesis is that of utility maximization.[34] An objective function for the hospital is derived from the utility functions of hospital administrators and staff physicians. Utility is ultimately seen by the proponents

equipment

Since the a

on the ran

existing st

hypothesis

C

cash flow

maximize

other than

is similar

An excess

additional

expanded.

(

hospital

they impl

somewha

fulness a

an impor

certain d

the hosp

does not

proponents of this theory as a function of the extensiveness of modern equipment and the professional prestige of physicians on the staff. Since the ability of the hospital to attract high caliber doctors depends on the range of its capital equipment as well as the quality of the existing staff, utility maximization reduces to a capital maximization hypothesis.

One final theory of hospital behavior is based on a version of cash flow maximization.[35] According to this hypothesis the hospital maximizes the difference between revenue and operating expenses other than depreciation costs. The basic premise behind this theory is similar to that of the quantity-quality maximization hypothesis: An excess of funds over costs is the objective of the hospital so that additional facilities may be added and the scope or quality of services expanded.

One distinguishing feature of all the existing theories of hospital behavior (except for the recovery-of-costs hypothesis) is that they imply cost minimizing behavior on the part of hospitals. This is somewhat surprising considering the widespread impression of wastefulness and inefficiency in the hospital industry. There is, however, an important qualification: What these theories predict is that after certain <u>desired</u> levels of quantity and quality of output have been set, the hospital will attempt to meet these goals at minimum cost. This does not assure, however, that these targets are set at levels where marginal

is that if t

will not bu

is availab

sense wou

benefit fro

suggested

and, there

economic

T

a synthesi

profit hos

ject to the

the short-

quantity m

ments ma

long-run.

even in a

Its resis

inherent

person's

in agree

marginal revenue equals marginal cost. What these theories imply is that if the medical staff asks for a $20,000 x-ray unit, the hospital will not buy one for $25,000. If, however, a somewhat inferior unit is available at $15,000, actual cost minimization in the economic sense would imply the purchase of that unit as long as the marginal benefit from the $20,000 machine was less than $5,000. None of the suggested theories[36] imply such behavior on the part of the hospital and, therefore, they do not preclude inefficient behavior in the economic sense.

The behavioral assumption adopted in this thesis is basically a synthesis of the theories outlined above. It is assumed that non-profit hospitals attempt to maximize the quality of their services subject to the constraint of meeting community demand up to capacity in the short-run. In the long-run, this theory approaches the quality-quantity maximization hypothesis, since many of the quality improvements may also serve to increase the quantity of services sold in the long-run.

The notion of quality has always been a source of problems even in areas much more developed than that of hospital economics. Its resistance to quantification and often even conceptualization is inherent in the subjective nature of the concept. Quite simply, one person's evaluation of an object or of an outcome is not necessarily in agreement with that of another individual. The problems of

evaluating c
economics.
has several
the perform
of hospital
ization of c
is not cert
lower qual

F

measure o
of the phys
The one in
time an in
changes in
This does
surate inc
can be ac
equipmen
hospital p
provemen
should no
the nonpr
as percei

evaluating different levels of quality are even more acute in hospital economics. Whereas in the case of most goods and services quality has several tangible aspects, such as the nature of materials used or the performance and durability of the particular product, the quality of hospital care also includes many intangibles such as the personalization of care or the psychic comfort of patients. For example, it is not certain whether rapid but painful treatment is of higher or lower quality than a slower but less painful process.

Fortunately, this thesis does not require a quantifiable measure of quality. After all, this question is best left in the hands of the physicians, primarily, and perhaps the hospital administrators. The one important fact that must be established is that at any point in time an increase in the level of quality of services without any other changes in the pattern of production requires an increase in costs. This does not mean that any cost increase is associated with a commensurate increase in quality, but rather that no quality improvements can be achieved free of cost. Improvements or expansion of capital equipment, higher employee-patient ratios, improved skill-mix of hospital personnel or higher calibre medical staff are all quality improvements which can only be accomplished at increased cost. We should now try to justify the assumption that the primary objective of the nonprofit hospital is to increase the quality of care, either real or as perceived by the various decision makers inside the institution.

T

The board

for the ins

the long -r

struction a

limited pa

tion. Shor

by the hos

makes mos

hospital, h

staff, on t

and discha

F

literature

and vested

long run

the board

since the

jective fu

tion or pr

quality o

the size

There are three sources of authority in the hospital hierarchy. The board of trustees is formally at the top, with legal responsibility for the institution. The trustees make major decisions dealing with the long-run goals and functioning of the hospital such as new construction and major service additions, but the trustees have only limited participation in the actual day-to-day operation of the institution. Short-run decisions concerning input and output levels are made by the hospital administrator and the physician staff. The former makes most of the decisions concerning the every-day operation of the hospital, hires the various inputs, and sets prices. The physician staff, on the other hand, has almost complete authority over admissions and discharges and the way in which the various inputs are used.

For a variety of reasons which are well documented in the literature,[37] trustees, administrators and physician staff have a strong and vested interest in continuous quality improvements as well as in long run quantity increases in terms of the size of the hospital. For the board of trustees satisfaction does not lie with pecuniary returns since they are not usually remunerated for their services. Their objective function, therefore, includes as principal elements the reputation or prestige of the hospital in the community and the quantity and quality of care provided.[38] Such prestige is, in turn, dependent on the size of the hospital, the number and quality of services offered, and the siz

is largely c

a wide rang

Th

notion of qu

his perform

fessional s

the staff, t

and the rep

equipped to

T

prestige o

and the ex

that physi

mand high

cated equ

wider cho

existence

working c

vity.

the quali

however,

and the size and professional caliber of the physician staff, which is largely determined by the existence of sophisticated equipment, a wide range of services, and the extent of teaching programs.

The objectives of the administrator also center around the notion of quality. Since he is not formally required to show a "profit," his performance must be judged by other criteria, such as the professional status of the physicians and specialists he helps attract to the staff, the prestige of the institution, the services it provides, and the reputation it enjoys concerning the quality of care it is equipped to offer.

The medical staff, finally, has an obvious interest in the prestige of the institution, the extensiveness and quality of equipment and the existence of a wide scope of services. It is probably true that physicians affiliated with the more prestigious institutions command higher fees for their services. The existence of highly sophisticated equipment and facilities, moreover, affords the physician a wider choice as to the proper method of treatment. Finally, the existence of highly skilled nursing and paramedical personnel improves working conditions for the physician staff and increases their productivity.

Considerations such as the above seem to argue in favor of the quality maximization hypothesis. The hospital's economic behavior, however, is also influenced by the desired amount of output. In the short run

has to mee

objectives

resource

physician

to appoint

the bylaws

strator to

the deliver

an employ

invited by

in the form

outside ad

on the con

almost su

prescribe

legal and

The phys

cross ad

internal

although

charged

those of

short run output acts as a constraint in the sense that the hospital has to meet community demand for services regardless of its other objectives. This is because short run output determination and resource utilization are almost completely in the hands of the physician staff. Although the board of trustees has the authority to appoint physicians and to delineate the extent of their practice, the bylaws of most hospitals require the trustees and the administrator to abide by medical staff recommendations with respect to the delivery of patient care. In most situations the physician is not an employee of the hospital; he is rather an independent professional, invited by the institution and granted practicing privileges. Although in the formal organizational chart of the hospital the physician is outside administrative lines of responsibility and without authority on the conduct of hospital itself, his authority over his patients is almost supreme. Only he can admit patients, make diagnoses, and prescribe therapy. The hospital, therefore, although a separate legal and producing entity, is particularly dependent on the physician. The physician's absolute control over his patients allows him to cross administrative lines of authority. This quite often creates internal problems, as, for example, for hospital employees, who, although formally responsible to the hospital manager, are also charged with carrying out doctor's orders, which may conflict with those of the administrator.[39] The ultimate consequence is that with a fixed hos

macy over

sions and

possible w

has been n

W

quality in

objective f

the quality

and physic

an overall

maximiza

the econo

I

that the h

costs at

ments or

amount o

profit st

objective

Baumol'

instituti

group s

a fixed hospital budget during any period in time, physician supremacy over the quantitative aspects of hospital output such as admissions and length of stay constrains the level of quality attainable to that possible with the existing budget and after all demand for services has been met.

While the quantity of output may act as a constraint on quality in the short run, it also becomes an element of the hospital's objective function when long run decisions are concerned. Part of the quality improvements as perceived by trustees, administrators, and physicians involve additions to plant, extension of services and an overall increase in size and capacity. In that sense long run output maximization can also be considered an important objective influencing the economic behavior of hospitals.

Both the short run and long run implications of the model are that the hospital must show a certain surplus of revenues over operating costs at the end of each period in order to undertake quality improvements or increases in capacity for the future. In other words, a certain amount of "profit" in the short run is not only consistent with the nonprofit status of hospitals but also necessary considering their long term objectives. This view of hospital economic behavior is supported by Baumol's general theory of behavior as it applies to all nonprofit institutions.[40] According to Baumol, nonprofit organizations as a group share at least two characteristics: (1) they earn no pecuniary return on i

purpose.

nonprofit c

them cons

pose behi

pecuniary

scope, an

the availa

goals con

these pro

unavaila

improve

or priva

by show

small (1

Patient

remaini

investm

return on invested capital and (2) they claim to fulfill some social purpose. The significant point is that the objectives of the typical nonprofit organizations are by their very nature designed to keep them constantly in need of funds since the quality and the social purpose behind their product become ends in themselves regardless of pecuniary considerations. Their goals of constantly increasing the scope, and quality of their products or services, therefore, require the availability of additional funds at the end of each period. These goals constitute, as Baumol puts it:

> bottomless receptacles into which limitless funds can be poured. Any well functioning nonprofit organization will always have a group of projects which it cannot afford to undertake and for whose realization it looks hopefully to the future.[41]

The question now arises, where will the funds necessary for these projects come from? Until recently, and largely because of the unavailability of detailed revenue data, it was thought that most hospital improvement and expansion was financed by philanthropy, either public or private. An excellent recent study,[42] however, dispells this notion by showing that even as early as 1966 donations represented a very small (1.8 percent) portion of total nonprofit hospital revenues. Patient revenue, on the contrary, represented 93.2 percent, with the remaining 5.0 percent coming from other sources such as earnings on investment, cafeteria sales, and rental of nonpatient facilities.[43] The conclusion

internally

other cont

I

maximize

cation is

revenues

reimburs

directly a

therefore

economic

through t

(1)

where

conclusion is that nonprofit hospitals generate enough revenue internally to more than cover total expenses without depending on other contributions for expansion.[44]

If now, as was asserted previously, hospitals attempt to maximize quality subject to meeting existing demand, the implication is that they would attempt to maximize the excess of total revenues over total costs during each period. The question of reimbursement becomes very important at this point because it directly affects the revenue side of hospital economics. We will, therefore, construct a simple algebraic model to demonstrate the economic incentives afforded by various reimbursement mechanisms through their effect on hospital revenue.

The total revenue equation for a hypothetical hospital is:

(1) $$R_t = k\, c_t\, Y_1 + p_t\, Y_2 + M$$

where

R_t = total hospital revenue in period t

Y_1 = total number of patients covered under a cost reimbursement scheme (Blue Cross, Medicare, Medicaid)

Y_2 = total number of patients covered under charge reimbursement (private insurance and self-pay patients)

M = total nonpatient revenue

c_t = average hospital cost per case

p_t = average charge per case

T

varies acc

paid, then

will be 1.

I

revenue e

(2)

where $y_{1'}$

revenue.

expressin

(3)

(4)

we can

(3.1)

or alte

(3.2)

where

*Sinc
now r

The parameter k is a cost reimbursement parameter which varies according to the formula used. For example, if full costs are paid, then k = 1. If a plus two percent factor is used, the value of k will be 1.02.

By dividing both sides of (1) by $(Y_1 + Y_2)$ we obtain the average revenue equation (2):*

(2) $$r = k\, c_t\, y_1 + p_t\, y_2$$

where y_1, y_2 are proportions of total cases and r is average patient revenue.

Finally, by dividing both sides by c_t we obtain equation (3) expressing average revenue as a percent of average costs:

(3) $$\frac{r}{c_t} = k y_1 + \frac{p_t}{c_t}\, y_2$$

Utilizing the identity:

(4) $$y_1 + y_2 \equiv 1$$

we can rewrite (3) as:

(3.1) $$\pi = k + y_2 \left(\frac{p_t}{c_t} - k \right)$$

or alternatively in terms of y_1,

(3.2) $$\pi = \frac{p_t}{c_t} + y_1 \left(k - \frac{p_t}{c_t} \right)$$

where $\pi = r/c_t$, the profit ratio.

---

*Since M is largely exagenous we have dropped it from (2) where now represents average patient revenue.

U

(3.1) that

a surplus

the hospi

higher th

plus is to

raising p

will have

still exis

Again the

of the av

for the

for on a

to meet

largely

very in

number

to keep

this ac

---

*This

Under full cost reimbursement where k = 1, we see from (3.1) that π will be greater than one, i.e., the hospital will have a surplus from patient care, as long as $\frac{p_t}{c_t} > 1$. In other words, the hospital must charge patients with indemnity coverage a price higher than average costs. In fact, the only way to maximize the surplus is to maximize $\frac{p_t}{c_t}$, either by lowering average costs or by raising prices to charge paying patients.

Under cost plus reimbursement, where $k > 1$, the hospital will have a surplus as long as $\frac{p_t}{c_t} > k$. If $\frac{p_t}{c_t} < k$ a surplus can still exist if the product $y_2 \left(\frac{p_t}{c_t} - k\right)$ is smaller than the plus factor.* Again the hospital can maximize the surplus by maximizing the excess of the average charge over average cost.

An interesting result is that in both cases incentives exist for the hospital to increase the proportion of its patients who are paid for on a charge basis. Since we have assumed that hospitals will try to meet any demand for their services and since admissions are largely in the hands of physicians, such incentives are probably not very important. If, however, Y_1 and Y_2 are taken to represent the number of patient days instead of cases we can see that incentives to keep charge-paying patients longer do exist.[45] To what extent this actually happens is not clear, however.

---

*This is because $k = (1 + s)$ where s is the percentage plus factor.

T

incentives

(3.1) or (3

except ind

surplus e

paying pa

pital care

To put it

of surplu

if any, fo

imburse

charge p

cost rei

Let us a

paramet

where c

The act

institut

The most important implication of the model is that efficiency incentives for the reduction of costs are nonexistent. As we see from (3.1) or (3.2) average hospital cost does not affect the profit ratio except indirectly by its influence on $\frac{p_t}{c_t}$. A hospital can increase its surplus either by lowering average costs or by raising prices to charge paying patients. Because of the virtual inelasticity of demand for hospital care, the second course of action is probably considerably easier. To put it differently, the hospital can determine the desired amount of surplus for a period and achieve it partly through the plus factor, if any, for the portion of its patient load covered under cost plus reimbursement and partly by an excess of prices over costs for its charge paying patients.

Let us now examine the revenue implications of replacing cost reimbursement by a particular type of incentive reimbursement. Let us assume that a plan is used which defines the reimbursement parameter k for the i^{th} hospital in (3.1) and (3.2) to be:

$$k_i = \frac{\bar{c}_t}{c_{it}}$$

where $\bar{c}_t$ = the average cost per unit of care for a group of hospitals. The actual method of grouping hospitals is not important at this point.

In other words, the parameter k is no longer determined by institutional agreements, but it is directly related to the hospital's cost p

group.

(3.3)

(3.4)

It can b

per uni

over co

pansion

high cos

than full

imburse

prices fo

it is doub

the reven

payments

enough, a

ment. S

measure

closely a

reviewed

case or

cost performance relative to that of other institutions in the same group. Equations (3.1) and (3.2), therefore, become:

$$\pi = \frac{\bar{c}_t}{c_{it}} + y_2\left(\frac{p_t}{c_t} - \frac{\bar{c}_t}{c_{it}}\right) \tag{3.3}$$

$$\pi = \frac{p_t}{c_t} + y_1\left(\frac{\bar{c}_t}{c_{it}} - \frac{p_t}{c_t}\right) \tag{3.4}$$

It can be seen from (3.3) or (3.4) that by keeping the average cost per unit of care low, a hospital can increase the excess of revenue over costs and use the surplus for quality improvements or expansion of its scope of services. Moreover, the disincentives to high costs are obvious. High average costs will result in less than full cost payments for the patients covered under incentive reimbursement. The hospital could, of course, attempt to raise prices for charge paying patients, but despite the inelasticity of demand it is doubtful that it could increase revenue enough to compensate for the revenue loss resulting from the low incentive reimbursement payments.

Although the incentives in keeping costs down are clear enough, a certain danger arises out of such a system of reimbursement. Since hospitals are rated and reimbursed according to some measure of costs, this measure must reflect relative efficiency as closely as possible. The various incentive reimbursement plans reviewed in this chapter, suggest the use of the average cost per case or per patient day as measures of hospital cost performance.

We show,
appropria
determine
a given se
an accura
given hos
the probl
new meas
differenc
of patien

We show, however, in Chapter IV that these measures of cost are inappropriate measures of hospital efficiency. Since efficiency is determined by the actual amount of output a hospital produces with a given set of inputs, the correct measurement of efficiency requires an accurate measure of the amount of patient care produced by a given hospital. For this reason, in the next chapter we will explore the problems caused by hospital output heterogeneity and suggest a new measure which adjusts for certain quantitative and qualitative differences and measures hospital output in terms of the actual amount of patient care produced.

a constar

hospital

is that t

definitio

sent us

differen

across a

adminis

put appl

pital ou

will der

differer

weights

hospita

CHAPTER III

THE DEFINITION AND MEASUREMENT OF HOSPITAL OUTPUT

The Problem of Defining Hospital Output

The definition and measurement of hospital output has been a constant source of conceptual problems in many studies dealing with hospital production and costs. First, it is not obvious exactly what it is that the hospital produces. Second, even if we decide on a certain definition, qualitative differences in the output of each hospital present us with serious conceptual and measurement difficulties. The differences are due to the intrinsic heterogeneity of hospital care across and within institutions and to the fact that hospital care is not administered instaneously but, rather, over time, and the rate of input application varies with time. In this chapter we will define hospital output as the weighted amount of patient care provided and we will derive a set of weights which account for certain qualitative differences in the hospital output of patient care. We will use these weights in order to construct 1) a costliness index as a measure of hospital costs and 2) a scalar measure of hospital output. These two concept

perform

prevent

wants t

servati

duct of

For tw

studies

is, so f

rates a

conclus

mortali

they ha

than re

ductio

perfor

by the

levels

of urb

concepts will be used in subsequent chapters to analyze hospital cost performance and efficiency for the purposes of incentive reimbursement.

Definition of Hospital Output

When a patient enters a hospital he generally seeks to either prevent or cure some ailment which threatens his health. What he wants to obtain, and what the hospital attempts to provide, is the preservation or restoration of his health. In this sense, the ultimate product of the hospital is improved health for the patients that it services. For two reasons this notion of hospital output is inappropriate for studies concerned with hospital productivity and costs. First, health is, so far, a basically unmeasurable concept.[46] The use of mortality rates as an index of health for example, would lead to the rather dubious conclusion that one of the reasons why many other countries show mortality rates considerably lower than those of the U.S. is because they have more and/or better hospitals. Second, hospital care, rather than representing health, is actually only one of the inputs in the production of health.[47] Clearly, the line of causality between the work performed by hospital inputs and the production of health is obscured by the presence of many other variables that influence population health levels, such as environmental and demographic variables, the degree of urbanization, work habits, and other non-hospital medical factors.

We r

secc

rese

dom

dire

they

do n

a fi

only

pati

We

pati

app

son

pro

pita

Fi

fac

se

We must, therefore, look for another concept, perhaps logically secondary to health, in order to define hospital output.

Community hospitals produce varying quantities of education, research, community services, outpatient care, and, their predominant output, inpatient care. These are the activities that result directly from the productive efforts of hospital inputs, and, as such, they logically constitute hospital output. However, these activities do not all take place in every institution. In the interest of achieving a first approximation of hospital output comparability, we will consider only the forms of output common to all hospitals. These are inpatient and outpatient care which, together, we will call patient care. We will therefore define hospital output as the number of units of patient care that the hospital provides for a period of time. This approach is almost exclusively used in the literature, although some authors do not consider outpatient care. We now come to the problem of defining and measuring patient care so as to measure hospital output in a meaningful way.

Measuring the Amount of Patient Care

Patient care is far from homogeneous among hospitals. First, institutions treat different mixes of cases according to their facilities and staff and the population composition of the areas they serve. Second, even similar cases often require different lengths of stay i

of patie

quently

pital is

a certai

represe

hospital

days sp

pital pr

measur

for exa

on cost

the othe

cost pe

tions a

adjusts

hospital

among

effect o

proach

of stay in different institutions, therefore representing different amounts of patient care. There are two basic measures of hospital output frequently used in the literature: the case, and the patient day. A hospital is seen as producing care for a variety of cases, or as producing a certain number of patient days of care. Cases can be conveniently represented by the number of admissions to, or discharges from, a hospital during a given time. Patient days are the total number of days spent by all patients during that time.[48] Unfortunately, any hospital production or cost study that used the number of cases as a measure of output makes the implicit assumption that a tonsillectomy, for example, uses the same amount of inputs or has the same impact on costs as a heart transplant. Measuring output by patient days, on the other hand, requires the additional assumption that input use or cost per day for a given case is constant. Since both these assumptions are difficult to justify, we must measure output in a way which adjusts for casemix and length of stay differences.

The Casemix Adjustment

In order to use the case or the patient day as a measure of hospital output, one must assume either that the casemix distribution among hospitals is identical or that casemix differences have no effect on hospital costs and optimal input combinations. Both approaches, although frequently used, simply assume the problem awa

data

for

the

four

thir

used

num

of o

othe

or f

sur

wei

fra

cos

mu

gre

ove

out

mu

away. In defense of existing research, it must be said that casemix data are rarely available[49] except in special situations and then only for a small number of hospitals. In one of the few attempts to handle the problem, M. Feldstein, using a sample of British hospitals, found that casemix differences alone account for approximately one-third of average cost variation.[50]

There are three ways in which casemix information can be used to arrive at a correct measure of output: 1) by including the number of patient care units in each category of care in some form of output vector; 2) by assuming that casemix is correlated with some other hospital characteristic such as size, location, teaching status or facilities and services; and 3) by creating a weighted output measure where case-types with high input requirements receive larger weights.

The first method, although theoretically justifiable, is fraught with econometric difficulties when used to estimate hospital cost or production functions. These difficulties arise from (a) multicollinearity among the explanatory variables, and (b) lost degrees of freedom due to the many independent variables.[51] Moreover, in the estimation of production functions, a scalar measure of output is usually required since the theory and the estimation of multi-product production functions is not yet fully developed.

wi

th

co

in

ta

ge

me

Al

ho

ch

di

co

Th

pi

pi

ho

pr

va

du

ou

in

A number of writers have assumed that casemix is correlated with other hospital characteristics[52] and have obtained estimates of the influence of casemix on costs by including these variables in their cost functions. R. Berry attempted to solve the problem by estimating cost functions for different groups of hospitals, each group containing identical facilities and services.[53] In this way one can make generalizations about each such grouping, but because no weighting mechanism is used, there is no way to compare groups directly. Also, in order to obtain a sufficient number of observations in each homogeneous group one must use a very large sample,[54] thereby including hospitals with very different accounting procedures and facing different input and output markets.

The assumption that hospital facilities and services are correlated with case-mix was also made by Saarthof and Kurtz.[55] Their measure includes seven services which are part of every hospital's operation, such as lab tests, x-rays, etc. They define hospital output to be the amount of each of these seven services the hospital provides. In order to integrate these services into a single product measure they derive a set of weights based on crude observations on the amounts of labor and materials going into the production of one unit of each service. This method, therefore, weighs output by the mix of intermediate services (or inputs) that are used in the actual production of treatments for the various cases. Although it adju

althou

prefer

find a

estim

the pr

shoul

produ

hospi

porti

cost

diffic

Saar

the c

inter

trea

stan

mea

sod

vice

wei

of c

it adjusts for differences in the mix of services and not casemix, and although the weights are chosen arbitrarily, this output measure is preferable to the case or the patient day.

A similar method is employed by Cohen.[56] He attempts to find a measure of output by weighing each intermediate service by its estimated average cost in dollars. Theoretically, this is similar to the previous approach since the cost of producing a unit of service should depend on the inputs and the production functions used in the production of these intermediate products. The many intangibles of hospital operations, however, together with the heterogeneity of reporting procedures make the econometric estimation of such average cost functions difficult and often inaccurate. In view of this it is difficult to say whether the Cohen approach is superior to that of Saarthoff and Kurtz.

Although facilities and services may be a good indication of the quality of care offered by different institutions, what we are truly interested in is to account for differences in the types of patient treated in the various hospitals while keeping the quality of care constant. In other words, the patient care output of a hospital should be measured in terms of the hospital's final product, expressed as episodes of illness treated rather than in terms of the intermediate services which produce this output. This can be done by the use of a weighted output index. In general terms, let X_{ji} represent the number of cases of type j treated by hospital i during a period of time. We th

wh

si

be

hi

Co

ou

in

ch

to

pe

ar

co

pr

kn

w

th

a

t

then want to construct a scalar measure of output:

$$Y_i = f(w_j X_{ji})$$

where w_j is the weight assigned to each case type. The weights for similar casetypes should be the same for all hospitals, and they should be derived in such a way as to assign greater values to case types with higher input requirements. One simple specific form of this measure is:

$$Y_i = \sum_{j=1}^{k} w_j X_{ji}$$

Conceptually, rather than being a strictly unidimensional measure of output, Y_i is the "mapping" of an n-dimensional space of output vectors into the one-dimensional space of a scalar. The problem now is with choosing an appropriate set of weights for the various X_{ji}.

When products are sold in competitive markets it is common to aggregate them by using prices as weights. In the perfectly competitive model prices depend in the long run on costs which, in turn, are derived from the production function and the input prices. Since the competitive assumptions are not met in hospital production, the use of prices as weights is unsatisfactory, especially since hospitals are known to apply differential pricing policies for different types of care with little regard to the average cost in each case category. We must therefore go directly to the cost side for our weights. The use of average costs as weights is a rough approximation, and it is based on the assumption that society values cases of different types in pr

in

di

th

pr

w

w

le

in

d

t

u

(

proportion to the average costs of producing treatments for these cases in the "average" hospital, i.e., that the average social costs of the different case types are proportional to the hospital average costs for these case-types.[57] On the basis of this assumption M. Feldstein proposes a measure of the weighted output ("work") of a hospital as:[58]

$$Y_i = \sum_{j=1}^{k} c_j X_{ji}$$

where c_j is the average cost of treating a case of type j. This study will use an expanded version of this method which will also adjust for length of stay differences in the treatment of similar cases by different institutions.

The Length of Stay Adjustment

Standard economic theory implicitly treats the firm's production process as instantaneous. In other words, studies of the technological relationship that transforms inputs into output do not usually include the time required for the production of one unit of output. In the case of the hospital, however, time will be shown to be a very important element which should influence the choice of the measure of output to be used.

Hospital output was previously defined as the amount of patient care provided by the hospital. One distinguishing feature of hospital output is that treatment for each case is produced over a pe

ar

an

am

co

Un

hos

Ho

the

537

ICI

the

low

len

1 s

dia

it r

to u

stay

curg

oppo

such

in th

period of time which is known as the "length of stay." The problem arises from the fact that length of stay varies both among case types and among hospitals for similar cases. If length of stay variability among hospitals were only due to differences in casemix, the problem could be solved with the adjustment shown in the previous section.[59] Unfortunately, there are substantial differences in length of stay among hospitals for identical cases. The Commission on Professional and Hospital Activities has produced a lengthy statistical study[60] reporting the mean, variance, and percentile distribution of average stays in 537 short-term general hospitals for each disease in the four-digit ICDA[61] classification system. In almost all diseases and operations the large variances as well as the substantial numbers of cases in the low and high percentiles indicate significant inter-hospital variations in length of stay for similar cases. For the sake of illustration, Table 1 shows the length of stay percentile distribution for five random diagnoses. Since the diagnosis breakdown is very detailed, and since it reasonable to expect most physicians in a given medical specialty to use similar production techniques, we would expect the length of stay distribution to be highly clustered around the mean. Even a cursory examination of the data, however, indicates precisely the opposite. In Table 1 we see that, even in a very specific disease such as malignancy of rectum, fifty percent of the patients stayed in the hospital for seven days or less while another forty stayed between

observ

examin

accura

TABLE

Diagnos

Influenz

Bronchi
chron
unspe

Acute c
occlu

Rheuma
fever

Maligna
of rec

medical

ences i

be appl

medica

between eight and 24 days and 9 percent between 25 and 55 days. We observe equally large length of stay variances in most of the diagnoses examined, evidence which leads us to believe that casemix is not an accurate reflection of length of stay.

TABLE 1. --Length of Stay Distribution for Selected Diagnoses

Diagnosis	ICDA Classification Number	Average Length of Stay (AL)	Variance	Percentiles			
				5th	50th	90th	99th
Influenza	480.0-483.0	4.9	12	1	4	9	18
Bronchitis, chronic & unspecified	501.0-502.9	6.4	27	2	5	12	26
Acute coronary occlusion	420.1	21.2	86	6	21	32	49
Rheumatic fever	400.0-402.1	12.9	120	2	10	26	57
Malignancy of rectum	154.0	10.8	142	1	7	24	55

Source: CPHA Length of Stay in PAS Hospitals. (Ann Arbor 1969), various pages.

The reasons behind such differences in length of stay can be medical, technological, and institutional.[62] There may be wide differences in recovery rates or in the ways in which certain treatments can be applied on different individuals. Factors such as age, previous medical history, income, a patient's family situation, and even certain d

in

in

le

th

d

o

a

th

o

lo

h

o

id

d

lo

"

h

is

a

th

ea

demographic factors can influence his or her length of stay. Second, in addition to the possibility of medical incompetence, there is the important element of difference in physician-view as to the proper length of stay.[63] Some physicians may require a longer period before they are sure that the treatment has been successful. Thirdly, if demand for beds is very high, patients may be discharged earlier in order to make room for more urgent cases. Alternatively, there is also reported a tendency for hospitals with empty beds to pressure their staffs to get them more business.[64] Finally, the current system of financing hospital care may induce some hospitals to keep patients longer than medical considerations would dictate.[65]

The fact that length of stay for similar cases differs among hospitals is one more reason for the inappropriateness of the number of cases as the measure of hospital output. Let us imagine two identical hospitals with identical numbers and types of cases but different average lengths of stay. Clearly, the one that keeps patients longer has produced more output in the sense that it has done more "work". One might argue that this is fallacious and that, if the two hospitals treat identical casemixes, the one with the longest stays is more inefficient in producing the same quantity of output. There are two reasons why this cannot be entirely true. First, assuming that hospitals use similar production functions for the treatment of each case-type,[66] this would mean that the marginal products of <u>all</u>

unp

un

so

du

the

no

ca

pr

so

its

mi

au

pr

qu

the

ma

po

su

pa

ho

mi

pro

inputs in the hospital with the longer stays fall to zero after a certain time during the course of treatment.[67] Although this may be true for some inputs, it is doubtful that it holds for all the factors of production. If a patient's medical history or age require him to stay in the hospital for 15 days while the average patient in the same diagnostic category only stays 11 days, it is difficult to say that no patient care was produced after the eleventh day. Second, the notion of hospital care is not a purely quantitative concept but, rather, contains some qualitative elements as well. If, for example, length of stay is itself an aspect of the quality of patient care this must be borne in mind in defining the hospital output.[68] As we saw in Chapter II, some authors do indeed contend that hospitals employ their inputs in the production of an output with two dimensions, namely, quantity and quality.[69] Because of these two considerations we will assume that the production function is an increasing function of time and that the marginal product of hospital inputs during the last day of care is positive. We will, therefore, conclude that a case with a length of stay of eight days in one hospital represents more output in terms of patient care than an identical case that stays for five days in another hospital.

The above would seem to indicate that if we adjust for case-mix differences among hospitals we would then solve the length of problem by letting the total number of patient days represent hospital

o
th
th
s
m
th
bl
th
m
la
to
of
af
as
Th
ho
si

wh
co

B w
es

output. Unfortunately, although it incorporates the time of production, the number of patient days is the product of two variables, namely, the average length of stay and the number of cases treated during, say, a year. Because of this, an observed one thousand patient days may represent one hundred cases staying for ten days each, or one thousand cases staying for one day. This would not present a problem if the average product of hospital inputs remained constant over the patient's stay. It is well known, however, that inputs are much more intensively used during the first days of care than during the last days of convalescence. Most of the x-ray procedures, laboratory exams, use of surgical facilities, and the most intensive use of the hospital's labor inputs takes place within the first few days after admission. After that time the rate of input application falls as medical services are increasingly replaced by "hotel"[70] services. The use of the patient day, therefore, would assign output values to hospitals following early discharge policies which relatively understate the actual amount of patient care they have provided.

The situation can be shown more explicitly in Table 2 which uses a hypothetical example of four hospitals with various combinations of cases and average lengths of stay.

If the case is chosen as the unit of output, hospitals A and B will be assigned identical output values. This would underestimate the amount of output in hospital B, which produced twice as many patient days. Conversely, if the patient day is chosen as the unit of output, hospital C will have 16 percent more patient days than hospital D. Again, this would underestimate the performance of hospital C which produced treatments for twice as many cases. Conversely, the number of patient days would overestimate the output of hospital D which produced treatment for only half as many cases.

TABLE 2. --Amounts of Output Measured by Cases, Patient Days, and Units of "Adjusted Patient Care"

Hospital	Cases	A/B and C/D	Average Length of Stay (AL)	Patient Days	A/B and C/C	Log AL	Adjusted Patient Care	A/B and C/D
A	2,000		6	12,000		1.792	3,584	
		1			2.00			1.39
B	2,000		12	24,000		2.485	4,970	
C	2,000		7	14,000		1.946	3,892	
		2			1.16			1.56
D	1,000		12	12,000		2.485	2,485	

Our solution to this problem is simple. Define a variable $Y_{ji}*$ representing the amount of "adjusted" patient care in case-type j as:

$$Y_{ji}* = Y_{ji} \log l_{ji}$$

where

Y_{ji} = number of cases of type j in hospital i

l_{ji} = average length of stay for case-type j in hospital i

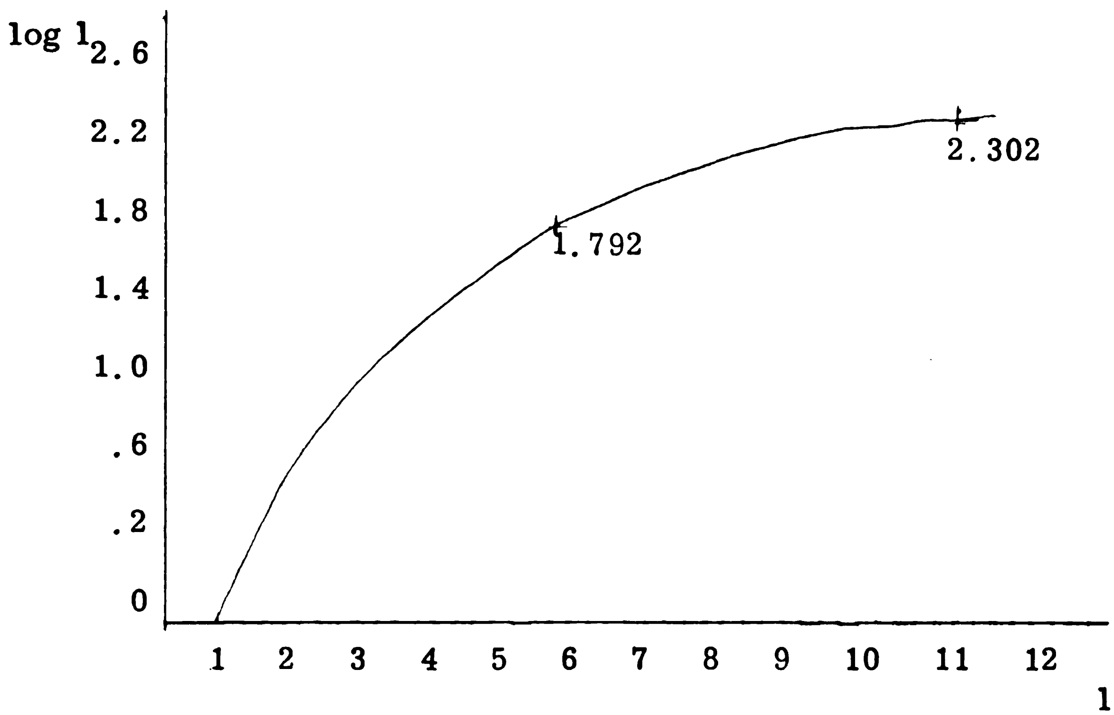


FIGURE 1. --The Logarithmic Transformation of Cases into Units of "Adjusted Patient Care".

The logarithmic curve shown in Figure 1 assigns higher values of output to hospitals with a higher length of stay. The rate of increase, however, is declining with time, reflecting the reduced rate of resource application. Although the logarithmic transformation used here is arbitrary, it has two desirable properties. First, it is monotonically increasing at a declining rate, and, therefore, it fits our theoretical expectations of a positive but declining marginal product of an additional day of care. Second, the logarithmic function is considerably easier to compute than other nonlinear functions with the above desirable properties.

The results of the transformation can be seen in Table 2. Hospital B is assigned a higher value of output because of its longer

AL but its output is only 39 percent higher than that of A rather than 100 percent as shown by the use of the patient day. Similarly, hospital C is assigned a value of output which is 56 percent higher than that of D as opposed to only 16 percent as shown by the use of the patient day. This 56 percent figure is more reasonable since hospital C treats twice as many patients and thus has a much higher percentage of "expensive" days, i.e., the first few days of treatment.

Our solution to the problem of hospital output heterogeneity therefore, involves two types of adjustment. First we adjust for case-mix differences by estimating a set of average cost weights (c_j) for the different case-types. Second, we adjust for length of stay differences among hospitals by multiplying the number of cases in each case-type by the logarithm of the average length of stay for that case-type in each hospital (l_{ji}). Our final scalar measure of output (Y_i^*) for hospital i is thus defined as:

$$Y_i^* = \sum_{j=1}^{k} c_j X_{ji} \log l_{ji}$$

where Y_i^* represents the total units of adjusted patient care produced by hospital i.

Our discussion on the definition and measurement of hospital output will serve as background for the development of the costliness index in the next chapter. The casemix and length of stay adjustments will be used to adjust cost per case. The measure of output derived here will also be used in Chapter VI in the estimation of a productivity index from a Cobb-Douglas production function.

CHAPTER IV

THE COSTLINESS INDEX AS A MEASURE OF HOSPITAL COSTS WHICH REFLECTS EFFICIENCY DIFFERENCES

Hospital Cost Measurement

As stated earlier, the purpose of incentive reimbursement is to provide hospitals with economic incentives for efficient operation by penalizing inefficiency and rewarding efficient use of resources. If some estimate of cost, therefore, is used as the standard for reimbursement, it follows that the cost concept used must bear a close relationship to productive efficiency. In this chapter it will be shown that average cost per case or per patient day, although often used, is an inappropriate reimbursement tool, and another measure of hospital cost performance, which is more closely related with efficiency, will be suggested.

There are two different measures of the cost of hospital care vhich are used most often: the average daily service charge ADSC) used by the Bureau of Labor and Statistics, and the average ost per patient day (ACPD) calculated by the American Hospital ssociation. The average daily service charge, which is part of the medical care component of the consumer price index, is an attempt to measure the price at which hospitals sell a day of in-patient care. It includes only the charge for room accommodations,[71] food service, routine nursing care, and minor medical and surgical supplies. The ACPD on the other hand, is a much more inclusive measure of the cost of hospital care since it also reflects all special services, drugs, and tests: it is calculated by dividing total hospital costs, excluding only capital investments, by the number of days of patient care.

There is no general agreement as to which measure reflects hospital per day costs in a more satisfactory way. The ADSC is criticized primarily for not incorporating many of the ancillary costs that a typical patient incurs as a part of his hospital stay, especially since specialized services represent a large and growing fraction of total costs. Furthermore, the daily service charge is sensitive to arbitrary changes in the allocation of total costs between room rates and other charges. Some writers, in fact, suggest that a significant part of recent increases in hospital costs as measured by the ADSC may reflect a shift away from a pricing policy which previously set the room rate below cost while other services were priced to yield a profit.[72] The ACPD, on the other hand, is criticized because it does not allocate costs between inpatient and outpatient care, making the implicit assumption that all

As

pro

cas

pro

enc

tre

ref

car

inc

ser

eas

as

lea

soc

roo

and

for

few

mea

hos

all hospital costs are incurred in the provision of inpatient care. As we will see later, our measure of hospital costs solves this problem by distinguishing between outpatient visits and inpatient cases and by incorporating both types of hospital costs. A second problem with the use of the ACPD is that it is sensitive to differences in accounting practices among hospitals, especially in the treatment of depreciation.[73]

Since our objective is to measure costs in a way which reflects differences in efficiency among hospitals, we clearly cannot use the average daily service charge. Hospital costs are incurred in the production of total patient care, including all the services not reflected by the ADSC. Moreover, the ADSC is easily subject to manipulation by the hospital. If it were to serve as the standard in an incentive reimbursement plan, it could actually lead to hospital inefficiency and a higher overall hospital bill for ociety. An institution, for example, could charge artificially low oom rates, recoup any losses from charges for other services, d, at the same time, reap further gains through financial rewards its seemingly efficient operation. In that case, there would be efficiency incentives and the total bill to society would increase.

Although cost per patient day is a much more comprehensive sure of hospital costs, it is also unacceptable as a measure of oital efficiency. As shown in Chapter II, the patient day is an f

b

w

a

f

i

a

c

O

v

re

to

ca

a

the

inappropriate measure of hospital output. Moreover, if an incentive reimbursement plan paid hospitals according to ACPD, inefficiency might actually be encouraged and the total hospital bill to society increased. A hospital could lower its ACPD, for example, by extending its typical length of stay, since the marginal cost of an additional day is lower than the average cost per day after the first few days of treatment.[74] In that case, the hospital could benefit from any financial rewards afforded by incentive reimbursement, but the total bill to society would increase unnecessarily. To this, we must also add the social costs of the misallocation of resources, and the possible pain or loss of life if hospital beds are not available for the treatment of other more urgent cases.[75]

For these reasons it would seem that the most logical choice is to compare hospitals according to the cost of the entire stay for an average case. M. Feldstein[76] also recommends the average cost per case as the most appropriate measure of hospital costs. One of the major advantages of this approach is that it could provide hospitals with financial incentives to reduce cost per case by reducing the length of stay, a move which in many cases may lead greater efficiency. It was previously shown, however, that the se is also an inadequate definition of the hospital product and that preferable measure of output should include information on both number of cases and the hospital's typical length of stay. For

tha
an
thi
enc

ide
in t
effi
are
thir
diff
mir
firn
with
unit
func
If s
mor
be l

effic
pital
The

hat reason, the proposed measure of hospital costs contains an adjustment for length of stay differences among hospitals. At this point, however, the relationship of costs to productive efficiency should be examined.

In a hypothetical situation where two identical firms produce identical products using the same kinds of inputs, any differences in unit costs would be reliable indications of relative differences in efficiency. Production costs are incurred as fixed and variable inputs are combined to produce certain quantities of output. All other things equal, therefore, unit cost differences are due either to differences in the production functions or to a failure to produce at minimum cost by one or both firms. More explicitly, if the two firms combine inputs in different ways, unit costs will be different with the firm using the most efficient technique experiencing lower unit costs. Similarly, if both firms use the same production functions, they may still display differences in technical efficiency. If some of the inputs employed by one firm (say, management) are more productive than in the other, unit costs in the first firm will be lower.

Unfortunately, the relationship between hospital costs and efficiency is not so clear. A great number of factors affect hospital costs, many of which are unrelated to the degree of efficiency. The subject of inter-hospital cost variation has been discussed

e
i
h
t
a
p
p
e
p
c

d
th
c
e
f
p
th
th
th
do
no

extensively during the past few years.[77] Previous research has identified several variables associated with cost differences among hospitals. The ones cited most often and on which evidence seems the most conclusive are: percentage occupancy of hospital beds, average length of stay, the existence of internship and residency programs, facilities and services offered, the diagnostic composition of the patient population or casemix and, finally, the efficiency of the hospital as a producing unit. The influence of hospital size on average costs has been analyzed often but no conclusive evidence exists that size in itself has any significant effect.

If any measure of hospital costs is to reflect efficiency differences among institutions this measure must be "purged" of the influence of all the factors which are not associated with efficiency. If a hospital displays high unit costs because it offers an extensive range of services, maintains specialized and expensive facilities, or offers medical education programs, it should not be penalized by the reimbursement mechanism. Although it is possible that medical education could be more efficiently carried on outside the hospital,[78] the fact remains that certain institutions are at this time forced to carry a substantial burden in the education of doctors, nurses, and other medical personnel for which they should not be penalized until other alternatives become available.

of

es

pa

as

as

av

In

se

a l

fa

Ev

ap

co

te

th

ex

ut

ei

pla

tot

the

on

fac

Certain studies have attempted to estimate the influence of facilities and services on hospital costs. One approach is to estimate the expected addition to average cost per case or per patient day attributed to the existence of a specific facility such as a blood bank or the provision of a certain type of service such as family planning. Such estimates are usually derived from average (or total) cost regressions with the use of dummy variables.[79] In 1969, however, the AHA listed thirty-five different facilities and services in its annual survey of hospitals. The estimation of such a large number of parameters requires a number of observations far in excess of the fewer than one hundred used in this thesis. Even if sufficient observations existed, however, the regression approach has several disadvantages. First, there is substantial collinearity between certain facilities and services since the existence of one quite often implies the existence of another. Second, this approach estimates the influence on average costs of the mere existence of a certain facility and not of the extent to which it is utilized. A partial solution to the first problem is to hold the effect of facilities and services constant by including as an explanatory variable in the cost regression a simple count of their total number.[80] This approach is of limited value since it makes the unrealistic assumption that all facilities have the same impact on costs. The second problem has yet to be dealt with in a satisfactory way.

faci

pro

dif

mac

cor

sho

to c

the

the

mea

inl

is,

due

In

suc

of

the

this

hos

lar

The difficulty of estimating separate cost figures for facilities and services will dictate an expedient approach to the problem. It is hypothesized that hospital costs are affected by differences in facilities, but the additional assumption is also made that the distribution of facilities and services is highly correlated with the location of the hospital. Hospital data clearly show that urban and metropolitan hospitals tend to be larger and to offer a greater number of specialized services. In Chapter V, therefore, the hospitals in the sample are grouped according to the degree of urbanization of their service areas, and separate measures of cost for each group are calculated. In this way the influence of differences in facilities and services on average costs is, hopefully, minimized.

The same approach is used to adjust for the cost differences due to the existence of medical internship and residency programs. In this case, it is fortunate that all twenty-one institutions with such programs are located in metropolitan areas, and compose one of the groups for which separate cost estimates are made. None of the hospitals in the sample are affiliated with a medical school, so this source of hospital cost variation is of no concern to this thesis.

A factor which many studies have identified as a source of hospital cost variation is the intensity of capacity utilization. A large part of a hospital's costs are essentially fixed at least in the intermediate run since most costs are determined by the size of the plant and the number of facilities and services.[81] Thus, the main staffing of the hospital is not directly related to the amount of patient care produced. For this reason, it is argued that an empty bed is seventy-five percent as expensive as an occupied bed, which implies that the marginal cost per day is only twenty-five percent of the average daily cost.[82] In fact, in one study it was estimated that the marginal cost of a patient day was from 21 to 27 percent of the average cost, depending on the type of patient treated (medical, surgical, etc.).[83] What this means is that in the treatment of a given patient the addition of an extra patient day will increase total case costs somewhat, but it will also decrease average cost per day for that case. It is possible, however, that at very high levels of utilization (say in excess of ninety-five percent of actual bed capacity) marginal cost may exceed average cost because of over-time labor requirements, scheduling problems, and other dis-economies of large scale production.

The degree of capacity utilization is an indication of the efficiency of use of existing resources, at least in the long run. Although, as shown in Chapter II, the amount of output produced (and thus the occupancy rate) is, in the short run, largely beyond the control of the agents who determine hospital capacity, a chronically low utilization rate should be an indication of long run inefficiency in the use of fixed resources such as hospital beds and facilities. For this reason it was decided that the measure of hospital cost performance should include no explicit adjustment for differences in occupancy rates. The rationale for this is that if hospitals are going to be rewarded or penalized by the reimbursement mechanism for their relative degree of efficiency, the reimbursement formula should contain built-in incentives for the socially efficient determination of capacity.

A word of caution is necessary here. It is well known that because of the random nature of demand for hospital care, hospitals are staffed and equipped for peak-load demand conditions, and that average occupancy is always lower than maximum capacity. The relative degree of variation in the census,*however, is greater for small hospitals than for larger institutions.[84] This is because small hospitals must operate at lower average occupancy in order to maintain the same probability of having available beds for unforseen changes in demand.[85] Similarly, certain rural hospitals must maintain a greater number of beds and facilities than would appear justified by the average daily census if they are the sole providers of hospital care for a fairly large but thinly populated area. If the measure of cost used by a reimbursement plan does not include explicit adjustments for differences in utilization rates,

*The census in any given day is the number of occupied beds.

special note of such systematic biases against certain small or rural institutions must be taken at reimbursement time. Other extreme utilization situations can also be adjusted for with the burden of proof on either the hospital or the reimbursing agency. Although this thesis cannot treat such situations explicitly because of the lack of sufficient data, instances will be noted where special reimbursement consideration may be appropriate.

We now come to the last two major factors responsible for hospital cost variation, namely, differences in patient mix or case-mix and differences in the length of stay. It was shown in Chapter III that differences in these two variables imply differences in the actual amount of output produced by the various hospitals. It follows, therefore, that for any given total cost, differences in the same variables would also imply differences in the unit cost of patient care. It will now be shown that adjustments for casemix and length of stay differences are necessary in order to arrive at a measure of hospital cost performance which reflects hospital efficiency and which can be used in an incentive reimbursement plan.

The Casemix Adjustment

If the hypothesis that casemix affects costs is true, this section will show that the economic incentives and disincentives built into a hospital reimbursement system may affect the pattern of care available to

or

qu

ce

in

le

of

ho

gr

im

re

hos

wa

the

is

ava

con

give

ity

inci

bur

so

to a community in an undesirable way, lower the quality of care, or fail to penalize inefficient methods of operation. Although the question of casemix has received increasingly wide attention in certain recent cost studies mentioned in Chapter II, the various incentive reimbursement plans treat the problem in a more or less cursory manner. For example, the plan by the Blue Cross of Western Pennsylvania simply establishes nine groups of hospitals based on location and the extent of their teaching programs.[86] Although such a grouping has often been used as an implicit adjustment for casemix differences, it has been shown recently to be inadequate.[87] A method often used is to group hospitals according to facilities and services.[88] Such a method was used in Saskatchewan but later discarded, perhaps because the relationship of facilities and services to efficiency of operation is not a clear one. It has been shown recently that the scope of available services is not necessarily a good proxy for the actual complexity of casemix.[89] More explicit attention to casemix was given in a recent reimbursement study,[90] but even there complexity of casemix was only approximated by length of stay and the incidence of multiple diagnoses in various types of cases.

According to the basic design of most incentive reimbursement plans where hospitals are reimbursed on the basis of some target cost or rate of cost increase, institutions with actual costs below the target amount will be rewarded with all or part of the difference. Similarly, hospitals with costs above the target may be penalized with a lower reimbursement. The implicit assumption behind such a method of payment is that cost differences at least among similar institutions are due to differences in the degree of efficiency. If, however, high average costs in some hospitals are due to a higher than average concentration of complicated (and therefore costly) cases, failure to take this into account would penalize institutions which may be otherwise operating quite efficiently. The net effect could be an increasing reluctance to treat such cases, with possibly deleterious effects on the overall quality of care available to a given community. This, for example, could be the case for certain urban hospitals which normally treat a higher than usual proportion of special cases. On the other hand, it is possible for certain hospitals to have low average costs both because of lower input prices such as wages and also because they treat a relatively inexpensive mix of patients. Although such hospitals would appear to function efficiently, this may not be the case. If indeed inefficiencies exist, failure to adjust for casemix would result in a reimbursement amount offering few incentives for more economical operation.

The importance of casemix to reimbursement is examined in a recent study of the Blue Cross incentive reimbursement plan of

in

m

pr

wi

fu

int

wa

tes

ev

bu

"w

mo

ing

whi

hos

Cas

bre

trea

da

use

of Western Pennsylvania.[91] Since the plan does not take casemix into account, the authors hypothesize that hospitals in which casemix is becoming more complex will face relatively more intense pressures to cut costs. Conversely, they reason that institutions with casemix changing towards less costly care would have additional funds to expend on other areas. By disaggregating hospital cases into common diseases, easy surgery, difficult surgery, and a four way classification of the 17 major ICDA groupings of diseases, they test the influence of casemix on relative rates of inflation. The evidence supports their hypotheses, and they conclude that a reimbursement plan which does not adjust costs for casemix differences "would put the administrator of a hospital with a casemix becoming more expensive under relatively unfair pressure."[92] In the following sections, therefore, a "costliness" index will be developed which attempts to take differences in the patient composition of hospital output into account.

Casemix Classification and the Data Used

The ideal casemix data set would consist of a detailed breakdown of the number of cases in various diagnostic categories treated by each hospital during a given time. Unfortunately, such data were not available to this study.[93] The alternative was to use data from the 1969 Michigan Hospital Survey conducted by the

M
h
P
th
c
T
p
S
b
r
is
to
v
T
c
o
p
v

ci
e
a
go

Michigan Department of Public Health. The Survey disaggregates hospital cases into Medical-Surgical, Obstetrics, Pediatrics, and Psychiatric. Since the M-S patients are further broken down into those under and over 65 years of age, a separate category was created out of the latter group which is loosely termed Geriatrics. To these five types of cases data were added on the number of outpatient visits taken from the 1969 American Hospital Annual Survey. The visits were converted into patient day equivalents by multiplying the number of visits by the ratio of outpatient revenue per visit to inpatient revenue per patient day. This method is used by the AHA to express outpatient visits in units equivalent to an inpatient day in level of effort.[94] On the average, this conversion amounts to four outpatient visits for one inpatient day. The data, therefore, represent a departmental mix of patients, or cases, requiring largely different types of treatment with little overlap except perhaps between medical-surgical and geriatric patients. Although not exactly a casemix classification in the conventional sense, use of the term is made throughout the text.

Questions may arise as to whether the breakdown is sufficiently detailed to account for the actual impact of casemix differences on hospital costs. Obviously there are casemix differences among the hospitals, especially within the medical-surgical category, that the above classification into six types of patients does not capture. There are, however, some good reasons why this particular classification of cases is chosen, besides the unavailability of more detailed data. This thesis attempts to derive average cost weights for the various types of cases which are then used in the construction of the cost and output indices. As shown in Appendix A, these weights are derived from average cost functions, the estimation of which becomes very difficult when a large number of independent variables (case-types) is used. First, serious multicollinearity problems have been encountered by other researchers.[95] Second, the use of detailed casemix data increases the probability of measurement error because of the ambiguity of assigning cases to the various case-types. Finally, the inclusion of a large number of independent variables in the cost functions makes the parameter estimates unreliable because of the limited number of observations available. The alternative method of hospital output disaggregation into the six types of patients, on the other hand, avoids all these statistical problems.

Although the classification of cases into six broad types does not allow adjustments for casemix differences within each category of care, it is still possible to capture a large part of their effect on costs. As shown later, the method of cost adjustment used also includes information on the length of stay for each patient-type except for outpatient visits. This approach has actually been used by some researchers in order to adjust for casemix,[96] on the assumption ti

a

w

to

in

In

m

a

th

in

on

of

on

cl

bu

pr

ar

ea

as

pa

a

e

that more complicated cases require longer stays. So, this should at least partially capture the effects on costs of casemix variation within the five relevant categories. It is still meaningful, however, to investigate the extent to which costs vary as a result of differences in the proportions of cases that belong to each of the six patient-types. In other words, the crucial hypothesis to be tested is that a typical medical-surgical case has a different impact on costs than, say, a typical obstetrical or pediatric case. If the hypothesis is true, the casemix adjustment used does indeed perform a big part of its intended function. One advantage behind this approach is that data on the six patient-types will be readily available for the creation of casemix-adjusted hospital cost measures to be used by ongoing or future incentive reimbursement schemes.

Perhaps the best justification of the specific case-type classification used here is that it is especially suited to a reimbursement formula which attempts to affect the efficiency of hospital operation. It is very likely that efficiency varies not only among hospitals but also among the various departments within each hospital.[97] At the same time, it is probably reasonable to assume that the efficiency of operation within each hospital department for different procedures is similar. In other words, if a hospital's surgical department is relatively inefficient, then inefficiencies will probably exist in both gall bladder operations and in heart surgery. The same hospital, on the other hand, may have an efficiently run obstetrical department. In that case, the disaggregation of total hospital output according to the six major departments of patient care is probably more relevant than the more detailed breakdown by procedure or diagnosis, especially since the six departments chosen are the main administrative centers involved in direct patient care. To the extent that departmental costs are related to efficiency, an incentive reimbursement scheme based on departmental measures of cost would provide strong incentives for hospitals to improve operations in badly run departments while rewarding the hospitals for economies in other departments.

Evidence of Differences in Casemix

Data from a sample of 94 Michigan short-term general hospitals are analyzed to determine the existence of differences in casemix.[98] The first step is to comput the means and standard deviations of the proportions of cases in each case-type. In order to measure casemix differences among hospitals we use Pearson's coefficient of variation which shows the standard deviation as a percent of the sample mean. Table 3 shows that hospitals display substantial variability in casemix with respect to five of the six case-types while in the medical-surgical category the standard deviation is seventeen percent of the sample mean proportion. The very high coefficient of variation in psychiatric cases must be interpreted cautiously since 45 percent of the hospitals treat no psychiatric cases while many of the others have only a few patients.

TABLE 3. --Casemix Proportions

Case-Type	Mean Proportions	Standard Deviation	Coefficient of Variation
Medical-Surgical	0.465	0.083	17.82
Obstetrics	0.130	0.062	47.52
Pediatrics	0.114	0.060	53.09
Geriatrics	0.170	0.054	31.99
Psychiatric	0.015	0.028	188.92
Outpatient	0.107	0.056	52.46

Besides the substantial casemix variations among hospitals, Table 4 shows that case proportions are also largely uncorrelated with each other. Although most of the fifteen correlation coefficients are significant at the 95 percent significance level, none of the case proportions shows very strong correlation with another. The most that can be said is that hospitals which treat many medical-surgical cases may tend to have somewhat fewer obstetric and pediatric patients. On the face of such evidence, therefore, the hypothesis that the casemix composition of output among hospitals, as defined by the six case categories, is constant and must be rejected.

TABLE 4. --Correlations Among Case Proportions

	M-S	OB	Ped.	Ger.	Psych.	Outp.
Medical-Surgical	1.000	-0.530	-0.575	0.076	-0.266	-0.216
Obstetrics		1.000	0.153	-0.265	-0.072	-0.188
Pediatrics			1.000	-0.285	0.110	-0.175
Geriatrics				1.000	-0.270	-0.350
Psychiatric					1.000	0.114
Outpatient						1.000

The Effect of Casemix on Cost Variation

After the hypothesis of similar casemixes among hospitals is rejected, the hypothesis that casemix differences are a significant factor in hospital cost variation must be tested. An approximate but simple test is given by the multiple correlation coefficient in a regression of average cost per case on the vector of casemix proportions for each hospital. More specifically, the multiple correlation coefficient, R^2, is an estimate of the proportion of total variation in average costs which is explained by variations in the casemix proportions. Table 5 shows the means for various

TABLE 5. --Effect of Casemix Variation on Selected Hospital Cost Components

Cost Item	Mean	Effect of Casemix (R^2)	Probability
Total	3,193,991	0.939	<0.0005
Total Payroll	2,385,853	0.951	<0.0005
Nursing	635,566	0.589	<0.0005
All Other Personnel	1,592,858	0.866	<0.0005
Supplies	510,441	0.641	<0.0005

hospital cost components as well as the R^2 and the probability (P) that the "true" R^2 is zero, or that casemix does not affect average costs. The high degree of correlation between casemix and costs is quite obvious.

The statistical problems associated with the estimation of the average cost functions are discussed in Appendix A. At this stage the point of interest is simply to establish that casemix differences contribute significantly to the variation of costs among hospitals. The high degree of correlation between casemix and the various cost components is obvious. It is possible, however, that the multiple correlation coefficients calculated from the average cost functions overstate or understate these relationships systematically. If some of the cost variation is due to variables other than casemix and these variables are positively correlated with any of the case proportions, the explanatory effect of the omitted variables will be attributed to casemix, and the R^2 will overstate the effect of casemix on costs. This is certainly true to some extent, although probably not extremely important. We calculated the correlation matrix of the six proportions together with some other variables which are believed to influence hospital costs such as size, utilization, location and the existence of teaching programs. Casemix did not seem to be systematically correlated with any of these variables, and, therefore, it is likely that the specification bias resulting from the omission of relevant explanatory variables in the average cost regressions is not very important. Moreover, whatever upward bias does exist is counteracted to some extent by the fact that, for reasons to be explained in Appendix A, the average cost functions were estimated in their linear forms. If the true cost function is nonlinear the estimated R^2 will therefore be an underestimate of the true value. Taking the net effect of these two possible biases and keeping in mind reservations about the reliability of the estimated R^2's, the conclusion must be that casemix has a substantial effect on hospital costs. The next step must, therefore, be the construction of a measure of costs which adjusts for casemix differences.

The Costliness Index

Adjusting for Casemix Differences

Martin Feldstein in his study of British hospitals suggests a measure of costs which takes casemix differences into account.[99] He defines the costliness index (C_i') for the i^{th} hospital as:

$$C_i' = \frac{\sum X_{ji} c_{ji}}{\sum X_{ji} c_j} \tag{1}$$

where

C_i' = Costliness of hospital i

X_{ji} = Number of cases in case-type j treated by hospital i during the year

c_{ji} = Average cost for one case of type j in hospital i

c_j = Average cost per case of type j in the whole sample

This index compares hospital costs for specific case-types with the sample average costs for the same case-types and weighs these costs by the hospital's casemix composition.

The magnitude of the costliness value for a hospital depends not only on the magnitude of the individual c_{ji}'s but also on the number of cases in each case-type. In the extreme case where a hospital's average costs for every type of care are higher than the sample average, the costliness index obviously has a value higher than one. It is likely, however, that in many hospitals certain departments have costs below the sample average while in other departments costs are higher. This is precisely where the particular method of casemix adjustment used in this thesis becomes relevant, since the effect that these interdepartmental relative cost differences have on the final measure of cost performance depends entirely on the proportion of total patient load treated by each department.

The extent to which differences in costliness imply differences in the efficiency of hospital operation should now be examined. As mentioned earlier, a number of factors are responsible for cost differences among hospitals. Although M. Feldstein's costliness index is adjusted for casemix, the average costs in the numerator are still affected by differences in input prices and other factors, one of which is the degree of efficiency with which the hospital operates. To that extent, all other things equal, higher costliness implies lower efficiency. Even this costliness index, however, does not account for all the other factors which determine hospital costs, and therefore, it is an imperfect indicator of relative hospital efficiency unless further adjusted for these other factors. In the beginning of this chapter certain of the institutional variables, such as differences in facilities and services, which have been found to affect crude average costs were examined. An indirect method of adjustment by grouping hospitals according to location was suggested on the assumption that the existence of such characteristics is highly correlated with the degree of urbanization of the hospital service area. Before this, however, a second major direct adjustment of hospital costs, dealing with the length of stay for patients in each case-type, must be performed.

The Length of Stay Adjustment

The analysis in Chapter III showed the importance of the length of stay in the definition of hospital output. Since output is defined as the amount of patient care over time, the length of stay directly affects the actual amount of output produced and, by extension, the cost of production. The costliness index, however, as derived in the previous section does not take this fact into account. M. Feldstein recognizes the trade-off between cost per case and length of stay, and he considers this as the strongest reason for measuring output in terms of the number of cases treated. He, thus, states that "hospitals should be free to select a combination of length of stay and cost per week and should be evaluated on the resulting cost per case.[100] Now, if length of stay varied only among case-types but not within, or in other words, if a hospital's length of stay is an accurate reflection of its casemix, as M. Feldstein claims, then the casemix adjustment would be sufficient. It was shown, however, in Chapter III, that there are substantial length of stay variations among hospitals in the treatment of similar cases and, therefore, substantial differences in the amount of patient care produced even between hospitals with similar casemixes.

For the sake of illustration, the hypothetical example shown in Table 6 may be used. Consider three hospitals A, B, C, with the same number of cases and average costs per case,

TABLE 6.--Effect of Adjustment for Length of Stay Differences

	X_{ji}	c_{ji}	c_j	C'	l_{ji}	$\log l_{ji}$	l_j	C_i^*
A	100	50	40	1.25	8	2.0794	7	1.17
B	100	50	40	1.25	7	1.9459	7	1.25
C	100	50	40	1.25	6	1.7917	7	1.35

Further, for the sake of simplicity assume that all three hospitals treat similar patients belonging to only one case-type. It can be seen from the table that their costliness value without adjusting for length of stay will be the same (C' = 1.25). In other words, an incentive reimbursement scheme which pays hospitals on the basis of costliness would treat all three hospitals in the same way. In terms of adjusted patient care, however, hospital A has produced more output than either B or C, and, since it has managed to do this at the same average cost per case, it is probably more efficient in the economic sense. It is obvious, therefore, that a desirable costliness index must somehow indicate these differences in efficiency.

Adjusting for Length of Stay Differences

A simple way to adjust for length of stay differences is an extension of the previous costliness index. Let us define:

(2) $$C_i^* = \frac{\sum X_{ji} c_{ji}}{\sum X_{ji} c_j} \cdot \frac{\sum X_{ji} \log l_j}{\sum X_{ji} \log l_{ji}}$$

where

l_{ji} = average length of stay for case-type j in hospital i

l_j = sample average length of stay for case-type j

This formulation of the costliness index adjusts casemix-adjusted hospital costs by the actual amount of patient care produced by a given institution. The second expression on the right-hand side of the equation is the ratio of patient care "expected" from a hospital on the basis of the sample average length of stay for each case-type, to the actual amount of care produced by the hospital. The reason behind the logarithmic adjustment is shown in Chapter III.

The fact that average cost per day for a given case is not constant is particularly important when hospital reimbursement is considered. If hospitals are paid an average per diem rate (even if this is adjusted for casemix differences), this results in underpayments for the first days of hospitalization and overpayments in the last days.[101] In that case we could reasonably expect hospitals to attempt to reap some financial gain by extending the period of treatment beyond the medically necessary length of time. The logarithmic adjustment used above is an attempt to deal with this problem by giving higher weights to the first few days of a patient's stay. A similar approach was used by the Philadelphia Blue Cross plan which for many years paid on a sliding per diem rate that was highest for the first day of stay and was reduced thereafter.[102]

If a hospital shows low (say, less than average) c_{ji}'s for one or more case-types this could be due either to particularly efficient operation of certain departments, to shorter than average stays, or, most likely, to a combination of both factors. The length of stay adjustment used here provides a means of distinguishing between these two factors and makes it possible to concentrate on relative efficiency differences. If a hospital achieves low costs per case by keeping stays shorter than average it will display a higher C_i^* value than a hospital with the same number of cases and average cost per case if the latter shows longer stays and, thus, provides more patient care. If, however, low average costs per case are achieved by an institution despite higher than average lengths of stay, this apparently high degree of efficiency will result in low values for c_{ji} and therefore will be reflected by low costliness.

An example of the actual effects of the length of stay adjustment on the costliness index is shown in the last column of Table 6. The presumably higher productivity of hospital A is reflected in the lower costliness index while hospital B is unaffected. Similarly, the higher costliness value for hospital C is due to the fact that it produced less output because of its shorter than average length of stay.

A most important point must be made here. The method of adjusting for length of stay differences shown above seems to offer economic incentives to hospitals to keep patients longer in order to improve their position on the costliness scale. As mentioned earlier, this is one of the strongest objections to using cost per patient day as the basis for reimbursement. This problem could be avoided by reversing the length of stay adjustment factor (2) and weighing the casemix-adjusted costs by $\frac{\sum X_{ji} \log l_{ji}}{\sum X_{ji} \log l_j}$. In that case a hospital is penalized for stays longer than average. There are, however, some good reasons why this is not a desirable formulation of the costliness index. The examination of the dimensions of hospital output showed that an extra day's stay for any given patient represents a certain amount of additional patient care for which the hospital should receive some credit. Moreover, longer stays do not necessarily improve (lower) a hospital's costliness since they also increase average cost per case, which is reflected by C_i^*. Finally, since the weight given to each additional day of care decreases because of the logarithmic transformation, the incentives for excessively long stays are weakened even further.

Besides the reasons mentioned above, there is a more fundamental argument in favor of the suggested length of stay adjustment because of the vital importance of the hospital product, it is important that the efficiency incentives embodied in the reimbursement mechanism do not also become incentives for the reduction of the quality of hospital care. Although most of the hospital literature seems to focus on the need to guard against unnecessarily long stays, there is also evidence that patients are sometimes discharged before their treatment is complete.[102] This thesis assumes that the consequences of such compromises in the quality of care are more serious than the missallocation of resources resulting from unusually long stays. For this reason, the costliness index is designed to penalize discharge policies oriented towards stays which are shorter than medically necessary as long as such stays are shorter than average.

A related reason why longer than average lengths of stay should not be directly penalized by C_i^* is due to the recognition of the heterogeneity of hospital output. As shown earlier, lengths of stay for different patients vary substantially even within very specific diagnoses or case-types because of differences in particular patient characteristics, recovery rates, number of complications, etc. It is very likely that certain institutions treat a higher proportion of patients who require longer stays than usual.[103] This consideration is particularly important, since the broad casemix classification used makes length of stay a partial surrogate for casemix differences <u>within</u> each case-type. The costliness index, therefore, is constructed so as to discourage shorter than average stays without penalizing a hospital which treats an unusual number of patients requiring longer hospitalization. If there, indeed, exist incentives for hospitals to prolong a patient's stay for other than medical reasons, other types of controls such as utilization control, recertification, and claims review can be used.[104] Such direct controls are specifically designed to prevent overutilization of hospital beds,[105] and they may be better suited to prevent unduly long stays than the payments mechanism because they can examine each case individually.

Estimation of C_i^*

The estimation of costliness index C_i^* requires data on c_{ji}, c_j, l_{jk}, and l_j. Unfortunately, data on c_{ji}, the actual average cost per case by case-type, do not exist for each hospital. There exist, however, data on $\sum X_{ji} c_{ji}$ which is the total cost of patient care in each institution reported by the AHA. These data do not include depreciation items, and, therefore, they approximate the total variable cost of patient care fairly closely.

Since data on c_j do not exist the average cost per case for each case-type must be estimated from another relationship. As shown in Appendix A, the c_j's are estimated from a regression of hospital average cost per case (c_i) on the proportions of cases in each case-type:

$$c_i = \beta_0 + \beta_j p_{ji} + \epsilon_i \quad i = 1 \ldots N, \quad j = 1 \ldots (k-1),$$

where $p_{ji} = \dfrac{X_{ji}}{\sum X_{ji}}$

The c_j's are then calculated as:

$$c_j = \beta_0 + \beta_j$$

$$c_k = \beta_0$$

Finally, information on l_{ji} and l_j is obtained by dividing the number of patient days in each casetype by the number of cases. The next step is the examination of the estimated values of C_i^*, their relationship to hospital observed costs, and the implications of using the costliness index as the basis for a hypothetical incentive reimbursement plan.

CHAPTER V

THE APPLICATION OF COSTLINESS IN INCENTIVE REIMBURSEMENT

This chapter analyzes the actual costliness values derived for a sample of 94 hospitals. Costliness is then compared to average cost in order to examine the implications of using the two as alternative measures of hospital cost performance by an incentive reimbursement scheme. A discussion of the sample and the data used can be found in Appendix B.

Table 7 shows means and standard deviations for observed relative costs (Cr) and costliness values (C*). Cr is the actual average cost per case as a percent of the sample average. C* is an index number representing the ratio of the observed total cost of patient care to the cost expected from a hospital on the basis of its casemix and typical length of stay for each case type. The two measures of hospital costs are not, therefore, comparable in terms of their absolute magnitudes. Fortunately, what is interesting is not the absolute values of Cr and C*, but rather their distribution and the way they rank hospitals on a reimbursement scale. The

TABLE 7. --Distribution of Relative Cost and Costliness

	Mean	Standard Deviation	Coefficient of Variation
Cr	99.6	23.44	23.53
C*	100.3	24.10	21.23

theoretical distinction between Cr and C*, for example, would be meaningless if the two variables varied together because then hospitals would tend to be in the same relative positions, a fact which would result in similar reimbursement amounts.

A simple test of the relationship between Cr and C* is the correlation coefficient from a regression of costliness on relative average cost:

$$C_i^* = \beta_0 + \beta_1 Cr_i + \epsilon_i$$

The correlation coefficient $R^2 = 0.415$ implies that approximately 41 percent of the variation in C* is due to differences in average costs among hospitals. As expected, relative cost does not by itself explain a very large proportion of differences in the actual cost of hospital care when casemix and length of stay differences among hospitals are taken into account.

An interesting finding is the low correlation (0.191) between costliness and hospital bed size. Average cost and size on the other hand shows a significantly higher correlation (0.570).[106] This is consistent with one of the few major questions on which hospital cost researchers are in substantial agreement, that is that size in itself does not have a significant effect on average costs and that the average cost curve is a flat U over most of the relevant range of output. The fact that observed average costs are usually higher in large hospitals simply means that these hospitals also treat a more complicated mix of cases, and are usually located in metropolitan or urban areas facing higher input prices and extreme demand conditions. From the correlations above it seems that the casemix and length of stay adjustments used in this thesis are fairly successful in purging costliness from the spurious correlation observed between size and average cost.

One final observation concerns the relationship of Cr and C* to total hospital costs. Again, the correlation between costliness and total costs is low (0.283) while that between average and total costs is substantially higher (0.690). The second relationship implies certain diseconomies of scale, a phenomenon not substantiated by any empirical evidence and probably due to the same spurious relationship as that between size and average costs.

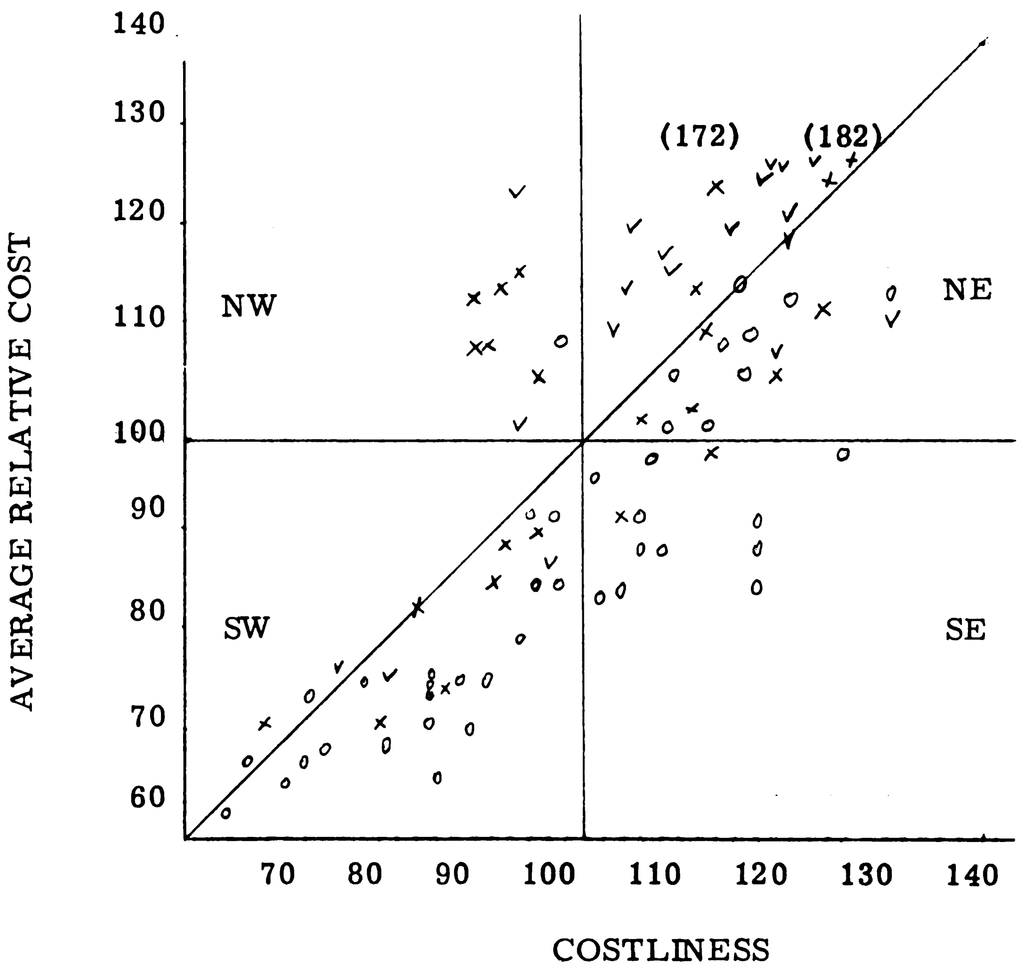
Table 7 shows the similarity of the distributions of Cr and C*. The standard deviations are almost the same and the coefficient of variation of Cr is only 10.3 percent higher than that of C*. Another indication of the similarity of the distributions of Cr and C* is revealed in Table 8 where hospitals are grouped according to the number of standard deviations above or below the mean Cr and C*.

TABLE 8. --The Distribution of Cr and C*

	$(\mu - 3\sigma)$	$(\mu - 2\sigma)$	$(\mu - \sigma)$	$(\mu + \sigma)$	$(\mu + 2\sigma)$	$(\mu + 3\sigma)$
Cr		15	35	30	11	3
C*	1	16	30	32	13	2

The fact that the distributions of Cr and C* are similar might suggest that they are interchangeable as measures of hospital costs. As shown below, however, this is not true when the specific values of Cr and C* for individual hospitals are examined.

Figure 2 plots the pairs of Cr and C* values for the 94 hospitals. If both costliness and relative cost measure hospital cost, even though they are not theoretically comparable in terms of their absolute values, the fact that the means and variances of their distributions are almost identical should cause the various points to lie near the 45^{o} line. In other words, hospitals with high relative cost



Location codes: o = rural hospitals, x = urban hospitals, v = hospitals in Detroit.

FIGURE 2. -- Relationship Between Cr and C* for 94 Michigan Hospitals

should also show high costliness and vice versa. The wide scatter of points in Figure 2, however, shows that this only very roughly approximates the actual case. The figure is separated into four quadrants by using the point given by the two means as the origin.

The points show that a substantial number of hospitals are in the northwest and southeast quadrants, that is, where Cr and C* move in opposite directions relative to their respective means. Even in the other two quadrants, however, where Cr and C* lie on the same side of their mean values the distances of the points from the 45° line indicate a substantial divergence between costliness and relative cost.

The data shown in Figure 2 are summarized in Table 9 where hospitals are grouped according to their position relative to the sample means of Cr and C*. The table shows that costliness splits the sample hospitals into two almost even groups, whereas

TABLE 9. --Hospital Ranking Relative to Mean Relative Cost and Mean Costliness Values

Relative Cost (Cr)				Costliness (C*)	
Cr $<\mu$		Cr $>\mu$		C* $<\mu$	C* $>\mu$
50		44		46	48
C*$<\mu$	C*$>\mu$	C*$<\mu$	C*$>\mu$		
37	13	9	35		

with relative cost as the measure more hospitals (50) are in the low-cost category. The distribution of C* is almost perfectly symmetric with the mean and median at 100, while the median for Cr is 98 or slightly less than the mean of 99.6. The fact that relatively more hospitals show less than mean relative costs is due to the existence of the few hospitals with unusually high relative costs shown at the top of the northeast quadrant of Figure 2.

Costliness, Relative Costs, and Efficiency

The only evidence presented to this point is that Cr and C* are not, in fact, equivalent measures of hospital costs. The analysis has not yet, however, touched on the main point, namely that costliness is a better measure of costs because it reflects differences in efficiency of operation. Unfortunately, there is no way to perform a rigorous and formal statistical test of this hypothesis in this chapter. The actual relationship of efficiency to costliness and average costs will be discussed and tested in Chapter VI. A look at the data at this point, however, can give some indications on the effects of the casemix and length of stay adjustments and on the actual financial implications of using costliness as a measure of cost and efficiency.

We examined the data for the 22 hospitals which show either low costliness and high relative costs (9 hospitals) or high C* and low Cr (13 hospitals). We looked at the occupancy rate (U) for each of these hospitals on the assumption that it is an indication of the efficiency of use of fixed resources. The relationship between Cr and U presents a seemingly paradoxical phenomenon. The nine hospitals with high relative costs also show a markedly high degree of efficiency (high U) at least in the use of fixed plant and equipment. This is shown by their much higher average utilization rate which runs at 81 percent of capacity. The low relative cost hospitals, on the other hand, appear to be less efficient in their use of plant as shown by the low utilization rate of 69 percent. This paradox does not occur when costliness is used as the measure of costs, since the differences in the efficiency of use of fixed resources are reflected by the differences in costliness in the two groups of hospitals.

Table 10 lists utilization, bed capacity, and location data for the 22 hospitals of interest. It is interesting to note that of the nine low costliness-high cost institutions only one is in a rural area while the rest are either urban or in metropolitan Detroit. The opposite phenomenon occurs in the thirteen high costliness-low cost hospitals where only two are located in urban areas. It is also shown that the average size in the first group is more than twice that in the second. As a subsequent section will show, there are good reasons to expect higher costs per case in urban and

TABLE 10.--Selected Data for Hospitals in Which Costliness and Relative Cost Diverge

Low C* - High Cr				High C* - Low Cr			
Hospital Number	Location[a]	Utilization Rate (%)	Beds	Hospital Number	Location	Utilization Rate (%)	Beds
2	0	74	89	1	0	63	172
23	6	92	120	3	0	65	90
33	3	88	125	4	0	70	146
52	2	82	276	6	0	59	187
55	2	82	276	13	3	65	70
58	6	88	297	15	0	79	132
81	2	74	296	36	0	61	44
82	2	85	257	43	0	86	55
93	6	79	283	46	0	57	43
		$\mu = 81$	$\mu = 232$	62	4	74	198
				86	0	73	160
				88	0	71	34
				89	0	80	82
						$\mu = 69$	$\mu = 108$

[a]The zero code identifies rural hospitals, codes 2, 3, and 4 are for urban institutions, and code 6 for Detroit hospitals.

metropolitan hospitals. The interesting finding at this point is that the nine hospitals in the first group manage to show low costliness despite their high average costs and that the other thirteen hospitals show high costliness even though they have low average costs per case. Since the actual cost per case is also included in the numerator of the costliness index, there must be some additional factors responsible for the reverse relationship between costliness and average cost. In fact, it appears that these factors are related to the case-mix composition and the lengths of stay characteristic of the patient populations.

If the first nine hospitals have a high proportion of cases of the more expensive case-types but are also characterized by efficient operation so as to have lower than average costs per case for these case-types, they should theoretically show high relative costs but low costliness. Moreover, if their casemix composition <u>within</u> each category of cases was characterized by a higher concentration of complicated cases with longer lengths of stay, again they should show low costliness and high relative costs. The arguments above could be reversed, of course, for the thirteen hospitals which show low costs, but inefficient production as displayed by their high costliness values. If low average costs are due not only to location[107] but also to less expensive casemix and short lengths of stay, the high costliness indicates 1) inefficient use of resources, and/or 2) the provision of less than average patient care to the average patient in each case category.

The previous analysis offers some good indications of the relationship between costliness and efficiency. It also shows that C* is a more accurate measure of efficiency than relative average cost. It must be said, however, that, theoretically, low costliness is neither a necessary nor a sufficient condition for a high degree of efficiency. An efficiently operating institution may still display high costliness if its average costs are abnormally high because of other reasons such as the existence of a large number of expensive facilities, the provision of many specialized services, and the existence of extensive teaching programs. On the other hand, an inefficient hospital could show a low costliness value if, for the reverse reasons, its average costs were significantly below the state average for similar case-types. Despite the possibility of such extreme cases, the theory behind the costliness index as well as some of the evidence in this study show it to be a satisfactory measure of hospital efficiency. The reservations mentioned above, however, do not permit any strong statements concerning efficiency for hospitals whose costliness values differ only by five or ten percent. Throughout this thesis, therefore, the more modest claim is made that the costliness index provides strong indications of differences in efficiency only for hospitals where the index values diverge substantially, say, in excess of thirty percent, and especially in cases where relative costs and costliness move in opposite directions in relation to their respective means.

Financial Implications of the Costliness Index

We will now examine the actual dollar implications of using costliness instead of average relative cost as the criterion in an incentive reimbursement scheme. We will assume a hypothetical situation where all hospitals are reimbursed by the same agency, and we will consider two alternative payment formulas. Under formula A the mean observed cost per case for the state ($\bar{C}$) is calculated from crude average cost per case figures. The relative cost (Cr_i) for each hospital is then calculated as:

$$Cr_i = \frac{C_i}{\bar{C}}$$

and it becomes the criterion which determines hospital reimbursement. The base on which it is applied is total hospital costs (TC) and the final payment (R) is determined by the formula:

(1) $$R = TC + TC\left[1/2(1 - Cr_i)\right]$$

In this way if a hospital has relative costs below the state average it receives its full costs plus a reward equal to one-half[108] of the difference between its relative average cost and the state average. If a hospital has costs higher than average it is, of course, penalized in the same way.

Formula B works in a similar way, except that costliness instead of relative average cost is used. Hospital payment under this method is:

$$R = TC + TC[1/2(1-C_i^*)] \tag{2}$$

Again low-cost hospitals are rewarded by one-half of the "savings" achieved or penalized by one-half of the excess over average costliness in the state.

We should point out that instead of using a fixed reimbursement factor, a third party could allow the percentage of payment over or below total hospital costs to vary for different relative cost and costliness intervals.[109] Our interest at this point is to demonstrate the net effect on hospital payments resulting from the use of formulas A and B. Since the distributions of Cr_i and C_i^* are fairly similar, our results do not depend heavily on the choice of the cost adjustment factor as long as it is the same in both formulas.

Empirical Results

As we see in Table 11 both formulas A and B would result in payments to hospitals during 1969 which would cover less than the total costs of operation. This means that the total savings resulting from lower reimbursement to high cost hospitals exceed

TABLE 11. --Reimbursement Amounts Under Formulas A and B and Total Hospital Costs

	Total Cost	Formula A	Formula B
Amount in $	$300,235,182	$274,545,889	$290,447,133
% of Total Hospital Costs	100	91.6	97.3

the total rewards to low cost institutions. In the long run, gains in efficiency induced by incentive reimbursement would lead us to expect precisely such an outcome. As P. J. Feldstein concludes:

> In summary, rewarding hospitals whose operating costs are below the mean and penalizing those whose costs are above it, would result in the total amount expended on hospital reimbursement being less than if total costs were reimbursed.[110]

Although we cannot be certain that both incentive reimbursement methods possess one desirable quality vis-a-vis payment of full costs, namely, the immediate containment of total outlays for hospital care, it is apparent that the reimbursement varies substantially between formulas A and B. If our sample is representative of the entire population of Michigan short-term general hospitals, we can see, on the basis of our calculations, that payments under formula A will be substantially less than if costliness is used as the reimbursement criterion. If hospitals in 1969 were paid on the basis of relative cost, they would incur a deficit of 8.4 percent of total costs, or more than three times the 2.7 percent deficit which would result if reimbursement were based on costliness.

The substantial difference between the savings possible under the two reimbursement formulas does not by itself suggest that the use of formula A is inappropriate. If we have reasons to suspect inefficiencies in the production of hospital services throughout the industry there is no theoretical reason why savings from increased efficiency in the next period should not be as high as 8 percent or more. The superiority of formula B, therefore, is based solely on the theoretical derivation of the costliness index and its relationship to efficiency. In fact, even if the savings possible under formulas A and B were reversed in our empirical results, we would still have to defend formula B as the appropriate reimbursement method.

Rather than adding to the theoretical validity of costliness reimbursement, our empirical results show that it may also be more practical than incentive reimbursement on the basis of costs. Given the already substantial rate of hospital cost increases, it is likely that reimbursement well below total operating costs would cause severe hardship for a number of institutions and would not only impede future quality improvements but would, in all likelihood, result in deterioration of present quality levels. If the total reimbursable amount under relative cost reimbursement for the industry is 91.6 percent of total costs, this implies that a number of hospitals would receive amounts well below 90 percent of costs. It is doubtful that any increase in efficiency would allow these institutions to maintain the scope and quality of their services at that reimbursement rate.

This point becomes even more interesting when we consider the actual dollar amounts of reverse reward and penalty payments under formulas A and B. In Tables 12 and 13 we see the dollar amounts reimbursable to selected hospitals under formulas A and B, as well as their total costs for 1969. Table 12 shows figures for the thirteen hospitals with below average costs and above average costliness while Table 13 shows the nine high cost-low costliness institutions. Although these hospitals are somewhat extreme cases, together they represent 23.4 percent of the sample. From our earlier analysis there are sound theoretical

TABLE 12.--Actual Reimbursement Amounts to Low Cost-High Costliness Hospitals ($)

Hospital Number	Total Cost	Payment Under Formula A	Payment Under Formula B	Reward Under Formula A	Penalty Under Formula B
1	3,397,010	3,543,590	3,046,608	146,580	-350,402
3	1,247,574	1,251,316	1,009,848	3,742	-237,726
4	843,019	887,867	829,741	44,848	- 13,278
6	3,098,330	3,201,039	3,063,163	102,709	- 35,167
13	1,417,340	1,492,742	1,364,048	75,402	- 53,292
15	727,409	780,728	715,952	53,319	- 11,457
36	1,158,270	1,167,594	1,129,892	9,324	- 28,378
43	445,074	448,656	430,297	3,582	- 14,777
46	3,125,789	3,182,209	3,062,960	56,420	- 62,829
62	3,195,919	3,349,962	3,173,547	154,043	- 22,372
86	436,430	457,465	425,737	21,035	- 10,693
88	869,532	890,276	850,763	20,744	- 18,769
89	765,341	773,482	760,246	8,141	- 5,095
Total				699,799	-863,371

reasons for expecting low cost-high costliness institutions to be less efficient than hospitals with another combination of relative cost and costliness. Relative cost reimbursement, however, will result in payments in excess of total costs (a reward) for the former while the latter will be penalized. As we see in Tables 12

TABLE 13.--Actual Reimbursement Amounts to High Cost Low Costliness Hospitals ($)

Hospital Number	Total Cost	Payment Under Formula A	Payment Under Formula B	Penalty Under Formula A	Reward Under Formula B
2	7,257,605	5,785,762	7,428,884	-1,471,843	171,279
23	1,447,649	1,146,755	1,469,146	- 300,894	21,497
33	2,294,520	2,289,930	2,398,805	- 4,590	104,285
52	5,158,128	4,759,404	5,933,652	- 398,724	775,524
55	5,881,440	5,544,727	6,431,648	- 336,713	550,208
58	11,202,195	9,661,333	11,448,083	-1,540,862	245,888
81	5,078,949	4,915,660	5,487,296	- 163,289	408,347
82	4,594,885	4,447,159	4,863,915	- 147,726	269,030
93	5,924,609	5,020,513	6,608,308	- 904,096	683,699
Total				-5,268,737	3,008,457

and 13 the dollar amounts of these rewards and penalties for most hospitals would be substantial.

We should point out that there is no clear theoretical reason why payments under formulas A and B should necessarily be less than total costs during any particular period of time. One possible explanation is that the average large hospitals in the sample tended to have higher mean costs and higher costliness than the average small hospital. As a result the absolute amounts of penalties

exceeded the amounts of rewards. Of course, if a mean cost per case weighted by bed size were used in the calculations, the total reimbursement amount would be much closer to total costs under both formulas. The fact that the cost figures were not weighted by bed size may also explain the fact that formula B resulted in higher payments to hospitals than formula A. As we see in Tables 12 and 13, total rewards under formula B were far in excess of total penalties while the reverse occurs under formula A. This is due to the fact that the hospitals in Table 13 are much larger than those in Table 12.[111]

The Influence of Location on Hospital Costs

We showed earlier that hospital costs often vary for reasons other than efficiency. We know from the existing literature that the geographical location of a hospital is usually a good proxy for many of the real variables which affect its costs. For a variety of reasons besides input productivity and output composition, hospitals in large metropolitan areas such as Detroit would be expected to have higher average costs than rural institutions. First, we know that salaries for hospital employees are higher in Detroit than in the rest of the state. For example, during 1969 average earnings for general duty nurses--the largest category of registered nurses--were $8,216 in Detroit while in the rest of the state they were $7,550.[112]

Secondly, metropolitan hospitals tend to have a wider range of facilities and scope of services.[113] Although in most instances their presence is related to demonstrable community needs, the end result is higher average hospital cost per case. Finally, urban hospitals are often subject to extreme demand conditions because of the higher population density in their service areas. Against all these factors contributing to higher costs in metropolitan areas we must consider some cost-saving factors. To the extent that metropolitan hospitals tend to be larger we can expect them to achieve certain economies from bulk transactions, especially on supplies, or from running their own hotel services, such as food and laundry. On balance, however, since labor costs are roughly 65 percent of total operating costs, and because of all the other factors, we still expect metropolitan hospitals to have higher average costs.

A payment method which is intended to promote efficiency without impairing the quality of services should make adjustments for the real variables which affect hospital costs but are not related to efficiency of operation. This is precisely the purpose behind the costliness index proposed in this thesis. If, however, costliness varies significantly among various geographical locations, then we have reason to suspect that even this measure of hospital costs should be adjusted in order to portray relative cost performance with sufficient accuracy. For precisely this reason advocates of the various Regional Average Cost incentive reimbursement schemes would assign hospitals to fairly homogeneous regional groups and determine the reimbursement amount on the basis of the group average.

In this section we consider the influence of location on both the average relative cost and the costliness of a hospital. The AHA data include information on the location of each sample hospital and on the population of the community in which it is located. There are seven size classes: non-SMSA areas, 50,000-100,000; 100,000-250,000; 250,000-500,000; 500,000-1,000,000; 1,000,000-2,500,000; and 2,500,000 and over. Of the 94 hospitals in our sample 47 are in the first or rural category. There were no hospitals in the second and sixth categories. Twenty-six were in categories three, four, and five (nine, eleven, and six respectively), and twenty-one in the last category, which is the Detroit area. We therefore classified hospitals in three classes: rural, urban, and Detroit.

For the reasons mentioned above we hypothesized that average relative costs will differ widely among the three classes of hospitals. We also hypothesized that the costliness values will differ less widely than the actual average costs. The reason for the latter hypothesis is that if our casemix adjustment operates properly, certain rural hospitals which have lower relative costs because operating costs in their region are low and because of an inexpensive casemix will, nevertheless, have a high costliness value if they are relatively inefficient in their operation. Conversely, certain urban hospitals with high labor costs and an expensive casemix may have lower than average costliness if they are efficient in treating these cases. Although we expect the casemix adjustment to even-out some of the cost differences which are due to location, the reasons mentioned earlier still lead us to expect differences in costliness values between rural and Detroit hospitals.

The figures in Table 14 support our theoretical expectations in a very convincing manner. Mean relative costs vary substantially among the three classes of hospitals. In order to make sure that this is not a random occurrence, we performed statistical t-tests on the equality of the means for each of the three pairs. All three mean relative costs proved to be different from each other at the 95 percent level of significance.

TABLE 14. --Distribution of Cr and C* by Location

		Relative Cost (Cr)			Costliness (C*)		
	N	Mean	Standard Deviation	Coeff. of Variation	Mean	Standard Deviation	Coeff. of Variation
Non-SMSA	47	90.1	16.2	17.9	97.9	26.6	23.9
100,000-1,000,000	26	104.0	21.9	21.0	102.3	19.3	17.2
2,500,000 and over	21	115.5	28.9	25.0	108.1	29.2	23.4

As expected, the differences in mean costliness values for hospitals in the three location classes were not as pronounced. Costliness in rural hospitals was almost as high as that in urban non-Detroit institutions. The only statistically significant difference (at the 95 percent level) was between rural hospitals and those in Detroit. Even there the difference in mean costliness was one-third of the difference in mean relative cost.

One conclusion arising out of these results is that adjusting for the location of the hospital is very important when average relative cost is used as a measure of the hospital's cost performance, and somewhat less important when costs are adjusted for casemix and length of stay. Even in that case, however, a provision should be made to distinguish between rural, urban, and Detroit hospitals.

Although the costliness index is less affected by differences in location because it concentrates more on efficiency differences, an important question still remains. Could we achieve an equal degree of homogeneity by grouping hospitals according to their location and perhaps according to some other characteristics and then use the within-group average relative cost as the standard for reimbursement? We saw in Chapter II that this is the approach suggested in many of the incentive reimbursement proposals. Can we then be fairly certain that relative average cost within each group is as satisfactory a measure of hospital efficiency as the costliness index?

First of all, we must reject this idea on theoretical grounds. If it is true that casemix and length of stay affect hospital costs and the actual amount of output produced we must, as we have shown, take these factors into account. The use of regional relative cost would be justifiable only if casemix and length of stay variations were much less pronounced within a region or a group of hospitals. Since we have no strong a priori reason to expect this to be the case we decided to look into the matter. We computed regional average costs and costliness for the three groups of hospitals

TABLE 15. --Number of Hospitals Where Regional Relative Cost and C* Move in Opposite Directions

	$Cr < \mu$ and $C^* > \mu$	$Cr > \mu$ and $C^* < \mu$
Non-SMSA	8	4
100,000-1,000,000	3	4
2,500,000 and over	4	3
TOTAL	15	11

according to the classification shown in Table 15. From the table we see that if the regional average relative cost were used to calculate the reimbursement amount twenty-six hospitals would be paid in a manner opposite to that suggested by costliness. Although regional relative cost adjusts for certain regional cost differences, therefore, its use seems to increase the number of cases in which financial rewards and penalties would be distributed in a manner unrelated to efficiency of operation,[114] at least as reflected by costliness. Moreover, the incidence of reverse incentives would be highest in Metropolitan Detroit hospitals where fully one-third of the institutions would be either rewarded for inefficiency or penalized for what might be high quality care.

An Alternative Classification

In order to examine the influence of location on hospital costs and costliness even further, we used a somewhat different classification of hospitals. Michigan Blue Shield currently divides the state into four "prevailing" areas characterized by relatively homogeneous physician billing practices. These areas are generally characterized as metropolitan, urban, suburban, and rural. A listing of the counties in each area is found in Appendix C. Area I is composed of metropolitan Detroit and Area IV is the whole of the upper peninsula. We computed average relative cost and costliness values for the hospitals in our sample which are located in each of the four areas. Mean Cr and C* values are shown in Table 16.

The implications of this alternative classification are similar to the previous ones. Both relative costs and costliness seem to differ among areas, although the differences in costliness are somewhat less pronounced. Suburban hospitals decidedly show the lowest average costs but their differences in costliness from urban hospitals and those in Detroit are much smaller. A possible explanation is that casemixes in suburban hospitals are less complicated requiring generally shorter stays and lower costs per case. Although the degree of efficiency implied by the costliness index is

TABLE 16. --Classification of Hospitals by Blue Shield Prevailing Areas

	Area I (Detroit)	Area II (Urban)	Area III (Suburban)	Area IV (U. P.)
Number of Hospitals	23	20	44	7
$\bar{C}r$	113.2	108.1	86.8	110.5
$\bar{C}*$	106.7	101.7	96.6	113.3

not as high as that shown by relative costs, suburban hospitals appear to be the most efficiently run institutions. Although the upper peninsula hospitals are too few for any definite statements, they appear to be uniformly high-cost institutions. Since there are reasons, such as lower wage rates, which would make us expect actually lower costs in these hospitals, we must conclude that considerable inefficiency exists, part of which is shown by the abnormally low utilization rates in these seven rural hospitals.[115] There may, however, exist certain extenuating circumstances which should be taken into account and with which we will deal in the remaining paragraphs of this chapter.

Reimbursement Implications

In Tables 12 and 13 we showed the reimbursement implications in dollar terms for the twenty-two hospitals where costliness and relative costs move in opposite directions. The reimbursement amounts were derived from average costliness and cost data taken from the whole sample. We now know that regional cost differences affect the costliness index although not as much as they affect average costs. In Table 17, therefore, we calculated the total reimbursement amounts for the four regional

TABLE 17. --Total Reimbursement Amounts Based on Regional Cost and Costliness Estimates (in millions)

	Area I	Area II	Area III	Area IV
Total Cost	$125.9	$100.3	$63.6	$10.2
Formula A	107.4	90.6	66.6	9.7
Formula B	116.0	99.8	65.0	9.2

groups of hospitals under formulas A and B in order to get an indication of the financial implications of the two formulas for the various groups of hospitals.

Both formulas would reward suburban hospitals with payments in excess of total costs. This seems justified in view of the apparently efficient operation of those institutions. The use of the costliness index in formula B, however, would reduce the financial rewards by half, since there are reasons to believe that part of the efficiency as shown by relative costs is due to simpler casemixes and shorter lengths of stay.

Urban hospitals would be severely hurt under formula A since they would be paid almost ten percent less than total costs. The costliness index for that group, on the other hand, reveals no serious inefficiencies and, accordingly, under formula B these hospitals would receive almost their full costs. The most devastating consequences of relative cost reimbursement, however, would be for the Detroit hospitals, where losses would run at almost 15 percent of total costs. Although certain savings can undoubtedly be effected in certain metropolitan hospitals through consolidation of facilities or plain improvements in efficiency, it would be unrealistic to expect such savings to lower total costs by fifteen percent from one year to the next without the substantial danger of compromising the quality of care. Even the significantly lower penalties imposed under formula B (8.5 percent of total costs) should probably be administered with caution and with individual attention paid to certain institutions.

Caution with reimbursement must also be exercised in the case of the seven upper peninsula hospitals, where both high average costs and high costliness seem to be the result of unusually low utilization rates. Although this means that certain fixed resources are underutilized, it does not mean that low occupancy rates are clear signs of inefficient operation. If demand in a particular service area is low, this does not automatically imply that the hospital should not be there. What it does mean, however, is that this hospital should not expand any further and possibly that it should either discontinue certain services or attempt some consolidation with other health facilities in the area. Since most of these hospitals serve large but sparsely populated areas, high costs may be an unavoidable consequence of the indivisibility involved in the construction and operation of a modern hospital.

Chapter Conclusions

We showed in this chapter that both when individual hospital cost performance is measured against a state average and when regional standards are employed, relative cost reimbursement would provide undesirable incentives or penalties for roughly a quarter of our sample of hospitals. To the extent that costliness is a satisfactory measure of hospital efficiency we showed that under average cost reimbursement some hospitals would be penalized not for inefficiency but for high costs due either to an expensive mix of patients, long but probably medically necessary stays, or to other factors beyond the hospital's control. We also saw that in many other cases such reimbursement would result in rewards for inefficient hospitals. Reimbursement on the basis of costliness, on the contrary, seems to reverse these situations by avoiding the undue imposition of penalties which might impair the quality of service and by offering rewards for actually efficient operation.

This last point becomes even more important when the total cost of hospital care to society is considered. We saw in Chapter II that capital spending in the hospital industry is determined not only by demand for the product but also by the availability of funds. In fact, in a recent paper Paul Ginsburg showed precisely that "total investment is not determined by demand for service but by hospital size and accessibility of hospital funds."[116] If that is the case then we can expect that if certain inefficient hospitals are rewarded, the excess funds will be spent in further expansion which does not represent efficient use of resources. To be sure, financial rewards to efficient hospitals will also be used for additions to plants and/or expansion of services but, and this is the important point, the additional resources employed will be used in a more efficient manner. In other words, if costliness is indeed a better measure of efficiency, quality improvements can be effected at a lower cost to society than under relative cost reimbursement. What remains now is to deal with this important point, namely, the formal relationship of costliness to efficiency. This will be the subject of Chapter VI.

CHAPTER VI

PRODUCTIVE EFFICIENCY AND HOSPITAL COSTS

The main theoretical justification in measuring hospital efficiency by costs lies in the inherent relationship of costs to productivity. Costs are incurred as hospitals employ certain fixed and variable inputs in the production of patient care. Given the quantities and prices of these inputs the average cost per unit of care depends on input productivity and the efficiency with which these inputs are used. In other words, hospital cost analysis is logically secondary to production function analysis as a way of analyzing efficiency in the production of hospital services. The intimate relationship of costs to productivity has been recognized in many recent empirical studies where, partly because of the difficulties encountered in the estimation of production functions, these have been estimated indirectly from the corresponding cost functions.

Since the basic purpose of any incentive reimbursement mechanism is to reward efficiency and discourage inefficiency it would make good sense to go directly to the production side and search for a measure of hospital efficiency. A few attempts have been made in this direction. Blue Cross of Southern California has grouped a sample of 26 hospitals by ownership, location, and size and utilizes labor productivity standards for each hospital department.[117] Another similar attempt is a plan by the Connecticut Hospital Association which uses statistics of production to determine a target budget for nine departments in each participating hospital.[118]

Although hospital cost studies have proliferated in the last few years, little research has been done on the production side. This is perhaps due to the difficulties of definition and empirical measurement of hospital output, or to a lack of satisfactory input data. One noticeable exception is in the 1967 study by M. Feldstein[119] where he estimates various forms of production functions and suggests two measures of productive efficiency derived from a Cobb-Douglas function. In this chapter we will develop and estimate two similar indices of hospital productive efficiency, and we will examine their reimbursement implications.

The difficulties and ambiguities associated with the estimation of production functions, and especially in the case of hospital care, limit the usefulness of productivity indices as empirical measures of hospital efficiency. They can be used, however, as indirect but useful tests of the suitability of the costliness index as a measure of cost performance and efficiency. If high productivity hospitals tend to have low costliness we can consider this further evidence that costliness is a satisfaction measure of efficiency. Moreover, if the association between high productivity and low relative costs is weaker than between high productivity and low costliness, this will imply that costliness is a superior measure of cost performance than relative cost. Our measure of output will be that of "adjusted patient care" as developed in Chapter III.

The Concept of Productive Efficiency

Productive efficiency is by no means a clear-cut concept. As Hall and Winsten[120] have pointed out, the appropriate definition adopted depends heavily on the use to which the various measures of efficiency are to be put. In general, however, what we need is a measure which summarizes efficiency differences among various firms and ranks these firms according to some criterion based on these differences.

Such a concept of efficiency has been proposed by M. J. Farrell.[121] His measure of "overall efficiency" is composed of two parts. The first he calls "technical efficiency", and it measures the extent to which the appropriate production function is used by a firm as compared with the other firms in the industry. The second is called "price efficiency", and it relates to the proper choice of input combinations.

These notions of efficiency can be best explained by the use of a simple diagram. In Figure 3 we assume a single output, two inputs X_1 and X_2 and constant returns to scale. The production function can then be completely described by isoquant QQ". If PP' is the input isocost line, the optimum input combination is given at point A. It is quite likely, however, that a firm produces inside its

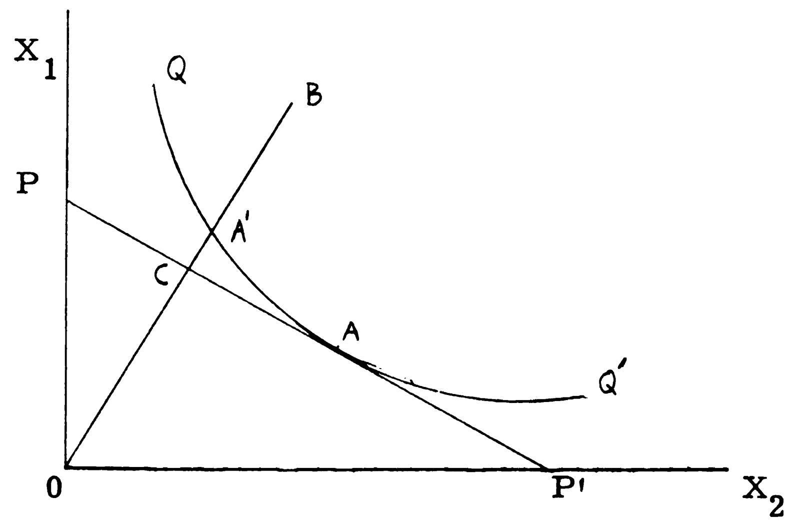


FIGURE 3. --Productive Efficiency

production frontier, namely, at point B. The ratio $\frac{OA'}{OB}$ measures technical efficiency or the extent to which the same output could be produced with fewer inputs. Alternatively, the ratio $\frac{OC}{OA'}$ measures price efficiency or the fraction of costs for which output could be produced with combination A instead of B. Overall efficiency is, thus, defined by Farrell as:

$$\frac{OA'}{OB} \cdot \frac{OC}{OA'} = \frac{OC}{OB}$$

Farrell estimates the actual values for the two measures of efficiency by constructing a "best practice" isoquant. Because this method uses only a small fraction of the total number of observations and because the estimation methods required are very complicated, we will adopt a variant of the approach suggested by M. Feldstein. We will derive the two measures of hospital efficiency separately by estimating the isoquant for a hospital of average productivity. The first measure is a productivity index (P*) which shows the ratio of actual hospital output to the output expected on the average from a given set of inputs. The second measure is an input efficiency index (I*), and it shows the difference in cost due to the fact that the hospital uses different input proportions from the average hospital. This separation of efficiency into two components was first made by Marschak and Andrews[122] who called them "technical" and "economic" disturbances. Of course, throughout this analysis we are assuming that hospitals face the same input prices. We will, however, examine productivity between urban and rural hospitals in an attempt to account for wage differentials.

We can demonstrate the relationships between costliness, productive efficiency and input efficiency with the help of Figure 4. Let QQ' represent the isoquant of a hospital with average productivity, and Q_1Q_1' that of a below average productivity institution.

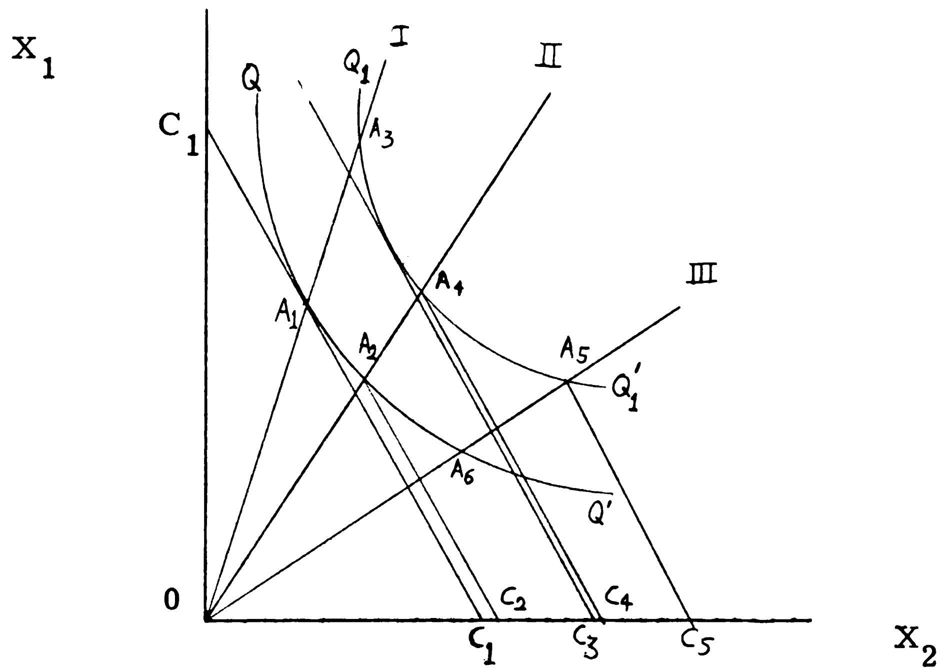


FIGURE 4. --Different Measures of Productive Efficiency

Both isoquants represent the same amount of output, but since the second hospital is less productive Q_1Q_1' lies above QQ'. Line C_1C_1 represents the input price ratio. The least cost input combination for the hospital with average productivity is A_1 with a total cost of C_1.[123] But since this hospital is not necessarily perfectly efficient in its choice of input combinations, it will produce with a different input ratio, say, A_2 or ray II and at a cost of C_2. If the less productive hospital used the same input combinations it would produce at A_4 at a cost of C_4. More than likely, however, it will select a different input ratio, and, if we assume it will choose that given by ray III, it will produce at A_5 at a cost of C_5.

We can now compare the overall efficiency of the less productive hospital at A_5 vis-a-vis that of the average productivity institution at A_2. The first measure of efficiency which will be represented by the productivity index (P_i^*) is a measure of the ratio of actual output to that expected on average from a given set of inputs. Alternatively, in terms of our figure, we can see the difference in productivity as the difference in total costs when the same amount of output is produced at the input ratios used by the average productivity hospital with different amounts of inputs as shown by A_2 and A_4. The productivity index for the less productive hospital is:

$$(1) \qquad P_i^* = \frac{C_2}{C_4}$$

The second measure of efficiency, or input efficiency index I_i^*, is shown by the difference in costs due only to the use of different input combinations. From the figure we see that:

$$(2) \qquad I_i^* = \frac{C_4}{C_5}$$

Finally, costliness as defined in Chapter IV essentially represents the ratio between actual cost and the cost of the average hospital for the same amount of output. We can therefore write:

$$(3) \qquad C_i^* = \frac{C_5}{C_2}$$

From (1), (2), and (3) we can easily derive the identity:

$$I_i^* = [C_i^* P_i^*]^{-1} \tag{4}$$

which will be used later to calculate the input efficiency index. Our first task, however, is to estimate the productivity index P_i^*.

The equation giving the amount of output that a hospital produces with a given set of inputs is the production function:

$$Y_i = f(X_i, \epsilon_i) \tag{5}$$

where Y_i is output, the X_i's are the physical amounts of inputs used and ϵ_i is a random error term implying that output varies among hospitals for the same amounts of inputs.

Since we have defined productivity as the ratio of actual output to that expected on the basis of the inputs used we can estimate the productivity index in a convenient way suggested by M. Feldstein. By estimating a specific form of the production function we can obtain a set of estimated values Y_i, which show the amount of output which would be expected from each hospital on the basis of its inputs if it were of "average" productivity. The productivity index (P*) could then be calculated as:

$$P^* = \frac{Y_i^*}{\hat{Y}_i^*} \tag{6}$$

In the next section we will estimate a Cobb-Douglas function of the form

(7) $$Y_i^* = A' \prod_{s=1}^{k} X_{is}^{a_s} \epsilon_i$$

and we will discuss the theoretical and statistical problems involved in the estimation of hospital production functions. According to our definition of hospital productivity we see from (6) and (7) that a convenient measure of P_i^* is:

(8) $$P_i^* = \hat{e}_i + 1$$

where $\hat{e}_i$ is the estimated residual from the production function regression. This method, of course, depends on an assumption of "neutral" productivity differences among hospitals which means that the output elasticities of the various inputs (α_s) are the same for all hospitals.[124]

Since the production function is estimated in its logarithmic form where the terms enter additively it is easy to show that the hospital with average productivity would have a P_i^* with the value of one.[125] All hospitals, therefore, with a P_i^* of less than one will be considered of less than average productivity, while a P_i^* greater than one indicates above average productivity.

The Production Function

Production function estimation is always beset with the difficulties of modeling a largely unknown and complex production process in such a way as to make it amenable to empirical estimation with our admittedly limited statistical tools. The statistical production model employed must be by necessity a compromise involving: 1) a sufficiently accurate description of the technical realities of production, 2) certain theoretical requirements imposed by economic theory, 3) statistical properties consistent with the methods of estimation used, and 4) a way of using the parameter estimates to test hypotheses of economic significance.

Depending on the purpose behind the estimation of a production function each one of the above considerations assumes a different weight in indicating the appropriate form of the production model used. If our purpose, for example, is to forecast future output with the maximum amount of reliability, the exact form of the function is very important. If the true form of the function $y= f(x)$ is as shown in Figure 5 and we estimate $y=g(x)$, the forecast error at, say, $x=x_o$ can be rather large as shown by $(y_o - \hat{y}_o)$. If on the other hand our purpose is to test hypotheses concerning returns to scale and input productivities over a wide range of input and output values, the smoothing function $y=g(x)$

may be an adequate approximation of y=f(x). The advantage of such an approximation as we will see is that it can be estimated as a linear statistical model.

It would not be efficient for us to attempt a full review of the literature on the theoretical problems of production function estimation. Suffice it to say that, written almost ten years ago, the definitive review article by A. A. Walters[126] contains no less than 345 references. We will simply give a brief, general discussion of production functions, establish the Cobb-Douglas function as a useful and practical approximation, and point to some of its characteristics and problems of estimation.

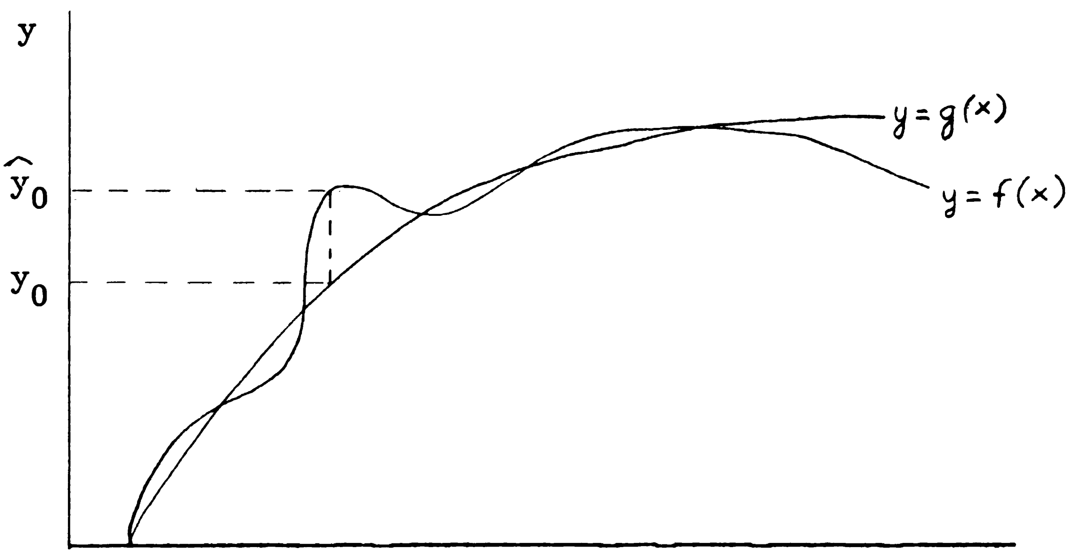


FIGURE 5. -- Two Hypothetical Production Functions

The Production Function and its Appropriate Form

Productive activity consists of the conversion of one or more inputs into a certain output. We can, therefore, imagine the activity as defining a set of points in the input-output space, which we call the production set. The boundary of the production set is defined by an equation relating output to inputs. Let us assume one output, y, and n inputs, $(x_1 \ldots x_n)$. The general form of the production function is then:

$$y = f(x_i) \tag{1}$$

In order to estimate empirically the parameters of $f(x_i)$ we must give it a specific algebraic formulation amenable to estimation. The actual form of the equation must be determined on the basis of the four considerations mentioned earlier.

If the production process is such that inputs are combined in fixed proportions, i.e., where there is only one production technique possible we can postulate a linear model. For n = 2 we can write:

$$y = ax_1 + bx_2 \tag{2}$$

Known as a Leontief function, the above function has been used fairly widely and especially in farm management studies.[127] It implies that the elasticity of substitution among inputs is zero and that there exist constant returns to scale. Furthermore, it does not allow for diminishing marginal returns for inputs, since it implies:

$$\frac{\partial y}{\partial x_1} = a, \quad \frac{\partial y}{\partial x_2} = b$$

i.e., constant marginal products for inputs x_1 and x_2. The linear function therefore is extremely specialized and its fairly wide use is due to the fact it yields rather easily first approximations of the sign and magnitude of certain parameters.[128]

Probably the most widely used function is the Cobb-Douglas (CD). In its general form it can be written as:

(3) $$y = Ax_1^{\alpha_1} x_2^{\alpha_2} \ldots . x_n^{\alpha_n}$$

The CD function allows for perfect substitutability among inputs. The elasticity of substitution is constant at all levels of output and equal to unity. If the function is estimated in its logarithmic form the regression yields direct estimates of the output elasticities $\alpha_1 \ldots \alpha_n$ of the outputs $x_1 \ldots x_n$. The scalar A is a technological constant.

The function is homogeneous of degree $\sum \alpha_i$. If $\sum \alpha_i$ <=> 1, there are increasing, constant, or decreasing returns to scale. Finally, the CD function allows for diminishing marginal physical products for the inputs. Specifically, if $\alpha_i < 1$, then

$$\frac{\partial^2 y}{\partial x_i^2} = \alpha_i(\alpha_i - 1)\ y x_i^{-2}$$

thus, the rate of change of the marginal product of x_i is negative.

There have been many criticisms of the CD function, which however seems to survive. A good exposition of the development of the CD function as well as of many of the criticisms leveled against it can be found in Heady and Dillion[129] and Walters.[130] The distinction between interfirm and intrafirm functions was made by Reder,[131] who showed the conceptual differences between the CD function and the theoretical production function. The empirical importance of the distinctions between the two types of functions is that when observations are beyond the geometric mean for a given input, the marginal value product of that input is likely to be overestimated. If observations on the other hand are below the geometric mean, the marginal product may be underestimated. Biased estimates may lead to erroneous conclusions concerning input use and optimal factor proportions.[132] Bronfenbrenner[133] showed that under competitive conditions the results obtained from interfirm observations should be the same as those derived from intrafirm data. Unwilling to base their estimates on the assumption of perfect competition, Mundlak[134] and Hoch[135] have estimated intrafirm functions from interfirm observations with a method which requires at least two observations on each intrafirm function.

The criticism has also been advanced[136] that the CD function does not allow for zero marginal product. Since the marginal product of input i,

$$\frac{\partial y}{\partial x_i} = \alpha_i \frac{y}{x_i} ,$$

is a decreasing but positive function of x_i, this implies that a firm with fixed capital can keep increasing output by increasing infinitely the amount of labor used. This objection to the theoretical implications of the CD function is, again, of little practical importance, because it is quite unlikely that a firm would hire any inputs to the points where zero or negative marginal products set in. Most researchers have found that the goodness of fit usually displayed by the CD function outweighs its restricted theoretical properties.[137] There are, however, certain practical problems which must be dealt with. For example, the fact that the function is not defined when inputs are entered at zero levels[138] places a limit to the extent to which inputs can be broken down into different categories. As a result we must restrict ourselves to inputs common to all firms in the sample.[139]

The theoretical objections to the Cobb-Douglas function have led to a search for other functional forms suitable to production function analysis. The most celebrated of the alternatives is the Constant Elasticity of Substitution (CES) function proposed by Arrow, Chenery, Minhas, and Solow.[140] The basic change introduced by the CES function was to allow for the elasticity of substitution to be constant at a value other than unity. The many problems involved in the estimation of the CES function make it of dubious value for our particular problem. The advantage of estimating a constant (but not unity) elasticity of substitution, (σ), is further diminished by the fact that it is assumed by the CES to be the same for all pairs of inputs. In the ACMS study the value of σ was not significantly different from one. In view of this and other evidence, whether $\sigma = 1$ is secondary in importance to the fact that both the CD and the CES function assume it to be the same for all pairs of inputs. Unfortunately, the difficulties in estimating a Variable Elasticity of Substitution (VES) function are far greater than those encountered in the estimation of CES functions.

Some other non-linear production functions such as the Spillman function, quadratic and square root functions are analyzed in Heady and Dillon.[141] These functions have certain desirable properties, but they are suited more to farm production situations.

In production processes involving more than two inputs, as in our case, the number of degrees of freedom required for estimation would be too large to make these functions useful.

Despite the many criticisms, the Cobb-Douglas function performs adequately when judged according to the four criteria listed in the beginning of this section. The goodness of fit has been often cited as one of its many advantages. Its computational simplicity and the economy in terms of degrees of freedom are important compensation for the drawback of unitary elasticity of substitution. Although a non-linear function, it can be estimated by linear statistical techniques, and the parameter estimates are readily interpreted in terms of concepts of economic interest. For all these reasons we decided to use the Cobb-Douglas in our estimation of hospital productivity. Before we go into the actual estimation, however, we must set up a statistical model in order to determine the appropriate estimation method.

The Statistical Model

The simple statistical model is composed of a stochastic Cobb-Douglas production function:

(1) $$Y_i = A \prod_{j=1}^{k} X_{ji}^{\alpha_j} e^{u_i}$$

The meaning of a stochastic production function is that a variety of unanticipated factors influence the level of output even when the levels of inputs are unchanged. In the case of hospital production of patient care a stochastic production function is particularly meaningful. Unexpected admissions, complications leading to stays longer than expected, epidemics and other factors introduce a substantial random element in the hospital output. This random element is represented by the disturbance term u_i which is assumed to be normally distributed with zero mean and finite variance. At this moment we will make no assumptions on the variability of the error variance.

Many studies have estimated production function parameters by Ordinary Least Squares techniques. It is fairly simple to show that single equation, least squares estimates of the parameters of (1) will be subject to simultaneous equation bias. Let us expand the model by including a set of input demand equations.

(2) $$X_{ji} = \alpha j \frac{Y_i c_i}{R_j p_j} e^{v_{ji}}$$

These equations show the levels of inputs that a hospital would hire if it pursued a policy of profit maximization, output maximization, or cost minimization. The parameters c_i and p_j

represent the marginal cost of output and the price of each input respectively. Since we showed in Chapter II that we cannot be certain of the actual objective function of the hospital we do not wish to base the input demand equations on any restrictive assumptions concerning hospital behavior. Hoch[142] has shown that with the introduction of R_j in the input demand equations, the behavioral assumption becomes a hypothesis to be tested and not an a priori statement. Thus, R_j represents an "average" for all hospitals deviation from optimality (profit maximization, cost minimization, etc.) due to the various constraints on the hospital's economic behavior. Individual hospital variations around R_j are assumed as part of v_{ji}. More specifically, the disturbance v_{ji} is introduced to allow for random, non-systematic errors on the part of individual hospital managers[143] in their attempts to adjust inputs so as to satisfy the necessary conditions for cost minimization.[144]

The simultaneous equations problem arises from the fact that in equation (2), X_{ji} depends on the actual level of output Y_i which includes the disturbance u_i. This, in turn, implies that the independent variables in the production function are correlated with the disturbance.[145] Since a necessary assumption for unbiased parameter estimates when ordinary least squares are used is the assumption of independence of the X_{ji}'s from the disturbance, we must establish this independence or use simultaneous equations techniques. An approach particularly suited to hospital production is suggested by Hoch[146] and analyzed by Kmenta et al.[147]

If we can assume that hospital managers minimize costs with respect to expected output and not actual output then the error of equation (1) does not enter into (2) and, therefore, simultaneous equation bias does not arise. If hospitals, in other words, determine input demand by differentiating anticipated output with respect to the X_{ji}, it can be shown that the observed values of the X_{ji}'s are not functions of u_i. This assumption is particularly well suited to hospital production conditions. We showed in Chapter II that hospital managers staff the hospital on the basis of a certain expected occupancy rate. The level of actual output, in fact, is beyond their control since admissions and length of stay depend largely on the decisions of the physician staff. This division of control over input and output levels, therefore, lends particular validity to the assumption that inputs are hired on the basis of expected output. The importance of this assumption, of course, is that we can estimate the parameters of the production function by single equation least squares on the logarithms of the variables. This will be the subject of the next section.

Estimation of the Production Function

We now come to the actual estimation of the production function. A Cobb-Douglas function was estimated for a sample of 94 Michigan hospitals. The output data used were the estimated values of adjusted patient care as explained in Chapter III. The specification of the inputs used in the production of patient care, however, presented us with certain problems.

Although the theory of production function estimation requires the specification of inputs in physical terms this is not always possible. The AHA data contain information on the number of beds, nurses,[148] interns, residents, and all other personnel. Two other variables, however, namely supplies and hospital assets, can only be included in money terms, which forces us to use a mixed specification of inputs both in physical and in value terms.

Since hospital production function studies are very scarce we have very little experience from which to draw. In one of the few such studies M. Feldstein[149] suggests that labor inputs should be aggregated by wage rates in order to achieve greater comparability among hospitals. This approach rests on the assumption that wage rates for different grades of nurses and other hospital personnel are fairly uniform among all the hospitals in the sample. Although this may be true in the case of the British Health Service with which Feldstein's study is concerned, we have shown[150] that significant wage differentials exist among the various regions of the state of Michigan. We have decided therefore to express the labor units in physical terms.

The categories of labor inputs for which AHA data are available are (1) Registered Nurses, (2) Licensed Practical Nurses, and (3) a general category called All Other Personnel which does not include physicians, administrators, interns, and residents. Data on both full-time and part-time personnel are included which we converted into full-time equivalents by assuming a conversion factor of two.[151] The use of physical units does not cause any problems in the case of RNs and LPNs since the groupings are fairly homogeneous, but the third category contains a number of occupations, the mix of which may vary among hospitals. Although such variability will introduce some bias into the individual coefficients of the production function, we believe that the bias resulting from the exclusion of this variable would be greater.

Unfortunately, data on a very important input, namely the number of physicians providing care in each institution, are unavailable. An attempt to obtain such data via a questionnaire to the sample hospitals was not successful. Although the response rate was satisfactory the double counting resulting from multiple appointments,[152] as well as the varied methods of reporting by the hospitals made the data unusable. The AHA data do include information on the numbers of interns and residents for twenty-one sample hospitals. Although interns and residents provide a significant part of patient care in these hospitals we cannot use them as a separate category of labor inputs. The specification of the Cobb-Douglas function does not permit entering any inputs at zero levels which would be the case for the hospitals which do not use interns and residents. If we decided to estimate two different production functions, on the other hand, the two sets of coefficients would not be comparable, and we would still face the problem of incomplete specification in the group of twenty-one hospitals since interns and residents do not provide all patient care.

Because of the lack of data we were forced to treat physicians as managerial rather than technical inputs. They are assumed, therefore, to determine the form of the production function as they decide on the way other inputs are used, but not to enter the production process as separate inputs.

Concerning the various capital inputs used in the hospital production function we faced even bigger problems than this variable usually causes in production function research.[153] Information on capital is at best scattered, and the only complete set of data was on the numbers of x-ray and cobalt treatment units available in each hospital and included in the Michigan Hospital Survey. Since these two inputs alone hardly express the total number of capital services, we have decided to use the total dollar value of hospital assets as an instrumental variable for capital. In order to do this we must assume that hospital assets are positively correlated with capital and that they are uncorrelated with the error term in the production function equation.[154] Both these assumptions seem fairly reasonable.

There are two more inputs that enter the hospital production process directly, namely beds and supplies. The number of hospital beds represents another capital variable and it obviously determines the amount of patient care a hospital can produce. Finally, the supplies variable includes certain drugs and dressings, x-ray films, etc., and it is expressed in money terms.

Empirical Results

We estimated a Cobb-Douglas function of the form:

$$Y_i^* = A \prod_{s=1}^{6} X_{is}^{\alpha_s} \epsilon_i$$

where

Y_i^* = the number of units of "adjusted patient care"

X_1 = beds

X_2 = registered nurses

X_3 = licensed practical nurses

X_4 = all other personnel

X_5 = assets (in thousands of dollars)

X_6 = supplies (in thousands of dollars)

The estimated coefficients were as follows:

Regression Number One

	Coefficients	Std. Errors of Coefficients	t -Values
Constant	4.889		
X_1	0.764	0.083	9.2
X_2	0.139	0.016	8.6
X_3	-0.016	0.021	0.8
X_4	0.219	0.007	31.2
X_5	0.011	0.026	0.4
X_6	0.000	0.003	0.3
R^2	0.9863		
F -Statistic	18.6		

The negative sign of X_3 runs contrary to our theoretical expectations of positive input elasticities. Since the variable also appears to be not statistically significant we looked for the possibility of multicollinearity between X_2 and X_3. Since the simple correlation coefficient between the two variables was 0.874 we

decided to lump X_2 and X_3 together and consider the two types of nurses as one hospital input. The new and final set of inputs therefore becomes:

X_1 = number of beds

X_2 = nurses (full time equivalents)

X_3 = all other personnel (full time equivalents)

X_4 = supplies (in thousands of dollars)

X_5 = assets (in thousands of dollars)

The estimated coefficients were as follows (the function was estimated in its logarithmic form):

Regression Number Two

	Coefficients	Std. Errors of Coefficients	t'-Values
Constant	5.020		
X_1	0.595	0.097	6.10
X_2	0.180	0.049	3.67
X_3	0.212	0.067	3.14
X_4	0.001	0.004	0.25
X_5	0.039	0.015	2.60
R^2	0.9805		
F-Statistic	20.4		

Variables X_1, X_2, X_3, X_5 are significant at the 99 percent level. X_4 is not significant at any level.

Regression number two is used to derive the residuals which form the productivity index as explained previously in this chapter.

Productivity and Costs

We now come to our main objective which is to assess the extent to which the costliness index is superior to relative cost as a measure of efficiency differences among hospitals. We examined hospitals which showed above or below average productivity to see whether they also showed low or high relative costs and costliness. There were 44 hospitals with above average productivity and 50 in the below average category. In 62 cases we found agreement between the productivity index and relative cost, in other words, either high productivity and low costs or low productivity and high costs. The agreement between productivity and costliness was higher. In 72 cases both indexes had the same implications for hospital efficiency. Binomial tests in both cases showed an extremely small probability that these results were due to chance.[155]

Although both Cr and C* seem to reflect productivity differences among most hospitals, the higher rate of success between P* and C* is significant. It seems, in other words, that the casemix and length of stay adjustments increase the extent to which our measure of hospital costs reflects differences in efficiency among

TABLE 18. --Cost and Productivity Performance According to Cr, Pr, C*, P* for 22 Hospitals

Hospital Number	Low Cr - High C*					Hospital Number	High Cr - low C*				
	Location	Cr	Pr	C*	P*		Location	Cr	Pr	C*	P*
1	0	-	-	+	-	2	2	+	-	-	+
3	0	-	-	+	-	23	1	+	-	-	+
4	0	-	+	+	-	33	3	+	-	-	+
6	3	-	+	+	+	52	0	+	-	-	-
13	4	-	+	+	+	55	0	+	+	-	-
15	0	-	-	+	-	58	6	+	+	-	+
26	0	-	+	+	-	81	2	+	+	-	+
43	0	-	+	+	-	82	2	+	+	-	+
46	4	-	+	+	+	93	6	+	+	-	+
62	0	-	-	+	-						
86	0	-	+	+	-						
88	0	-	-	+	-						
89	0	-	+	+	+						

hospitals. Furthermore, a closer look at the data revealed certain interesting relationships which lend additional validity to the costliness index.

We took a look at the 13 hospitals which showed low relative costs and high costliness and the nine institutions with high costs and low costliness. In Table 18 we show hospital performance according to Cr, Pr,[156] C*, and P*. A plus sign signifies above average costs or productivity while a minus sign indicates below average values for these variables. In the previous chapter we suggested that some of the discrepancies between Cr and C* were due to differences in efficiency. The results shown in Table 18 seem to bear this out to a considerable extent. We hypothesized that the low relative costs in the first thirteen hospitals were due to reasons other than efficiency and that, in fact, efficiency was low as shown by the higher than average costliness. For nine out of the thirteen hospitals this appears to be the case as indicated by the below average productivity. The fact that costliness represents efficiency better than relative cost is seen more clearly in the nine hospitals with the low costliness values where seven out of nine actually show above average productivity. We included Pr in the Table above, incidentally, in order to show that failure to adjust for casemix in the productivity index will obscure the relationship between costliness and productivity.

We can also look at the relationship between productivity and costliness in a different way. If C* is a better measure of efficiency than Cr, we would expect that in the cases where high (low) relative costs are associated with high (low) productivity, C* would correct some of the apparent inconsistency. As a matter of fact, in the 32 hospitals in this category above (below) average productivity was reflected in low (high) costliness in fifteen cases.

One final observation was made when we classified hospitals by location. As we saw earlier the costliness index indicated that the incidence of inefficiency was higher in rural hospitals than in urban or metropolitan institutions. In Table 18 we see that of the thirteen hospitals which appear as inefficient on the basis of costliness the nine which actually show below average productivity are all rural hospitals. One the other hand, of the nine low costliness institutions, the seven which also showed high productivity were urban or metropolitan hospitals.

We now have additional evidence to support our previous claim that reimbursement on the basis of costliness will be more equitable and will include more realistic efficiency incentives than average cost reimbursement. If hospitals are paid on the basis of a target average cost the first thirteen hospitals in Table 18 would receive some sort of financial reward. In nine of these cases such reimbursement is not justified on the basis of productivity. The costliness index would, of course, penalize inefficiency in these nine hospitals, but some care should be given to

the four institutions where penalties do not seem warranted by the above average productivity. The advantages of the costliness index are much more pronounced in the nine hospitals where its use would prevent actually efficient hospitals from being penalized for treating expensive casemixes.

Input Efficiency and Costs

As we mentioned earlier the input efficiency index I* measures the extent to which hospitals use efficient combinations of inputs in the production of patient care. Our definition of I* allows us to calculate it indirectly by using our estimated costliness and productivity values and by using the identity (4) earlier in this chapter.

The notion of a certain degree of inefficiency in the hospital industry finds a certain amount of support in our results. There are 53 hospitals with below average input efficiency while only 41 hospitals combine inputs more efficiently than the average institution. Further analysis of the I* values, however, did not prove very enlightening.

Perhaps surprisingly there seems to be a very weak association between productivity and input efficiency. The estimated simple correlation coefficient is only 0.128. We might have expected that hospitals which combine inputs inefficiently,

i.e., in other than the "optimum" proportions, would also show below average productivity for these inputs. This may, in fact, be the case in other industries where input and output decisions are made by the same management. Good management would probably hire inputs at or near the optimum proportions and also use them with a high degree of productivity. As we saw in Chapter II, however, input and output decisions are made by different agents in the hospital industry. The hospital administrator can affect input efficiency because he is responsible for budgetary allocations (input combinations). Productivity, on the other hand, is more in the hands of the physician staff which decides on the way inputs are used and the amount of output produced. Whether or not, therefore, inputs are hired in optimum combinations does not guarantee that they will also be used in the most productive way. Further evidence of the lack of association between I* and P* is the fact that the two measures of productive efficiency move in opposite directions in fifty-five hospitals. This again is explained by the fact that most hospital inputs are fixed at least in the short run while productivity may change depending on physician demand for beds or the nature of patient care that a hospital provides at a given point in time.

Perhaps these are also the reasons behind the observed weak relation between input efficiency on the one hand and costliness and relative costs on the other. The simple correlation between I* and Cr was 0.093 while that between I* and C* was -0.074. Although the negative association between I* and C* goes along with our theoretical expectation, both correlations are too weak for any definite statements.

The apparent lack of association between input efficiency and productivity on the one hand and hospital costs on the other should not be taken to mean that cost reductions cannot be achieved by a more efficient allocation of hospital inputs. Although the scope of this study does not allow us to examine the point further, it is possible that both productivity improvements and cost containment can result from changes in input proportions such as between physicians and paramedical personnel or by substituting capital for certain forms of labor. Our results, however, show that cost containment can be best achieved by increases in the productivity of hospital inputs because of the much stronger association of productivity with average costs and especially with costliness.[157]

Chapter Conclusions

Although examination of the input efficiency index proved inconclusive, the productivity index showed that costliness is a superior measure of hospital costs and efficiency than average relative cost. In the majority of the cases where costliness and relative costs have opposite reimbursement implications the use of costliness is justified by the productivity index. We view this as further evidence that hospital costs should be adjusted for casemix and length of stay differences if incentive reimbursement on the basis of costs is to have the correct efficiency incentives without compromising the quality of care.

The final point is of a somewhat normative nature. If our analysis is incorrect and costliness is actually a poor measure of efficiency then reimbursement on the basis of C* will reward some high cost institutions which, nevertheless, appear to have high productivity. It will also penalize a few low cost institutions which, however, also seem to be less productive. On the other hand, if our analysis is correct, the potential danger from ignoring casemix differences is much greater. As we have already shown, in that case not only would the efficiency incentives be dulled, but serious quality deterioration may occur if hospitals with expensive casemixes are penalized.

CHAPTER VII

CONCLUSIONS, IMPLICATIONS, AND POLICY RECOMMENDATIONS

The basic purpose of this thesis has been to develop and examine a measure of hospital costs which is closely related to the efficiency of operation within the hospital. The rationale behind the development of such a measure of costs is found in its possible use as a tool for the determination of incentive reimbursement of hospitals by the various third parties. There is a growing realization that the spectacular increase in the price of hospital care during the last decade is intimately related to the prevailing methods of reimbursement. An examination of the hospital economic behavior revealed that the nonprofit status of the industry, the nature of the hospital product, and the preoccupation of hospital decision-makers with the quality of care provide few incentives for efficient operation. We further showed that the inherent tendencies toward higher costs in the hospital industry are reinforced by the current reimbursement methods which provide payments of full costs or more, thus making productive efficiency essentially a secondary consideration.

Although the hospital industry has been traditionally characterized by lack of competition in the usual economic sense, we believe that the reimbursement system should not buffer hospitals from all the constraints and pressures of a competitive marketplace. More specifically, the method of payments should provide strong incentives for hospitals to behave in an efficient and economical fashion, given certain predefined standards for the quality and scope of their services. It is precisely such a realization that has recently caused great interest in the concept of incentive reimbursement.

The purpose of incentive reimbursement is to make payments to hospitals in a manner which induces greater cost consciousness and stimulates a concern for efficiency. Accordingly, it has been suggested that high cost institutions should be reimbursed at less than full costs and low cost hospitals rewarded with payments in excess of the full cost of the services rendered. The expectation of such a reimbursement scheme is that hospitals with high costs will be forced to reduce their total expenditures in order to insure their viability. While it is also expected that low cost hospitals will use their additional revenues for desired quality improvements or expansion of services, it is believed that the rates of cost increases for the industry as a whole will be lower than the ones prevailing during recent years.

Since the main objective of incentive reimbursement is to encourage the efficient operation of hospitals, it follows that the standard of reimbursement used must be closely related to efficiency. The thesis shows that the various incentive reimbursement plans proposed or currently in operation use estimates of the average cost per case or per patient day as the basis for determining reimbursement rewards or penalties. We also showed, however, that these cost estimates are greatly affected by a variety of factors, many of which are not related to efficiency. Two of the most important variables in this category are the hospital's patient mix (or casemix) and the average length of stay. Our analysis showed that both variables affect the qualitative and quantitative aspects of hospital output. The same factors, therefore, also affect a hospital's average cost per unit of care if the actual amount of patient care produced is used to represent hospital output.

The main objective of this thesis is to derive a measure of hospital costs which bears a closer relationship to the efficiency of operation within a given hospital. We therefore had to adjust average costs per case for certain factors which cause hospital costs to vary but are unrelated to efficiency. Some of these variables such as regional wage differentials, differences in facilities and services, and the existence of teaching programs were adjusted for in an indirect way by classifying hospitals according to the degree of urbanization of the area they serve. There were two factors, however, for which direct adjustments were made, namely, casemix and length of stay.

We adjusted average costs per case for casemix differences by disaggregating hospital output into six types of cases: medical-surgical cases, patients 65 years of age or older, obstetrics, pediatrics, psychiatric cases, and outpatient care. Cost weights for each case-type were estimated from a regression of average cost per case on the proportions of cases in each case-type. A "costliness" index was then calculated from the ratio of actual costs per case to the costs expected from the regression equation and weighted by the hospital's casemix composition. As a result, the index reveals cost differentials among hospitals, but it adjusts for that part of the cost differences which is due to differences in casemix. The reason behind the casemix adjustment is that hospitals which show high average costs per case not because they are inefficient but because they tend to treat patients in expensive categories of care should not be penalized by the reimbursement mechanism.

The rationale behind the length of stay adjustment lies in the fact that differences in the length of stay reflect differences in the actual amount of patient care produced. A measure of hospital output is developed in the third chapter which transforms hospital cases into units of "adjusted patient care". This is done by multiplying the number of cases in each case-type by the natural logarithm of the average length of stay for each case-type in each hospital. The costliness index is then adjusted for length of stay differences through multiplication by the ratio of the adjusted patient care which a hospital would show if its length of stay in each case-type were the same as the average for all hospitals, to the amount of adjusted patient care actually produced by the hospital. The final costliness index thus adjusts for both casemix and length of stay differences among hospitals.

The costliness index values were calculated for a sample of 94 Michigan short-term general hospitals and compared with observed average costs per case for their reimbursement implications. We found that for roughly a quarter of the sample hospitals the two measures of cost had the opposite implications. That is, hospitals which would be rewarded if incentive payments were based on their cost performance as measured by average costs per case, would be penalized if costliness were used as the reimbursement standard, and vice versa. Moreover, the differences in the calculated reimbursement amounts under the two payment methods were considerable. We then adjusted the observed average costs and the costliness index for differences in factors such as regional wage differentials, the existence of teaching programs, etc. Again costliness and observed average costs showed significant differences in their reimbursement implications for the hospitals in the sample.

From the derivation of the costliness index we have theoretical reasons to believe that costliness is a superior measure of hospital efficiency than the observed average costs per case or per patient day. If this is indeed the case, then failure to use the costliness index[158] in an incentive reimbursement plan would provide some inefficient hospitals with additional funds while penalizing certain other efficient institutions which treat expensive or complex cases. The result would be increased inefficiency in the hospital industry and/or reduction in the quality of care.

Because of the significance of these results we decided to provide a test, admittedly rough, of the relationship between the two measures of costs, on the one hand, and efficiency on the other. We therefore estimated two measures of hospital efficiency, namely, a productivity index and an input efficiency index. The two indices were calculated from the residuals of an estimated Cobb-Douglas production function, where adjusted patient care was used as the measure of output. Our results show that costliness bears a relationship to productivity which is considerably higher than that between productivity and average cost per case. Unfortunately,

the results from an analysis of the input efficiency index were inconclusive. The productivity index analysis, however, confirms our theoretical expectations that costliness is superior to unadjusted average cost per case as a measure of hospital efficiency. Based on that result, and for the reasons given elsewhere in this thesis, we believe that costliness is a suitable measure of hospital costs for the purposes of incentive reimbursement.

Implications and Recommendations

The potential effectiveness of the costliness approach depends upon the final form it assumes and the regulatory milieu within which it operates. As presented, the costliness index is not a reimbursement system in itself but rather an indicator of hospital efficiency; there are a variety of ways in which this indicator could be utilized. It could, for example, be used as a guide for hospitals, planning agencies, and the economic stabilization program. It could be employed as an instrument for insurance regulation, particularly regarding the relationship between hospitals and Blue Cross. The information obtained in tabulating the costliness index could be released to the public in an effort to bring pressure to bear on inefficient hospitals. Or finally, the index could be used as the basis for hospital reimbursement. Clearly, the last alternative is the most likely to yield significant results.

Assuming that the costliness reimbursement technique is correctly specified and properly administered it has the potential to promote more efficient operation within individual hospitals as well as a more rational allocation of resources across the whole system of hospitals.

Any mandatory system of incentive reimbursement will, of course, require the establishment of incentives and safeguards designed to mitigate the financial "shock" of implementation. One way to deal with this problem is suggested by the incentive reimbursement scheme used by Blue Cross of Western Pennsylvania.[159] Adopted in 1966, this program established nine hospital groups based on each institution's location and teaching program, and called for reimbursement on the basis of a 10 percent range about the group mean. That is, if the average cost per patient day for a given institution were above 10 percent of its group mean, the hospital would be paid only the mean plus 10 percent for each covered patient day. Instead of using an average cost base as in the Pennsylvania case, the costliness formula could be substituted for reimbursement purposes. Thus, a hospital which operated above 10 percent of the mean efficiency as defined by the index would be reimbursed the mean plus 10 percent. The opposite would apply in the case of the inefficient hospital.

An alternative approach would be to use nonsymetrical incentive payments wherein the efficient hospital is provided full incentive payments (based upon its costliness index), but the inefficient hospital is reimbursed on a sliding scale of penalties down to some predetermined minimum, e.g., 95 percent of incurred costs. By setting a floor below which penalty differentials are not imposed, the institution would be protected from the initial disruption involved in having to make massive adjustments in its methods of operation; yet the reimbursement differential could still be large enough to provide incentives for greater efficiency in the future.

This approach to incentive reimbursement could, of course, be quite expensive depending upon the range of relative efficiency levels exhibited by the hospitals in the group. A potential method of avoiding such additional expenses would involve loan financing. Under this system the costliness reimbursor would impose penalties for relative inefficiency in the form of loans or deferred payments to be offset against net revenues accrued in subsequent periods from any increase in efficiency. If, in the process, it is found that certain hospitals have not shown sufficient improvement, further loan payments could be reduced or eliminated.

A final way to mitigate the shock of implementing costliness reimbursement would be to incorporate a time dimension in the reimbursement mechanism. In this case the inefficient hospital which demonstrates progress toward efficiency over time could be reimbursed at a progressive rate. That is, an institution whose rate of improvement in operational efficiency is above average over a certain period, yet is still classified as inefficient, would receive a proportionately higher payment than is indicated by its costliness index. On the other hand, for the efficient hospital sliding toward inefficiency, a regressive rate of reimbursement would be applied. Implicit in this concept is the recognition that costliness reimbursement can (and perhaps should) reflect the rate of increase in efficiency as well as the level of efficiency.

Apart from the actual implementation of costliness reimbursement, there are a number of implications for the hospital industry associated with this approach to incentive reimbursement. First, as costliness reimbursement progresses over time, it can be expected that the range of efficient and inefficient hospitals will narrow and cluster around the mean of the costliness index. This follows naturally from incentives which would drive grossly inefficient hospitals out of business or force them to emulate more efficient institutions. In addition, since hospitals with highly efficient costliness ratings would be rewarded with excess funds which could be expended for "unnecessary" (i.e., economically

nonsupportive) services or invested in luxury or experimental capital to improve existing services, we might expect their costliness ratings to fluctuate or even deteriorate to some extent.

Second, a hospital's relative costliness is not independent of the rest of the industry. Although an institution may succeed in lowering its costs absolutely, it will not improve its position on the costliness scale unless it manages to operate more efficiently than the hospital of "average" efficiency. This is likely to induce a healthy sense of competitive cost consciousness within the hospital system. One word of caution is in order, however. It is possible under costliness reimbursement that a hospital with a costliness index, say, of 1.05 in the initial period might improve in absolute efficiency each year but still remain at the same relative position over time. Clearly a hospital which is 5 percent less efficient than the average is not grossly inefficient nor totally lacking in effective management. However, it is equally clear that such a hospital cannot operate year after year on revenues which fail to cover costs. As mentioned previously, the most effective way to overcome this situation is to set a full cost reimbursement range around the mean such that hospitals in this situation are not penalized.

Third, it can be expected that individual hospitals will react in a variety of ways to any losses or gains obtained through costliness reimbursement. It may even prove possible for some very inefficient hospitals to escape the consequences of their poor performance although penalties are imposed. This would be the case, for example, in communities which can mount successful fund raising drives for hospital support. A similar situation might arise if the hospital could tap large philanthropic contributions. In general, however, we would expect the inefficient hospital to attempt first to subsidize any losses incurred through the costliness reimbursement by increasing the level of charges to private-pay patients and patients covered by indemnity insurance. This tactic would be impossible if all reimbursors used the costliness approach, or even if the costliness reimbursor's policyholders represented a substantial portion of the hospital's clientele. But if the costliness reimbursed population is small relative to the hospital's total patient load, then a minor increase in the level of charges might be sufficient to overcome any losses; and in this case little impact should be expected from the implementation of any type of incentive reimbursement.

Aside from the possibility that some hospitals might be able to escape the intended effects of costliness reimbursement altogether, the approach is likely to induce changes in the services provided by individual hospitals. For example, if the degree of inefficiency is inversely related to the degree of specialization of services available within an institution, then costliness reimbursement might well spark a movement toward greater hospital

specialization and/or the establishment of pooling arrangements among hospitals. Significant economies could be achieved, for example, if underutilized, high fixed cost specialties were eliminated. In most areas the supply of such services could be made available from one or two hospitals. Specialization along these lines would not only relieve the inefficient hospitals of a heavy financial burden, but would also benefit the institution which assumes responsibility for providing specialized services. Because of the costliness incentives we would expect efficient institutions to assume greater responsibility for those patient groups and/or services which require technologically sophisticated equipment and skills. Furthermore, costliness reimbursement would provide efficient hospitals with the surplus revenue needed for expansion and improvement. Conversely, it could be expected that inefficient hospitals would increasingly accommodate those patients needing more routine services, since by minimizing the range and complexity of services available, these institutions could protect themselves from unanticipated shifts in demand and any consequent decline in reimbursement revenue.

Finally, it should be noted that in extreme cases of inefficiency, an institution may be forced to stop providing hospital care altogether. This need not be a catastrophic occurrence for the community if the remaining hospitals in the area are making

adjustments in their mode of operations as suggested above. Furthermore, the prospect of an inefficient hospital having to shut its doors may dramatize the issue of the economic and social criteria which must govern the existence of an institution that is woefully inefficient and unable to make the necessary structural and administrative changes required under incentive reimbursement.

APPENDICES

APPENDIX A

ESTIMATION OF THE COST WEIGHTS FOR THE SIX CASE-TYPES

ESTIMATION OF THE COST WEIGHTS FOR THE SIX CASE-TYPES

Our solution to the problem of hospital output heterogeneity is based upon the assumption that an important part of casemix variation among hospitals can be captured by separating cases into the six broad case-types of medical-surgical (MS), obstetrics (OB), pediatric (P), geriatrics (G), psychiatric (PS) cases, and outpatient (OUT) visits. The hypothesis to be tested is that the average costs per case in each case-type differ significantly from one another.[160]

In estimating average cost weights for each one of the six case-types that compose hospital output, we could estimate the linear total cost function:

$$(1) \qquad C_i = \alpha_o + \sum_{j=1}^{6} \alpha_j X_{ji} + u_i$$

where C_i is the total cost of providing patient care, X_{ji} is the number of cases of type j treated by hospital i, α_o is a constant term accounting for any fixed cost elements included in C_i, α_j is the marginal cost of a case in type j, and u_i is a random error. The weights assigned to each case-type (w_j) are estimated from the average costs (c_j)

$$w_j = c_j = \frac{\alpha_o}{\Sigma X_{ji}} + \alpha_j$$

Unfortunately, previous experience with total cost regressions has proved highly unsatisfactory.[161] Our estimated equation for C_h yielded the following coefficients

$$(1)\quad \widehat{C}_i = \underset{}{4174} + \underset{(98.3)}{772.3\ MS} + \underset{(147.9)}{31.2\ OB} - \underset{(203.3)}{303.8\ P} + \underset{(278.5)}{875.7\ G} + \underset{(315.6)}{1567.6\ PS} + \underset{(147.4)}{616.6\ OUT} \qquad R^2 = .9344$$

which were not very satisfactory. The coefficient of obstetrics is low and insignificant, while the negative coefficient of pediatrics offends the theoretical expectation of nonnegative marginal costs. There are three possible explanations for the poor estimates in this total cost regression: heteroskedasticity, multicollinearity, and incorrect specification of the cost function.

Problems of Estimation

Heteroskedasticity

One of the necessary conditions for efficiency in least squares estimation of a linear equation is that of constant error variance or homoskedasticity. When this assumption is not satisfied, ordinary least squares estimates are unbiased and consistent but inefficient, and the estimated standard errors of the coefficients are biased. Because of the large size differences among hospitals there is a strong a priori likelihood that error variances are correlated with output.[162] The primary reason for estimating the parameters is to test the hypothesis that average costs are different for each case-type. We must, therefore, use a heteroskedastic model, since these tests utilize the estimates of the parameter variances. If we divide both sides of the total cost function by the total number of "adjusted"[163] hospital cases we can estimate the average cost function:

$$C_i/X_i = \beta_0 + \sum_{j=1}^{6} \beta_j p_{ji} + v_i \quad (2)$$

where p_{ji} is the proportion of total cases in type j, $\beta_0 = \frac{\alpha_o}{x_i}$, $\beta_j = \alpha_j$ and $v_i = u_i/x_i$. We assume that v_i is distributed normally with zero mean and constant variance. This model is also used by M. Feldstein, who claims that the use of an average cost function reduces the likelihood of heteroskedasticity. Let us now examine model (2) as a solution to the serious degree of multicollinearity apparent in model (1).

Multicollinearity

When some or all of the independent variables in a regression are highly correlated, it is difficult and often impossible to isolate separate influences and to obtain a reasonably precise estimate of the relative effects of each variable. In other words, the parameter estimates are biased and inefficient. As shown in Table 19, our six output variables are highly correlated. A possible reason is that they are all highly correlated with hospital size.

TABLE 19.--Correlations Among Annual Number of Cases in Each Case-type and Size of Hospital

	Mean	1	2	3	4	5	6	7
1. Med/Surg	3543	1.000	.653	.717	.856	.475	.789	.965
2. OB	1113		1.000	.784	.679	.219	.456	.693
3. Ped.	965			1.000	.671	.452	.543	.753
4. Ger.	1115				1.000	.280	.584	.893
5. Psych.	173					1.000	.582	.540
6. Outp.	30024						1.000	.782
7. Size (in beds)								1.000

The classic solution to the problem of collinearity is to increase the sample size. For obvious reasons this is impossible in our case. Even if it were possible, a large sample would be of little help if a stable underlying structure generates additional collinear data. Another solution is to reduce the number of parameters to be estimated. This will only help if the excluded variable(s) is (are) highly correlated with variables left in the equation.

Moreover, excluding a variable which is theoretically believed to belong in the equation will lead to specification bias and biased parameter estimates.

The best known procedure for reducing the dimension of a set of variables is principal component analysis. The basic problem with this procedure--as with other such statistical methods--is that the estimated regression coefficients make interpretation difficult because the choice of variables that enter the regression is made on statistical grounds (i.e., the explanatory power of variables or combinations of variables),[164] and, therefore, the estimated regression does not follow from any theoretical construct. For example, if hospitals incur costs for the production of treatments in k case-types and if the cost regression includes only k-n explanatory variables, we cannot conclude that the parameter estimates are average or marginal costs for cases in these case-types.

One solution to the collinearity problem is given by the heteroskedastic model (2). In the regression of average cost per case on the proportions of total cases that belong to each case-type, p_{jj}, we see from Table 20 that the collinearity among case proportions is much lower than among the absolute numbers of cases. Casemix proportions were used by Feldstein to estimate equation (2) with 28 case-types, but the effort yielded unsatisfactory results in the form of high standard errors and theoretically unacceptable (negative) coefficients. A regression with 9 case-types[165] gave positive and generally significant parameter estimates with $R^2 = 0.275$, but as we shall see later these coefficients were obtained by the use of an incorrect statistical model.

TABLE 20. --Correlations Among Proportions of Cases in Each Case-type

	Mean Proportion	1	2	3	4	5	6
1. Med/Surg	0.465	1.000	-.508	-.560	.064	-.251	-.242
2. OB	0.125		1.000	.125	-.259	-.090	-.188
3. Ped.	0.114			1.000	-.285	.110	-.160
4. Ger.	0.170				1.000	-.270	-.350
5. Psych.	0.015					1.000	.114
6. Outp.*	0.107						1.000

*Outpatient visits in inpatient case equivalents.

From model (2) we can estimate the average cost per case of type j as:

$$c_j = \beta_0 + \beta_j = w_i$$

using a linear cost function. The linear cost function implies that the average cost of each case-type is constant, or alternatively, that the total cost function is a linear combination of the individual case costs. This is a reasonable assumption since our cost variables do not include any capital costs. Other rationales for the linear cost function are that it is a useful approximation to a nonlinear function and, that in the case of the hospital industry, possible sources of nonlinearity (such as general economies of scale) are not very important.[166]

Since we ideally wish to estimate a variable average cost function, the output weights for each casetype are designed to reflect the relative use of <u>variable</u> inputs only (i.e., total payroll, employment benefits, expenses for supplies, and purchased services). For this reason, the cost function should not include a constant term. The final equation is thus:

$$C_i = \sum_{j=1}^{6} \gamma_j P_{ji} + \varepsilon_i \tag{3}$$

Equation (3) is actually equivalent to the one employed by M. Feldstein who does estimate a constant term. Let us rewrite equation (2):

$$c_i = \beta_0 + \sum_{j=1}^{6} \beta_j p_{ji} + v_i \tag{2}$$

Since the p_{ji} variables are expressed as proportions of a total, we have:

$$\sum_{j=1}^{6} p_{ji} = 1 \tag{4}$$

In order to estimate (2) we must omit one of the p_{ji}'s, the effect of which on c_i will be captured by the constant term.[167] In other words, we estimate

(3.1) $$c_i = \beta_0 + \sum_{j=1}^{5} \beta_j p_{ji} + v_i$$

and compute the average costs c_j as:

$$c_j = \beta_0 + \beta_j$$

$$c_6 = \beta_0$$

Using the identity (4) we can rewrite (3.1) as:

$$c_i = \beta_o \sum_{j=1}^{6} p_{ji} + \sum_{j=1}^{5} \beta_j p_{ji} + v_i$$

which yields

(5) $$c_i = \sum_{j=1}^{5} (\beta_0 + \beta_j) p_{ji} + \beta_0 p_{6i} + v_i$$

and which in turn is exactly equivalent to equation (3).

The coefficients (5) and (3) are related as follows:

(6) $$\gamma_j = \beta_0 + \beta_j \qquad j = 1 \cdots 5$$
$$\gamma_6 = \beta_0$$

This shows that models (2) and (3) are equivalent and that neither one involves the estimation of a constant term. We estimated both equations with the following results:

$$(2) \quad c_i = 166.5 + 432.4\ MS + 290.1\ OB + 132.0\ P + 1586.8\ PS$$
$$\qquad (246.05) \quad (249.78) \quad (256.54) \quad (409.25)$$

$$+ 133.2\ OUT \qquad R^2 = 0.177$$
$$(241.99)$$

where the excluded variable is Geriatrics (G). When this variable is included the estimated equation without a constant gives:

$$(3)' \quad c_i = 498.9\ MS + 356.6\ OB + 198.6\ P + 166.5\ G + 1653\ PS$$
$$(81.5) \quad (156.1) \quad (164.5) \quad (179.9) \quad (384.1)$$

$$+ 199.7\ OUT \qquad R^2 = .9399$$
$$(174.9)$$

We see from the regression coefficient that models (2) and (3)′ are equivalent. The coefficient of (G) in (3)′ is equal to the constant in (2) and the linear relationships between the coefficients specified in (6) hold to two decimal places. The variances of the coefficients, however, and the R^2's are different in the two estimated equations because of the different computation methods used in the two regressions. More specifically, when a zero intercept is chosen, all variances, covariances, and correlations are computed about the origin rather than about the mean. Since mathematically, neither equation includes a constant, we adopted estimation equation (3)′ as the correct estimation method without, however, attaching much importance to the high R^2. We see from equation (3)′ that all coefficients have the right signs, and all but one are highly significant. Geriatric cases are not significant at any acceptable level and the low coefficient is suspicious. We, nevertheless, decided to retain the model specification as shown by (3)'.

Interpretation of the Coefficients

Since we have assumed a linear cost function and since there is no constant in equation (3) the regression coefficients are marginal and average costs per case of each case-type. The order of magnitude of the results is borne out by a priori expectations except in the case of psychiatric and geriatric patients.

We calculated average costs per day for each of the case-types by dividing average costs per case by the sample average length of stay for each case-type as shown in Table 21. These results compare well with our a priori expectations of average daily costs. Discussions with hospital administrators led us to expect higher daily cost for OB than MS or P. The only surprising result is the very high cost of psychiatric cases. One possible explanation is that this coefficient captures not only the variable costs but also the large fixed costs of maintaining a psychiatric unit. For this reason, we ran a regression adding a dichotomous variable to indicate the existence or lack of such a unit. The coefficient of this variable was very small and insignificant. The extreme stability

of the psychiatric coefficient in many specifications of the cost function not reported here, as well as its high level of significance oblige us to take it on face value despite our reservations.

On the basis of the above results we must conclude that costs per case differ substantially with casemix. We performed the various tests for the equality of the various coefficients. All coefficients were significantly different from each other except for those of MS and OB. We nevertheless decided to maintain the six case-type classification. Since both coefficients are significant, the distinction between MS and OB is theoretically valid.

TABLE 21. --Average Cost Per Day and Per Visit by Case-type

	Average Cost Per Case ($)	Average Length of Stay	Average Cost Per Day or Visit ($)
Medical-Surgical	374.6	9.50	39.36
Obstetrics	353.9	3.80	92.89
Geriatrics	166.5	12.60	13.21
Pediatrics	186.8	4.37	42.50
Psychiatric	1706.0	5.71	299.12
Outpatient	658.8	32.00	20.51

APPENDIX B

THE SAMPLE DATA FROM 94 MICHIGAN SHORT-TERM HOSPITALS IN 1969

THE SAMPLE DATA FROM 94 MICHIGAN SHORT-TERM HOSPITALS IN 1969

The primary source of the data used in the present study was obtained from the American Hospital Association 1969 Annual Hospital Survey. A tape containing information on 251 Michigan hospitals was sorted to obtain data for short-term general hospitals with complete information on costs, inputs, and outpatient care. Because of the difference in the reporting period for various hospitals we were forced to drop from the sample hospitals with a reporting period of less than 365 days. To ensure comparability within the sample, hospitals with a reporting period ending on or before March, 1969 or extending beyond September, 1969 were also excluded. The final sample consisted of 94 hospitals ranging in bed size from 25 to 716. For thirty, the reporting period ended on June 30, 1969 while for the remaining sixty-four the data covered the year ending on September 30, 1969.

Hospitals are classified by the AHA according to type of service provided and form of control. Our sample includes only short-term medical and surgical hospitals, classified by the AHA as hospitals with fifty percent or more of the total patients staying less than thirty days. In the interest of homogeneity among hospitals we dropped all institutions which, although classified as short-term, allocate a substantial portion of their

bed complement to long-term care. Of the 94 hospitals in the sample, 28 had long-term beds, and the highest percentage of long-term to total beds was 17 percent.

Seventy-four hospitals in the sample are nongovernmental, nonprofit. Twelve of these are church-operated. Of the 20 governmental but nonfederal hospitals, five are managed by counties, twelve by city governments, and three are operated by a hospital district or authority. Since there is general agreement in the hospital literature that the type of control has very little influence on the economic behavior of the short-term nonprofit institutions, the sample was not stratified by this factor.

A supplementary set of data was obtained from the 1969 Michigan Hospital Survey conducted annually by the Michigan Department of Public Health, Office of Facilities Planning and Construction. Because the data are published only in summary form in the Michigan State Plan for Hospital and Medical Facilities Construction, the detailed data on patient utilization in selected categories were hand copied from the original questionnaire submitted by more than 140 hospitals. These data included the numbers of admissions and patient days in the following categories of care: general medical-surgical for persons under and over age 65, obstetrics, pediatrics, and psychiatric care.

To our knowledge, the only source of more detailed casemix data is the Commission on Professional and Hospital Activities in Ann Arbor, Michigan. A detailed breakdown by the four-digit ICDA code of case data from a large number of participating U.S. hospitals is collected by CPHA. Unfortunately, the data are virtually inaccessible to most health researchers. Bound by contractual obligations to member hospitals, CPHA requires a written release of information from each hospital. The obvious difficulty of obtaining such permission for a large enough sample, the high monetary cost of retreiving the information, as well as the participating hospitals' reluctance to divulge information, forced us to abandon this data source.

The AHA and Michigan Hospital Survey data were merged on one tape, checked for accuracy and consistency, and then adjusted to convert part-time labor inputs into full-time equivalents. The principal difficulty involved in merging the two tapes arose due to the fact that the Michigan Hospital Survey covers the full 1969 calendar year. This led to certain inconsistencies with the AHA data, especially in the total numbers of admissions and patient days. We were forced to drop from the sample the hospitals showing any sizeable differences. For the 94 institutions in the final date file, we assumed that the number of patients treated from June through December, 1968 was equal to the number treated during the same period in 1969. Based on this assumption, MHS patient care data were used to derive the output variables. All other data were taken from the AHA tape.

APPENDIX C

COUNTY CLASSIFICATION BY PREVAILING BLUE SHIELD AREAS

COUNTY CLASSIFICATION BY PREVAILING BLUE SHIELD AREAS

Area I:

Wayne	Macomb	Washtenaw
Oakland		

Area II:

Bay	Midland	Muskegon
Saginaw	Kent	

Area III:

Emmet	Roscommon	Shiawassee
Cheboygan	Ogenaw	Allegan
Presque Isle	Iosco	Barry
Charlevoix	Mason	Eaton
Antrim	Lake	Livingston
Leelanau	Osceola	Huron
Otsego	Clare	Tuscola
Montmorency	Gladwin	Sanilac
Alpena	Arenac	Lapeer
Benzie	Oceana	St. Clair
Grand Traverse	Newaygo	Berrien
Kalkaska	Mecosta	Van Buren
Crawford	Isabella	Cass
Oscoda	Montcalm	St. Joseph
Alcona	Gratiot	Branch
Manistee	Ottawa	Hillsdale
Wexford	Ionia	Lenawee
Missaukee	Clinton	Monroe

Area IV:

Keweenaw	Baraga	Delta
Houghton	Marquette	Schoolcraft
Ontonagon	Dickinson	Luce
Gogebic	Menominee	Mackinac
Iron	Alger	Chippewa

APPENDIX D

ESTIMATION OF THE PRODUCTION FUNCTION: EMPIRICAL RESULTS

ESTIMATION OF THE PRODUCTION FUNCTION: EMPIRICAL RESULTS

The estimation of the production function for our sample of 94 Michigan hospitals was based on a Cobb-Douglas function. The output data were the estimated values of "adjusted patient care" explained in Chapter III. The specification of the inputs used in the production of patient care, however, presented certain problems. Although the theory of production function estimation requires the specification of inputs in physical terms, this is not always possible. The AHA data contain information on the number of beds, nurses, interns, residents, and all other personnel. But two other important variables--supplies and hospital assets--can only be included in money terms, which means that a mixed specification of inputs is unavoidable. This problem has been noted in other studies. In one of the few such studies relating directly to the hospital industry, M. Feldstein[168] suggests that labor inputs should be aggregated by wage rates in order to achieve greater comparability among hospitals. This approach rests on the assumption that wage rates for different grades of nurses and other hospital personnel are fairly uniform across all sampled hospitals. Although this may be true in the case of the British National Health Service (with which Feldstein's study is concerned), we have shown that significant wage differentials exist among the various regions in the state of Michigan.[169] We decided therefore to express the labor units in physical terms.

The categories of labor inputs for which AHA data are available are (1) registered nurses, (2) licenses practical nurses, and (3) a general category entitled "all other personnel" which excludes physicians, administrators, interns, and residents. Data on both full-time and part-time personnel are included in the AHA tapes (part-time employees were converted into full-time equivalents by assuming a conversion factor of two). The use of physical units does not cause any problems in the case of RNs and LPNs since the groupings are fairly homogeneous, but the third category contains a number of occupations the mix of which may vary among hospitals. While such variability introduces some bias into the individual coefficients of the production function, it was thought likely that the bias resulting from the exclusion of this variable would be even greater.

Unfortunately, data on the number of physicians providing care in each institution proved unavailable. An attempt to obtain such data via a questionnaire to the sample hospitals was not successful. Although the response rate was satisfactory, the double counting resulting from multiple staff appointments, as well as the varied methods of reporting employed by the hospitals, made the data unusable. The AHA data includes information on the numbers of interns and residents for the twenty-one teaching hospitals in the sample. But although interns and residents provide a significant amount of patient care in these hospitals, they cannot be used as a separate category of labor inputs because the specification of the Cobb-Douglas function does not permit entering any inputs at zero levels (which would be required for the hospitals which do not have intern and residency programs). Nor was it feasible to estimate two different production functions, one for teaching and one for nonteaching hospitals, since interns and residents do not provide all patient care. For these reasons we were forced to treat physicians as managerial rather than technical inputs. Doctors are thus assumed to determine the form of the production function in terms of their decisions regarding the use of other inputs, but they do not enter the production process as separate inputs.

Even larger problems were faced in estimating the various capital inputs used in the hospital production function. Information on capital is at best scattered, and the only available data relating to physical units were the numbers of x-ray and cobalt treatment units available in each hospital as tabulated in the Michigan Hospital Survey. Since these two inputs alone hardly express the total number of capital services, we decided to use the total dollar value of hospital assets as an instrumental variable for capital. This assumes that hospital assets are positively correlated with capital and are uncorrelated with the error term in the production function equation.[170] Both of these assumptions seem fairly reasonable.

The two remaining inputs in the hospital production process are beds and supplies. The number of hospital beds represents another capital variable (and is included in money terms in the assets variable) and is used as a measure of hospital capacity. Finally, the supplies variable includes certain drugs and dressings, x-ray films, etc., and it is expressed in money terms.

We first estimated a Cobb-Douglas function of the form:

$$Y_i^* = A \prod X_{is}^{\alpha_s} \, \mathcal{E}j,$$

where

Y_i^* = the number of units of "adjusted patient care"

X_1 = beds

X_2 = registered nurses

X_3 = licensed practical nurses

X_4 = all other personnel

X_5 = assets (in thousands of dollars)

X_6 = supplies (in thousands of dollars)

The estimated coefficients are shown in Table 22.

TABLE 22.--Results from Regression Number One

	Coefficients	Std. Errors of Coefficients	t-Values
Constant	4.889		
X_1	0.764	0.083	9.2
X_2	0.139	0.016	8.6
X_3	-0.016	0.021	-0.7
X_4	0.219	0.007	31.2
X_5	0.011	0.026	0.4
X_6	0.000	0.003	0.0
R^2	0.9863		

The nagative sign of X_3 is contrary to the theoretical expectation of positive input elasticities. Since the variable also appears to be statistically insignificant we examined the possibility of multicollinearity between X_2 and X_3. The simple correlation coefficient between the two variables was 0.874, and it was decided to lump X_2 and X_3 together and consider the two types of nurses as one hospital input. The new and final set of inputs is thus defined as:

X_1 = number of beds

X_2 = nurses (full-time equivalents)

X_3 = all other personnel (full-time equivalents)

X_4 = supplies (in thousands of dollars)

X_5 = assets (in thousands of dollars

The coefficients estimated in logarithmic form are shown in Table 23.

TABLE 23. --Results from Regression Number Two

	Coefficients	Std. Errors of Coefficients	t-Values
Constant	5.020		
X_1	0.595	0.097	6.10
X_2	0.180	0.049	3.67
X_3	0.212	0.067	3.14
X_4	0.001	0.004	0.00
X_5	0.039	0.015	2.60
R^2	0.9805		

Variables X_1, X_2, X_3, and X_5 are significant at the 99 percent level. X_4 (supplies) is not significant at any level. This second regression was used to derive the residuals which form the productivity measures utilized in Chapter VI.

FOOTNOTES AND REFERENCES

FOOTNOTES AND REFERENCES

[1]D. P. Rice, B. S. Cooper, "National Health Expenditures, 1929-1971," Social Security Bulletin, XXXV, No. 1 (1972), 14-15.

[2]M. S. Mueller, "Private Health Insurance in 1970: Population Coverage, Enrollment, and Financial Experience," Social Security Bulletin, XXXV, No. 2 (1970), p. 15.

[3]See, for example, Gerald Rosenthal, "The Public Pays the Bill," The Atlantic, July, 1966; and Herman M. Somers and Ann R. Somers, Medicare and the Hospitals (Washington, D.C.: The Brookings Institution, 1967).

[4]Martin S. Feldstein, The Rising Cost of Hospital Care Washington, D.C.:. Information Resources Press, 1971), p. 76.

[5]Ibid. p. 77.

[6]Department of Health, Education, and Welfare, A Report to the President on Medical Care Prices (Washington, D.C.: Government Printing Office, 1967).

[7]i.e., the patient pays part of the additional cost.

[8]Of course, to the extent that such premiums affect the consumer's disposable income, they also affect his demand for hospital care, although not significantly.

[9]Philanthropy, however, has greatly decreased as a percent of total hospital revenue in recent years. See p. 34.

[10]For a discussion of the conditions determining hospital demand see G. D. Rosenthal, The Demand for General Hospital Facilities (Chicago, Ill.: American Hospital Association, 1967).

[11]Private conversations with Michigan Medicaid officials,

[12] M. V. Pauly and D. F. Drake, "Effect of Third-Party Methods of Reimbursement on Hospital Performance," in Empirical Studies in Health Economics, H. E. Klarman, ed. (Baltimore, Maryland: Johns Hopkins Press, 1970).

[13] Karen Davis, Theories of Hospital Inflation: Some Empirical Evidence (Washington, D.C.: The Brookings Institution, 1971).

[14] Ibid., p. 20.

[15] Department of Health, Education, and Welfare, Social Security Administration, Office of Research and Statistics, Reimbursement Incentives for Hospital and Medical Care (Washington, D.C.: Government Printing Office, 1968), p. iii.

[16] A survey of various proposals can be found in L. P. Hardwick, S. B. Meyers, and L. Woodruff, Incentive Reimbursement: Prospects, Proposals, Plans, and Programs, Research Series No. 6 (Feb., 1969), Blue Cross of Western Pennsylvania; also see: Technical Work Group on Health Care Costs, "Payment Mechanisms and Reimbursement Controls," pp. 367-445, in Rising Medical Costs in Michigan: The Scope of the Problem and the Effectiveness of Current Controls (Lansing, Michigan: Department of Social Services, November, 1972); W. E. Seago, The Effect of the Medicare Principles of Reimbursement on the Allocation of Hospital Costs Between the Medicare Program and Non-Medical Patients (Athens, Georgia: University of Georgia, 1970).

[17] The Boston Consulting Group, Reimbursing Hospitals on Inclusive Rates, Reprinted by U.S. Department of Commerce, National Technical Information Service (Washington, D.C.: U.S. Government Printing Office, 1971).

[18] Department of Health, Education, and Welfare, Social Security Administration, Office of Research and Statistics, Medical Care Costs and Prices: Background Book (Washington, D.C.: Government Printing Office, January, 1972), p. 137; C. P. Hardwick, H. Wolfe, "Evaluation of an Incentive Reimbursement Experiment," Medical Care, X, No. 2, pp. 109-117.

[19] Judith R. Lave, Lester B. Lave, and Lester P. Silverman, A Proposal for Incentive Reimbursement for Hospitals (Pittsburgh, Pa.: Carnegie-Mellon University, 1971), p. 2.

[20]Ibid., p. 4.

[21]Larry J. Shuman, Harvey Wolfe, and C. Patrick Hardwick, "Predictive Hospital Reimbursement and Evaluation Model," Inquiry, IX, No. 2 (June, 1972), pp. 17-33.

[22]Ibid., p. 30.

[23]Saul Waldman, "'Average Increase in Costs'--An Incentive-Reimbursement Formula for Hospitals," in Department of Health, Education, and Welfare, Reimbursement Incentives for Hospital and Medical Care, pp. 39-48.

[24]Karen Davis, Economic Theories of Behavior in Nonprofit, Private Hospitals (Brookings Institution, Washington, D.C., 1971).

[25]Mary Lee Ingbar and Lester D. Taylor, Hospital Costs in Massachusetts (Cambridge, Mass.: Harvard Univeristy Press, 1968), p. 74.

[26]American Hospital Association, Factors to Evaluate in the Establishment of Hospital Charges (Chicago: American Hospital Association, 1966).

[27]Millard F. Long, "Efficient Use of Hospitals," in The Economics of Health and Medical Care (Ann Arbor: University of Michigan Press, 1967); also Herbert E. Klarman, The Economics of Health (New York: Columbia University Press, 1965); Melvin W. Reder, "Some Problems in the Economics of Hospitals," American Economic Review, LV (May, 1965).

[28]Klarman, The Economics of Health, p. 121.

[29]The model, therefore, also makes the implicit assumption that demand is fairly elastic. Of course, another way to increase the amount of output sold is by increasing the quality of care or by offering more "frills" like color television and other luxuries.

[30]See, for example, Walter J. McNerney and Study Staff, Hospital and Medical Economics (Chicago: Hospital Research and Educational Trust, 1962), p. 923.

[31] Edward M. Kaitz, Pricing Policy and Cost Behavior in the Hospital Industry (New York: Frederick A. Praeger, Inc., 1968), p. 38-39.

[32] Karen Davis, Net Income of Hospitals, 1961-1969 (Washington, D.C.: Social Security Administration, 1970).

[33] J. P. Newhouse, "Toward a Theory of Nonprofit Institution: An Economic Model of the Hospital," American Economic Review (March, 1970).

[34] Melvin W. Reder, "Some Problems in the Economics of Hospitals."

[35] Karen Davis, A Theory of Economic Behavior in Nonprofit, Private Hospitals, Unpublished doctoral dissertation, Rice University, May, 1969.

[36] Except, perhaps, for the quantity maximization hypothesis, Even there, however, it is reasoned that superior equipment will attract more and better doctors in the staff and thus increase the number of patients in the long run.

[37] J. P. Newhouse, "Toward a Theory of Nonprofit Institution: An Economic Model of the Hospital;" also, R. D. Fraser, "The Economics of Nonprofit Institutions: The Hospital," Canadian Perspectives in Economics (Collier-Macmillan Canada, Ltd., 1972).

[38] R. D. Fraser, Ibid., p. 4.

[39] H. D. Mauksh, "The Organization Context of Nursing Practice," in The Nursing Profession. Fred Davis, ed. (New York: John Wiley and Sons, 1966), pp. 116-130.

[40] W. J. Baumol and W. G. Bowen, "On the Performing Arts," American Economic Review Proceedings (May 1965), pp. 495-502.

[41] Ibid., p. 497.

[42] Karen Davis and Richard W. Foster, Community Hospitals: Inflation in the Pre-Medicare Period (Washington, D.C.: Government Printing Office, 1972).

[43] Ibid., p. 44.

[44] Charlotte Muller, Paul Worthington, and George Allen, Sources and Management of Capital in New York City Voluntary Hospitals (New York: City University of New York, 1972), p. 179.

[45] As long as the average charge per day is higher than the average daily cost.

[46] For a discussion of the definitional problems inherent in the concept of health and the difficulties of measuring it, see Richard Auster, Irving Leveson, and Deborah Sarachek, "The Production of Health, An Exploratory Study," The Journal of Human Resources, IV, No. 3 (January, 1969), pp. 413-436; also Mary Larmore, An Inquiry into an Econometric Production Function For Health in the United States (Ann Arbor, Michigan: University Microfilms, 1967).

[47] And probably not the most important. See R. Auster et al., Ibid.

[48] Although these are measures of inpatient care, some attempts to include outpatient visits have been made. Notably, the American Hospital Association data include a measure of "adjusted" patient days where an outpatient visit is weighted as one-fourth of an inpatient day. The problems of comparing outpatient visits with either cases or patient days and the arbitrary aggregation of inpatient and outpatient care are obvious.

[49] An impressive body of casemix data for a large number of hospitals is collected by the Commission on Professional and Hospital Activities in Ann Arbor, Michigan. Unfortunately, although the data are available to researchers, CPHA is bound by contract not to identify the individual hospitals, unless prior written approval by the institutions is obtained. The difficulty of obtaining this approval for a sizeable sample of hospitals is considerable especially in view of the reluctance of most institutions to subject their data to public scrutiny.

[50] Martin S. Feldstein, Economic Analysis for Health Service Efficiency (Amsterdam: North-Holland Publishing Co., 1967).

[51] There are also the estimation difficulties associated with inverting large matrices. See: Judith R. Lave, Lester B. Lave, and Lester P. Silverman, A Proposal for Incentive Reimbursement for Hospitals; also Martin S. Feldstein, Ibid.

[52] In the absence of casemix data this correlation cannot, naturally, be established.

[53] R. E. Berry, Competition and Efficiency in the Market for Hospital Services, The Structure of the American Hospital Industry (Ph.D. Dissertation, Harvard University, Cambridge, 1965).

[54] Berry, for example, used fifteen percent of all short-term general and special non-Federal hospitals in the U.S. Ibid.

[55] Saarthoff, D. and Kurtz, R., "Cost Per Day Comparisons Don't Do the Job," The Modern Hospital, Vol. 94, No. 4 (October, 1962).

[56] H. A. Cohen, "Variations in Cost Among Hospitals of Different Sizes," Southern Economic Journal, Vol. 33, No. 3 (Jan., 1967), pp. 355-366; also H. A. Cohen, "Hospital Cost Curves With Emphasis on Measuring Patient Care Output," in Empirical Studies in Health Economics, H. E. Klarman, Ed. (Baltimore, Maryland: Johns Hopkins Press, 1970).

[57] Martin S. Feldstein, Economic Analysis for Health Service Efficiency, pp. 30-31, 96, for an exposition of the underlying assumptions.

[58] Ibid., p. 96.

[59] Ibid., p. 24. Feldstein assumes that length of stay variations among hospitals are due only to casemix differences. This, as will be shown, is not supported by available evidence.

[60] See Commission on Professional and Hospital Activities, Length of Stay in PAS Hospitals (Ann Arbor, Mich.: Commission on Professional and Hospital Activities, 1969).

[61] International Classifications of Diseases, Adopted for Indexing Hospital Records by Diseases and Operations, U.S. Public Health Publication 719 (Revised, 1962).

[62] For a discussion of the reasons behind interhospital differences in length of stay see: School of Public Health and Administrative Medicine, Columbia University, Prepayment for Hospital Care in New York State (New York: Columbia University Press, 1960), pp. 251-257; I. H. Hayes and H. Becker, eds., Financing Hospital Care in the United States, Vol. II, Prepayment and the Community (New York: The Blakiston Press, 1955), pp. 290-293; also J. G. Zimmer, "Length of Stay and Hospital Bed Misutilization," paper delivered at American Public Health Association annual meeting, Atlantic City, New Jersey, 1972.

[63]M. W. Reder, "Economic Theory and Nonprofit Enterprise," American Economic Review, Vol. 55 (May, 1965), p. 475.

[64]See the remarks of Dr. R. E. Trussel, Commissioner of Hospitals of New York, in "Conference on Research in Hospital Use," Jan. 22-23, 1963, Chicago (U.S. Dept.) HEW, Public Health Service, 1963.

[65]Under current government and third party reimbursement formulae hospitals can charge average cost per day which after the first few days is higher than the marginal cost of an additional day of care.

[66]This is a weak assumption because it is true that some hospitals manage to shorten the length of stay by more efficient operation or the use of timesaving capital equipment. We believe, however, that the argument is generally true for the majority of hospitals.

[67]Or even to the less plausible conclusion that the sum of the input elasticities becomes negative after a number of days.

[68]Even though it is known that some patients stay in the hospital beyond the point warranted by strict medical considerations, quite often psychic benefits occur which can be seen as a net addition to patient care or hospital output.

[69]Although, as Martin S. Feldstein points out in Economic Analysis for Health Service Efficiency, p. 24, it is easy to exaggerate the extent to which differences in length of stay imply differences in quality of care.

[70]"Hotel" services include the use of beds, board, and routine nursing services.

[71]More specifically, the ADSC reflects the charge for the most common type of accomodation (i.e., a two-bed room).

[72]See, in particular, H. M. Somers and A. R. Somers, Medicare and the Hospitals (Washington, D.C.: The Brookings Institution, 1967); P. J. Feldstein and S. Waldman, "Financial Position of Hospitals in the Early Medicare Period," Social Security Bulletin, 31 (October, 1968), pp. 18-23.

[73] Martin S. Feldstein, The Rising Cost of Hospital Care (Washington, D.C.: Information Resources Press, 1971), p. 8.

[74] For further evidence on this point see Martin S. Feldstein, Economic Analysis for Health Service Efficiency Chapter 3.

[75] This should not be taken to mean that hospitals will act in a socially reckless fashion. It is simply to show that economic incentives to inefficiency and moral irresponsibility will actually exist if an incentive reimbursement plan measures costs by the average cost per patient day.

[76] Martin S. Feldstein, The Rising Cost of Hospital Care, p. 8.

[77] Because of the great number of hospital cost review articles we will not give a thorough review of the literature. For examples of cost studies and surveys of the relevant work see R. E. Berry, Jr., "Returns to Scale in the Production of Hospital Services," Health Services Research, 2 (Summer, 1967), pp. 123-139; W. John Carr and Paul J. Feldstein, "The Relationship of Cost to Hospital Size," Inquiry, IV (June, 1967), pp. 45-65; Harold A. Cohen, "Variations in Cost Among Hospitals of Different Sizes," Martin S. Feldstein, Economic Analysis for Health Services; Mary Lee Ingbar and Lester D. Taylor, Hospital Costs in Massachusetts; J. R. Lave, "A Review of Methods Used to Study Hospital Costs," Inquiry, 3 (May, 1966); J. K. Mann and D. E. Yett, "The Analysis of Hospital Costs: A Review Article," Journal of Business, Vol. 41, No. 2 (1968).

[78] See Technical Work Group on Health Care Costs, Rising Medical Costs in Michigan: The Scope of the Problem and the Effectiveness of Current Controls (Lansing, Michigan: Department of Social Services, Nov., 1972), pp. 151-154.

[79] Mary Lee Ingbar and Lester D. Taylor, Hospital Costs in Massachusetts.

[80] See W. John Carr and Paul J. Feldstein, "The Relationship of Cost to Hospital Size."

[81] Paul J. Feldstein, An Empirical Investigation of the Marginal Cost of Hospital Services (University of Chicago: Graduate Program of Hospital Administration).

[82] Ray E. Brown, "What Do We Mean by Hospital Costs," Hospital Forum, Vol. 1, No. 4 (May, 1961).

[83] Paul J. Feldstein, An Empirical Investigation of the Marginal Cost of Hospital Services, p. 49.

[84] See W. John Carr and Paul J. Feldstein, "The Relationship of Cost to Hospital Size," p. 42.

[85] See Mark S. Blumberg, "'DPF Concept' Helps Predict Bed Needs," Modern Hospital, Vol. 97, No. 6 (December, 1961), pp. 75-81.

[86] See Judith R. Lave, Lester B. Lave, and Lester P. Silverman, A Proposal for Incentive Reimbursement for Hospitals for an extensive analysis of this plan.

[87] Ibid. p. 5.

[88] See Ralph E. Berry, Jr., "Product Heterogeneity and Hospital Cost Analysis," Inquiry, Vol. 7, No. 4 (March, 1970), pp. 67-75.

[89] See comment following R. E. Berry. Ibid.

[90] See Larry J. Shuman, Harvey Wolfe, and C. Patrick Hardwick, "Predictive Hospital Reimbursement and Evaluation Model," Inquiry, Vol. IX, No. 2 (June, 1972).

[91] See J. R. Lave, Lester B. Lave, and Lester P. Silverman, A Proposal for Incentive Reimbursement for Hospitals.

[92] Ibid., p. 10.

[93] See footnote 49 for discussion.

[94] American Hospital Association, "Guide Issue" of Hospitals (Chicago, Illinois: American Hospital Association, 1971), p. 446.

[95] J. R. Lave and L. B. Lave, "Hospital Cost Functions," American Economic Review, Vol. 60, No. 3 (June, 1970).

[96] See L. J. Shuman, H. Wolfe, and C. P. Hardwick, "Predictive Hospital Reimbursement and Evaluation Model," Inquiry, Vol. IX, No. 2.

[97] See Mark V. Pauly, "Efficiency Incentives and Reimbursement for Health Care," Inquiry, Vol. 7, No. 1 (1970), p. 125.

[98] A detailed description of the sample and of the other data used is found in Appendix B.

[99] See Martin S. Feldstein, Economic Analysis for Health Service Efficiency, p. 26.

[100] Ibid., p. 24.

[101] Irwin Walkstein, "The Legislative History of Hospital Cost Reimbursement," in Department of Health, Education, and Welfare, Social Security Administration, Office of Research and Statistics, Reimbursement Incentives for Hospital and Medical Care (Washington, C.D.: Government Printing Office, 1968), p. 5.

[102] Blue Cross Member Hospital Contract (Philadelphia, Pa.: The Associated Hospital Service of Philadelphia, July, 1958).

[103] T. Fitzpatrick, D. Riedel, and B. Payne, "Character and Effectiveness of Hospital Use," Hospital and Medical Economics, Vol. 1, edited by Walter J. McNereney (Chicago: Hospital Research and Education Trust, 1962), p. 474.

[104] Such direct institutional controls and their possible effects are analyzed extensively in Technical Work Group on Health Care Costs, Rising Medical Costs in Michigan: The Scope of the Problem and the Effectiveness of Current Controls, pp. 342-364.

[105] Medicare's approach to these problems was to require hospital committee utilization review and physician certification of the medical necessity of services. These requirements established a procedure for review of admissions, duration of stay, and services furnished. After 20 days of continuous service, the utilization-review committee determines in all cases whether further services are needed.

[106] Since relative hospital cost is average cost divided by a constant the correlation between Cr and bed size is the same as that between average cost and size.

[107] We should remember that location is used as a surrogate for various factors responsible for differences in costs such as size, facilities and services, and wage differentials.

[108] The use of the fifty percent factor in determining rewards and penalties is completely arbitrary and intended only for purposes of example. An actual reimbursement plan could use various modifications of the formulas above such as rewarding (or penalizing) by 100 percent of the savings (or excess costs).

[109] See the Western Pennsylvania proposal, Chapter II, p. 20.

[110] Paul J. Feldstein, "An Analysis of Reimbursement Plans," in Department of Health, Education, and Welfare, Social Security Administration, Office of Research and Statistics, Reimbursement Incentives for Hospital and Medical Care (Washington, D.C.: Government Printing Office, 1968), p. 24.

[111] See Table 10.

[112] See Technical Work Group on Health Care Costs, Rising Medical Costs in Michigan: The Scope of the Problem and the Effectiveness of Current Controls (Lansing, Michigan: Department of Social Services, Nov., 1972), Table

[113] The ideal casemix adjustment would take care of such differences. Unfortunately, previous attempts have been unsuccessful and, therefore, as we mentioned in Chapter II, we will have to rely on location as a proxy variable for differences in facilities and services.

[114] In the earlier case such "reverse" payments would occur in only twenty-two hospitals.

[115] The average utilization rate for these seven institutions was 57 percent.

[116] Paul B. Ginsburg, "Resource Allocation in the Hospital Industry: The Role of Capital Financing," Social Security Bulletin, 35, No. 10 (October, 1972), p. 30.

[117]C. P. Hardwick and H. Wolfe, "Evaluation of an Incentive Reimbursement Mechanism," Medical Care, X, No. 2, pp. 109-117.

[118]Ibid., p. 113.

[119]Martin S. Feldstein, Economic Analysis for Health Service Efficiency.

[120]M. Hall and C. B. Winsten, "The Ambiguous Notion of Efficiency," Economic Journal, 1959, p. 75.

[121]M. J. Farrell, "The Measurement of Productive Efficiency," Journal of the Royal Statistical Society, Series A, Part III, 1957, p. 253.

[122]J. Marschak and W. H. Andrews, "Random Simultaneous Equations and the Theory of Production," Econometrica, 1944, p. 143.

[123]Actually, $C_1 \ldots C_5$ represent the physical amounts of input X_2 used. They would, however, represent the total cost of production if multiplied by the fixed price of X_2.

[124]This assumption is also used by M. Feldstein.

[125]Since the observation for the average productivity hospital would lie on the regression line the residual would be zero, and, therefore:

$$P_i^* = 1 + 0 = 1.$$

[126]A. A. Walters, "Production and Cost Functions: An Econometric Survey," Econometrica, Vol. 31, No. 1-2, (January-April, 1963).

[127]E. Heady, T. Dillon, Agricultural Production Functions, Iowa State University Press, 1961.

[128]E. Heady, ed., Resource Productivity, Returns to Scale, and Farm Size., (Iowa State University Press, 1956) p. 148.

[129]E. Heady, T. Dillion, Agricultural Production Functions.

[130]A. A. Walters, "Production and Cost Functions: An Econometric Survey."

[131] Melvin Reder, "An Alternative Interpretation of the Cobb-Douglas Function," Econometrica, Vol. II, 1943, p. 259.

[132] Cecil B. Haver, "Economic Interpretation of Production Function Estimates," in Resource Productivity Returns to Scale and Farm Size, E. Heady, ed., p. 145.

[133] M. Bronfenbrenner, "Production Functions: Cobb-Douglas, Interfirm, Intrafirm," Econometrica, Vol. 12, 1944, p. 35.

[134] Yair Mundlak, "Empirical Production Functions Free of Management Bias," Journal of Farm Economics, February 1961.

[135] Irving Hoch, "Estimation of Production Function Parameters and Testing for Efficiency," Econometrica, Vol. 23, 1955.

[136] H. O. Carter, "Modification of the Cobb-Douglas Function to Destroy Constant Elasticity and Symmetry," Resource Productivity Returns to Scale and Farm Size, E. Heady, ed., p. 168.

[137] Knud Rasmussen, Production Function Analyses for British and Irish Farm Accounts, (Nottingham, University of Nottingham Press, 1962).

[138] Since the function is estimated in its logarithmic form, if $x_i = 0$, $\log x_i$ is undefined.

[139] This, of course, is consistent with the necessary assumption of similar production functions for all firms in the cross-section study.

[140] K. J. Arrow et al., "Capital-Labor Substitution and Economic Efficiency," Review of Economics and Statistics, Vol. 43, 1961, pp. 225-235.

[141] E. Heady and T. Dillon, Agricultural Production Functions, pp. 83-96.

[142] Irving Hoch, "Simultaneous Equation Bias in the Context of the Cobb-Douglas Production Function," Econometrica (October, 1958), pp. 556-578.

[143]A. Zellner, J. Kmenta, and T. Dreze, "Specification and Estimation of Cobb-Douglas Production Function Models," Econometrica (October, 1966), p. 325.

[144]The meaning of the cost minimization hypothesis as applied to hospital economics was explained in Chapter II, pp. 27, 28.

[145]T. Marschak and W. H. Andrews, "Random Simultaneous Equations and the Theory of Production."

[146]See Irving Hoch, "Simultaneous Equation Bias in the Context of the Cobb-Douglas Production Function."

[147]See A. Zellner et al., "Specification and Estimation of Cobb-Douglas Production Function Models."

[148]There are three types of hospital employees generally called "nurse" but there are great differences in their training and status: (1) the registered nurse (RN or "professional nurse") is licensed by the state after training in one of three types of schools--a three-year hospital school of nursing, a four-year baccalaureate program, or a two-year program in a community college; (2) the licensed practical nurse or LPN (called licensed vocational nurse in some states), requiring a year of training at a vocational school; and (3) the nurse's aide (sometimes called "attendant," "ward maid," or "male orderly") who usually has no formal training other than that received on the job and is not licensed.

[149]Martin S. Feldstein, Economic Analysis for Health Service Efficiency, pp. 96-97.

[150]See Chapter V, p. 109.

[151]This is the method used by the AHA. See 1971 Guide Issue of Hospitals, p. 448.

[152]We could not obtain information on the amount of time physicians spent in each hospital.

[153]A. A. Walters, "Production and Cost Functions pp. 23-24.

[154]J. Kmenta, Elements of Econometrics (New York: Macmillan, 1971), p. 309.

[155] For a discussion of the Binomial Test see W. T. Conover, Practical Nonparametric Statistics (New York: John Wiley and Sons, 1971), pp. 77-78.

[156] Pr is a productivity index derived from a production function estimated with simply the number of cases as the dependent variable.

[157] These conclusions agree with those arrived at by M. Feldstein.

[158] Or some other measure of costs which adjusts for differences in casemix and length of stay as well as for the other factors which affect costs but are not related to efficiency.

[159] See Chapter II, p. 20.

[160] A second assumption which will not be tested is that these differences in average costs are due to production function differences among case-types. We assume, in other words, that production functions are identical among hospitals, except for the time required for the treatment of each case.

[161] Martin S. Feldstein, Economic Analysis for Health Service Efficiency; J. R. Lave, L. B. Lave, and L. P. Silverman, Aggregation in Regression Analysis Applied to Hospital Cost Estimation.

[162] Jan Kmenta, Elements of Econometrics.

[163] Adjusted hospital cases are the sum of cases in the five inpatient categories plus the number of outpatient visits. We divided outpatient visits by thirty-two to convert them into case equivalents.

[164] For an analysis of this and other statistical methods used to avoid multicollinearity see J. R. Lave, L. B. Lave, and L. P. Silverman, Aggregation in Regression Analysis Applied to Hospital Cost Estimation.

[165] General medicine, pediatrics, ear, nose, and throat, traumatic and orthopedic surgery, other surgery, gynecology, and others.

[166] Martin S. Feldstein, Economic Analysis for Health Service Efficiency, p. 140.

[167]Otherwise the X'X matrix used in the computation of the β_j's will be singular.

[168]Martin S. Feldstein, Economic Analysis for Health Service Efficiency.

[169]See Chapter V, p. 109.

[170]A. A. Walters, "Production and Cost Functions: An Econometric Survey,", pp. 23-24.