1981

Doctors' Incentives regarding Length of Stay - An Extension and Test of the Pauly-Redisch Hypothesis

Stephen E. Margolis

ECONOMICS AND LAW WORKSHOP

81-23

DOCTORS' INCENTIVES REGARDING LENGTH OF STAY - AN EXTENSION AND TEST OF THE PAULY-REDISCH HYPOTHESIS

by

Stephen E. Margolis

Major funding for the Centre for Economic Analysis of Property Rights has been provided by The Bureau of Corporate Affairs, Consumer and Corporate Affairs, and by the Academic Development Fund, The University of Western Ontario. The views expressed by individuals associated with the Centre do not reflect official views of the Centre, The Bureau of Corporate Affairs, or The University of Western Ontario.

Subscriptions to the Workshop papers and the Working Paper Series are $40 per year for institutions and $25 per year for individuals. Individual copies, if available, may be purchased for $3 each. Address all correspondence to John Palmer, Centre for Economic Analysis of Property Rights, The University of Western Ontario, London, Ontario, CANADA N6A 5C2

DOCTORS' INCENTIVES REGARDING LENGTH OF STAY -

AN EXTENSION AND TEST OF THE PAULY-REDISCH HYPOTHESIS

by

Stephen E. Margolis*

August 1981

*Assistant Professor of Economics, University of Western Ontario.
I would like to acknowledge helpful comments from Stan Liebowitz and a referee. Able research assistance was provided by Laurie Dunk.

ABSTRACT

Pauly and Redisch's hypothesis of non-profit hospitals is extended to show the implications of the cooperative model regarding lengths of stay for hospital patients. Empirical implications of the cooperative model are shown and contrasted with the null hypothesis of atomistic behavior. Empirical tests are presented which reflect favorably on the cooperative hypothesis. Estimates suggest that seventy to eighty percent of small changes in the number of hospital beds per capita will be accommodated as adjustments in average length of stay. Such a result may warrant a major re-evaluation of the expected effect of current hospital regulations.

The fact that the majority of U.S. hospitals are not-for-profit institutions has led economists to speculate about the maximand implicit in the behavior of hospitals. One response has been an extension of theories of bureaucracy to hospital administration (q.v., Miqué and Bélanger 1974). Newhouse has constructed a theory of hospital behavior which posits that quality and quantity enter the objective function of hospital administrators. A strikingly different approach is taken by Pauly and Redisch (1973). Pauly and Redisch argue that the objective of a hospital is maximization of collective profits of the doctors who practice on the hospital staff. They suggest that the separation of the patient's bill into two separate components (payment to doctor and to hospital) creates an illusion of the existence of separate organizations with separate interests. Their theory has substantial intuitive appeal given the strong influence of medical professionals on the administration of a hospital, the authority of doctors to give orders to hospital employees, and the obvious dependence of the hospital on its doctors for generating business. While the Pauly-Redisch hypothesis has gained substantial recognition among health economists, there is little empirical demonstration of its validity.

This paper extends the Pauly and Redisch hypothesis to one aspect of physicians decisionmaking, the decision regarding the length of stay of patients in hospitals. It is assumed here that the decision regarding length of stay is made by the physician. Such an assumption is simplifying, though not strictly necessary. It is widely argued that patients allow physicians to make certain decisions for them, since the patient is not often sufficiently informed to make these decisions himself. In most cases the physician must "sign out" the patient. Where the patient elects to sign himself out of a hospital, the physician is absolved, to some degree, of responsibility for the patient's condition. Finally, the physician may influence the pace of treatment and in so doing, restrict the patient's options still further.

Beyond contributing to a theory of hospital behavior, the issue of length of stay is crucial for present policy purposes. In both the United States and Canada various jurisdictions have been reducing the availability of hospital beds as a means of controlling medical costs.[1] Such policies follow from the view that with the presence of insurance and the inability of patients to determine their wants, many of the usual disciplines of the market appear to be absent from the medical industry. A constraint on the availability of hospital beds is assumed to be an avenue for controlling the number of procedures performed by the medical profession.[2] The empirical findings presented here, however, suggest that an adjustment in the number of hospital beds is largely accommodated by an adjustment in length of stay, and has little effect on the number of hospitalizations. Although hospitalization is costly, the marginal day for a given patient is likely to be cheap, as the demands of a patient on the last day of recuperation are probably substantially less than his demands early in his stay. Treatment of a recuperating patient outside the hospital also imposes costs, costs which in some cases may exceed those which would occur inside the hospital. Further, a constraint on the availability of beds may impair the ability of the medical system to cope with emergencies. Thus the savings available from regulating the number of hospital beds may be far less than has generally been expected.

---

[1] For a discussion of certificate of need legislation, see Sloan and Steinwald.

[2] The relation between hospital availability and use is referred to in the medical literature as Roemer effects. See Roemer.

The paper proceeds as follows. In the next section, two models of physician behavior are presented. The first model assumes atomistic behavior on the part of physicians. In that case, each physician treats his patients as he sees fit, given his patients' demand for treatment and his production technology. The second model assumes the cooperative profit maximizing behavior of the Pauly and Redisch model. Here physicians behave as though they take into account the opportunity cost of a bed-day for the staff of the hospital. Such collective maximization would require a high degree of cooperation. In the atomistic model, physicians use hospital beds in their own self-interest. Given that the capacity constraint will occasionally be binding, such behavior is contrary to collective profit maximization. In the cooperative model, the full costs of bed use are internalized. In Section II an empirical version of the model is specified. That section concludes with a brief comment on some previous empirical studies of length of stay. Section III reports a test of the model which consists of estimation of a reduced form which can distinguish between cooperative and atomistic behavior.

I. The Model

Atomistic Behavior

A first approach to physicians' behavior, argued here to be naive, postulates that physicians on a hospital staff maximize their own private profits, taking prices and availability of a hospital's resources as given. Once a patient is admitted to the hospital, the physician behaves in a manner consistent with private profit maximization, as though his decisions will have no effect on future availability of beds to him. Such an assumption may not reflect behavior perfectly even in the non-cooperative case, since the doctor's own decisions may affect subsequent availability of beds to him. However, the benefits to a single doctor of his own conservation of hospital facilities are likely to be both small and remotely connected with his decisions.[3]

In both of the models introduced below, the unit of output is a "cure attempt" or simply a "cure". It is not necessary that each hospitalized patient receive one cure attempt. Seriously ill patients may receive several "cures". Also quality variations will be ignored here by subsuming them as variations in output. So, for example, if an ordinary hernia repair is one unit of output, a very nice one might be 1.5. A doctor chooses Q, the number of cure attempts, and H, the length of stay. The doctor faces a demand function for his services which relates output to the total payments for a cure attempt. This amounts to an assumption that the patient is indifferent between paying a dollar to the doctor and paying a dollar to the hospital. The doctor's choice of Q thus determines T, the total charge per patient that will allow Q units of output. It is assumed that doctors share the market equally and face identical individual demand functions. This assumption is not essential but will simplify the analysis. Demand is written $T = T(Q)$. While it seems fairly natural to assume that doctors' products are sufficiently differentiated that they face downward sloping demand functions, nothing here requires that, i.e., $\frac{\partial T}{\partial Q} \leq 0$.

Hospital charges are written as $H_o + aH$, where H is the length of stay and a is the cost per day for all costs which vary in proportion to length of stay. H_o represents all other hospital charges; that is, costs which are fixed with regard to length of stay. H_o would include charges for operating rooms, tests, blood, some drugs, etc. It is assumed here that a and H_o are treated as parameters by physicians. This assumption is natural enough for

---

[3]Atomistic behavior is not an entirely imaginary straw man. Harris provides a description of the surface structure of hospital organization which specifically characterizes doctors' decisionmaking as non-cooperative. Harris offers: "It is the constant non-cooperative scramble to expand one's own defensive position which drives the hospital to bigness and betterment." Although Harris' position is not entirely clear on this issue, he emphasizes the split between hospital administrations' and physicians' objectives. That emphasis clearly conflicts with Pauly's and Redisch's contention that the split is illusory.

atomistic case, but will not operate in the cooperative case. Also, no specific assumption is necessary regarding the determination of a and H, but the non-profit status of hospitals suggests that they are chosen so that average cost is covered. Once T is determined through choice of Q, the compensation to the doctors is determined as a residual: $P = T - H_o - aH$.

It is often argued with regard to medical services the conventional demand functions do not apply. If patients are fully insured, they have little incentive to avoid expenditures, so demand facing a physician could be perfectly inelastic. While zero elasticity is not assumed here, it is not fundamentally in conflict with this model. So long as there is some constraint which bounds T, the analysis goes through. Even under full insurance, an upper bound might be provided through review by third parties or peers.

A physician's private cost function is written C(Q,H). That is, the total cost of a medical practice is a function of the number of cure attempts performed and the number of hospital days used for each cure. We assume that marginal cost is positive, $C_Q(\quad) > 0$. Marginal cost may be constant or increasing, $C_{QQ} \geq 0$. The presence of H in the cost function is peculiar to this model and requires some discussion. The physician is aided in his practice by the use of some hospital resources. Clearly some medical procedures are impossible without hospital inputs. Others would be quite costly. For this reason it is assumed here that $C_H(Q,H) < 0$ for all H less than some H^*. Obviously this is not true for all types of practice. Where the assumption fails, physicians would not use hospitals. Since this is a model of hospital use, the assumption is not restrictive, such cases are simply excluded from consideration here. It is not necessary that all physicians face the same demand or cost functions. It is conjectured here, but again not necessary, that H* is large. Physicians will benefit in a number of ways from keeping their patients hospitalized. Their private requirements for office space and other facilities are reduced, they can order tests and drugs without much difficulty, and they can see patients at their own convenience. Further, the caring services of nurses and other professionals are to some extent substitutes for the physician's own inputs. We would expect $C_H(Q,H)$ to become positive at some H, because the physician must continue to visit a hospitalized patient and because the patient's condition may become adversely affected by an overly extended hospital stay.

With the assumptions made above, we may write the physician's maximizing problem as

$$(1) \quad \max_{H,Q} \pi = T(Q)Q - C(Q,H) - (H_o + aH)Q$$

which yields the first-order conditions:

$$(2) \quad \begin{aligned} &A) \quad \frac{\partial \pi}{\partial Q} = T(Q) - H_o - aH + QT'(Q) - C_Q(Q,H) = 0 \\ &B) \quad - aQ - C_H(Q,H) = 0 \end{aligned}$$

The first condition is conventional, that marginal revenue equals marginal cost. The second condition states that physicians' use of hospital facilities is such that the marginal reductions in their costs available from using hospital facilities equals the use price charged by the hospital.

The atomistic case assumes no specific accommodation by doctors to the scarcity of hospital beds. Where the constraint is binding, we might expect that beds are simply rationed on a first-come first-served basis, with perhaps a distinction made between elective and non-elective cases or between emergency and non-emergency cases. It might be more agreeable to build a model of intermediate naiveté, so that doctors anticipate the rationing and set price accordingly. However, to do so would require some specific assumption about doctors' perception regarding the basis of rationing. So long as the rationing scheme does not penalize doctors for length of stay, a less naive model would only alter condition 2A. Implications regarding length of stay are unchanged; once a patient is admitted the physician uses the hospital to maximize his private gain. On the other hand, any rationing scheme which did address length of stay could usefully be regarded as falling under the cooperative case.

Cooperative Behavior

The atomistic model of physicians' behavior presented above serves as the null hypothesis for this discussion. The alternative which is promoted here is that doctors behave cooperatively in their use of hospital resources. The form that this cooperation takes is not explicit, but the possibilities are numerous. Hospitals could have implicit or explicit quotas for doctors, doctors could adopt "standards of professional practice" or administrators could attempt to persuade physicians to conserve resources during periods of relatively high demand. There are a number of possible enforcement mechanisms including, of course, differing availability of beds. Doctors who behaved uncooperatively regarding length of stay might find that they will have difficulty in obtaining admission for their patients. Alternatively, the hospital could behave as the collusive agent through the appropriate choice of prices a and H_o. Then doctors could make decisions as in the atomistic case and still behave in a manner consistent with collective profit maximization. While such a solution would avoid certain problems of collusion, complications in distribution of profits would arise.

The cost and demand structure is similar to that used for the atomistic model. Here the maximization incorporates explicitly the scarcity of hospital beds. This is represented by the constraint $HQN \leq \bar{B}$ where H is again the length of stay per cure attempt, Q is the number of cure attempts, N is the number of doctors and $\bar{B}$ is the total number of patient days that the hospital could produce in a given time period. Doctors will maximize collective profits under this aggregate constraint. Throughout this analysis N is assumed to be exogenous.

In the atomistic case, a and H_o were taken to be parametric for the physician. Such an assumption is less natural for the cooperative case. If hospitals have significant fixed costs and if prices are set equal to average cost, then physicians should expect a to fall as the occupancy rate rises. Cooperative behavior should internalize these effects. Since the potential externality is pecuniary, appropriate behavior would "see through" the illusion of average cost pricing. The model is greatly simplified if we do the same. In what follows, hospital prices are ignored and doctors are assumed to behave as if they face the true costs of using hospital resources. For convenience, suppose that <u>variable</u> hospital costs can be adequately approximated by two components, costs which are fixed per case, $\tilde{H}_o$, and costs which vary with the length of hospitalization $\tilde{a}H$. These two components are analogous to H_o and a for the atomistic case. Thus hospitals are assumed to produce patient days at constant marginal cost, up to the maximum, where marginal cost becomes prohibitively large. Fixed hospital costs are treated as if they were paid by physicians through lump-sum transfers and otherwise are ignored here.

Under the present assumptions, it is convenient to think of the total payment for medical services as being paid to the physician, who then compensates the hospital for variable costs in each case. Fixed costs are covered by lump-sum transfers. Such a specification is faithful to the formulation by Pauly and Redisch. Again let T be the total payment for medical services so that the physician's compensation per case is now $\tilde{P} = T - \tilde{H}_o - \tilde{a}H$. Again assuming that all physicians are identical, we write the collective maximization for the physicians on a hospital staff as:

$$\max_{Q,H} \pi = N[T(Q)Q - C(Q,H) - (\tilde{H}_o + \tilde{a}H)Q]$$

To maximize this we write the Lagrangian

$$L' = N[T(Q)Q - C(Q,H) - (\tilde{H}_o + \tilde{a}H)Q] + \lambda(B-HQN)$$

Exploiting the assumption that N is exogenous, the problem is simplified by dividing through by N yielding:

$$L = T(Q)Q - C(Q,H) - (\tilde{H}_o + \tilde{a}H)Q + \lambda(\frac{B}{N} - HQ)$$

The necessary conditions for the representative physician are:

(5)

A. $\frac{\partial L}{\partial Q} = T(Q) + QT'(Q) - C_Q(Q,H) - (\tilde{H}_o + \tilde{a}H) - \lambda H = 0$

B. $\frac{\partial L}{\partial H} = - C_H(Q,H) - \tilde{a}Q - \lambda Q = 0$

C. $\frac{\partial L}{\partial \lambda} = (\frac{B}{N} - HQ) \geq 0 \qquad \lambda \geq 0 \qquad \lambda(\frac{B}{N} - HQ) = 0$

The first condition can be expressed more meaningfully as:

(5A') $$T(Q) + QT'(Q) = C_Q(Q,H) + \tilde{H}_o + \tilde{a}H + \lambda H$$

so that the left-hand side is recognizable as marginal revenue. The right-hand side is marginal cost plus a term which reflects the scarcity value of beds. Rearranging terms to get λ:

(6) $$\lambda = \frac{T(Q) + QT'(Q) - C_Q(Q,H) - \tilde{H}_o - \tilde{a}H}{H}$$

so λ equals the profits to be had from treating an additional case divided by the number of days required to treat a case. Under the assumption that all physicians are identical, λ can be interpreted as the system-wide shadow price of beds. Notice that λ is positive, so that from equation (5A') we can see that marginal revenue must be larger where the constraint on bed days is binding.

The second necessary condition may be rewritten

$$(5B') \qquad -C_H(Q,H)/Q = \tilde{a} + \lambda$$

Here the left-hand side is the effect on the physicians own average costs of increasing the hospital stay per cure. Where $\lambda = 0$, the right-hand side is simply the hospital costs generated by an increase in the typical hospital stay. Where the constraint is binding,the marginal savings to doctors from hospital use must be greater, which requires that H is smaller.

Equations (5A') and (5B') indicate the two testable contrasts between the atomistic and cooperative cases. The cooperative model implies that when hospitals are at or near capacity, physicians should respond by raising the total charge to patients and by reducing the average length of the hospital stay, each relative to outcomes away from the capacity constraint. The section below considers the observable counterpart of this model in order to provide a test of physicians behavior with regard to length of stay.

II. An Empirical Model of Length of Stay

As presented above the cooperative case would appear to suggest a simple relationship between the relative scarcity of hospital beds and the length of stay. There are, unfortunately, a few complications. There is first the problem of measuring the relative availability of hospital beds.

The second problem is simultaneity: availability of beds affects length of stay, length of stay affects availability. These concerns, which would arise in any empirical implementation of this model, are considered below. The specifics of the estimation are discussed in Section III.

The micro model presented above suggests that physicians will behave differently when the hospital constraint is binding. A very literal empirical interpretation would suggest that individual hospitalizations be examined and related to the economic circumstances which prevailed at the time of hospitalization. In addition to imposing outrageous requirements for data, such an approach would probably not be an entirely faithful representation of the economics of the situation. The approach taken here is to aggregate over time by representing the constraint on hospital beds through the annual average occupancy rate for hospitals (the unit of observation will be a state, as is discussed below). There are several advantages to this aggregation over time. First, it should approximate to a degree the doctors' awareness of the constraint. It is unlikely that daily fluctuations in availability would have an immediate effect on behavior, except under emergency conditions. While it is assumed here that doctors behave as if they pursue collective profit maximization, it is likely that they rely on simple signals rather than explicit and frequent collusion. Second, rational behavior would not reflect just instantaneous availability, but also the implications of decisions made today on future availability. Third, adaptations to economic conditions are likely to be, in large part, adjustments to the conventions which constitute standard medical practice. A final consideration is the ease of collection and processing of the data.

This representation requires an assumption about the connection between the occupancy rate and the scarcity of hospital beds. Such a connection seems intuitive enough, but can be further supported. A high occupancy rate should be interpreted by physicians as predicting a high probability that the constraint is binding. Second, the simple measure of capacity that is generally used (number of beds multiplied by the number of days) may overstate the real capacity of hospitals. As this "accounting" capacity is approached, difficulties in scheduling both patients and staff begin to arise. Thus, presence of an occupancy rate that is high, though still less than one, may reflect conditions of excess demand.

The implications of the two models in terms of occupancy rates is represented in Figure 1. The line AA' represents the behavioral relationship between occupancy rate and length of stay in the cooperative case. Physicians respond to conditions of relatively high demand by shortening the average length of stay. Behavior in the atomistic case is represented by the horizontal line BB', where length of stay is invariant with respect to occupancy rate.

A minor variant of the atomistic model would allow for hospitals to bill at average cost. If average cost falls as occupancy rate rises, then equation (2B) requires that physicians choose longer stays when occupancy rates are high. This would require that the behavioral relationship is upward sloping. The empirical implications of the atomistic and cooperative models remain distinct under these assumptions.

The determining relationship is an identity involving length of stay and occupancy rate. The relationship is

$$(7) \qquad R \equiv \frac{A \cdot L}{B} \quad \text{or} \quad L \equiv \frac{BR}{A}$$

where R is the occupancy rate, L is the average length of stay, A is the number of admissions, B is the number of patient days which hospitals could produce in a given period. Where the number of admissions can be taken as exogenous, as is appropriate in the atomistic case or in the cooperative case when demand price elasticity is zero, the identity can be represented by straight lines like OD or OC in Figure 1. For the cooperative case, one avenue of adjustment is in the doctor's price, which influences the number of admissions. As the occupancy rate increases, the number of admissions falls, so the relationship will not be linear, as shown in OF. Solutions are represented by intersections of the behavioral relationship (AA′ or BB′) with the identity (OC or OD). These intersections (a,b,c,d) are stable. A point like f could not be sustained: Physicians would be choosing a length of stay that was not optimal and would curtail stays to lower the occupancy rate.

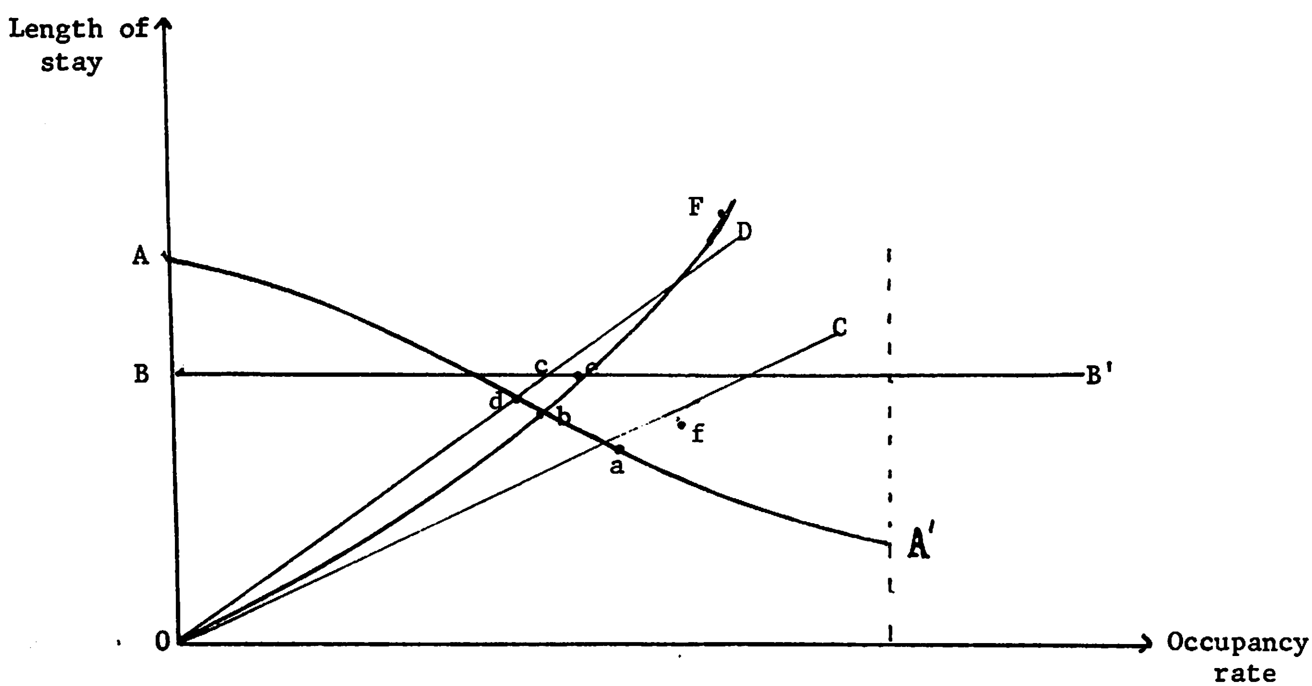

Figure 1

Shifts in the curves due to exogenous changes are represented in Figure 2 for the atomistic case and Figure 3 for the cooperative case. Increases in insurance coverage will increase admissions and therefore rotate the identity downward (from OD to OC or from OH to OG). Increased insurance coverage would also decrease the marginal cost of hospital days to patients so the behavioral relationships would shift upward (from BB to MM' or from EE' to NN'). Better insurance coverage therefore has an ambiguous effect on length of stay, while unambiguously increasing the occupancy rate, <u>ceteris paribus</u>. Increasing the number of doctors will also have ambiguous effects on length of stay. If more doctors lead to more procedures, the identity will rotate downward. The effect on the behavioral relationship is unclear. If an increase in the number of admissions implies that the average case is less serious, then the behavioral relationship (MM' or NN') is shifted downward. So, for example, an increase in the number of doctors would move the equilibrium from a point like <u>a</u> to a point like <u>b</u> in Figure 3. If however increased presence of physicians is associated with more ambitious or complicated treatments, the behavioral relationship shifts upward, moving the equilibrium from c to d in Figure 3.

A central concern here is the relationship between the number of beds and length of stay. Increasing the number of beds rotates the identities toward the vertical axis. There is movement along the behavioral relationship, but no movement of that relationship itself. So for the cooperative case, increasing the number of beds unambiguously increases the average length of stay. Notice that there is no effect on length of stay in the atomistic case (where average cost pricing is assumed, increasing the number of beds will decrease length of stay).

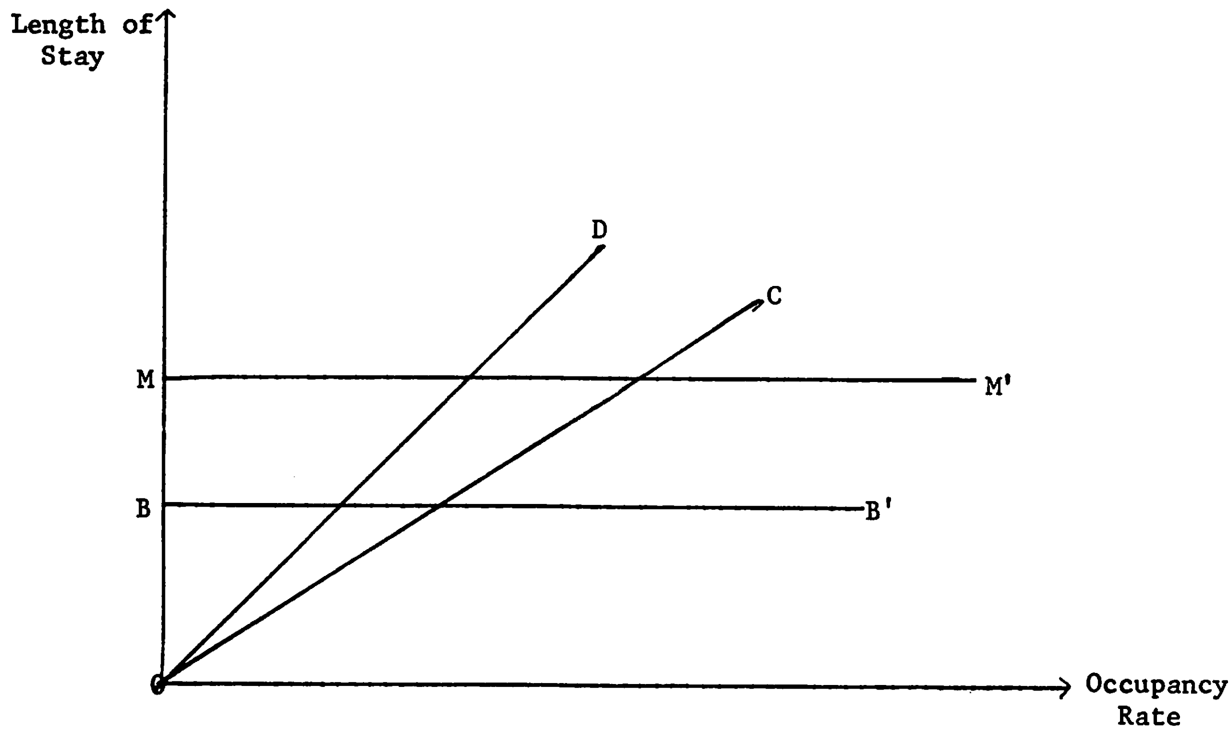

Figure 2

Shifts in the Atomistic Case

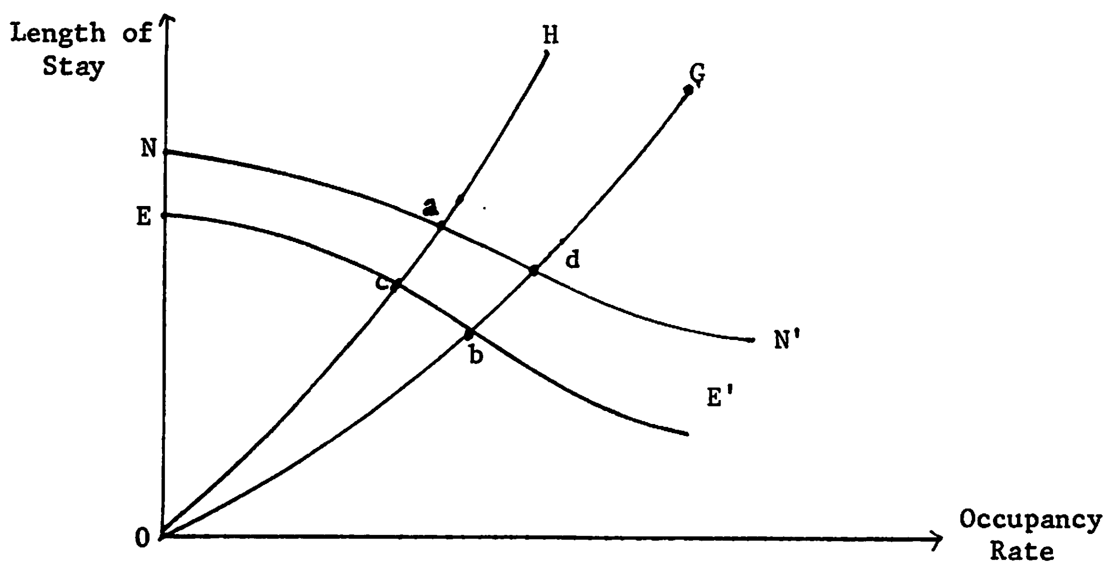

Figure 3

Shifts in the Cooperative Case

Given the concern here for length of stay, the empirical test of the model consists of estimating the reduced form relationship between the number of beds per capita and the average length of stay. An alternative would be direct estimation of the behavioral equation shown in Figures 1 to 3. Such a procedure would add needless complexity in estimation. As estimated, the reduced form can readily be interpreted and provides a clear test of the model and yields results which bear directly on the policy discussion. Simple estimation of the relationship between occupancy rate and length of stay does not provide a test of the model. Observations would be the intersections as in Figures 2 and 3, so that the behavioral relationship is not revealed by simple regression of these two variables.

The empirical specification developed above facilitates discussion of the existing empirical studies of length of stay. In light of the present model, these studies can be usefully divided into two groups, those which use occupancy rate and those which use the per capita endowment of beds as independent variables.

Two studies which are representative of the first group are those by Rafferty (1971) and Davis and Russell (1977). Rafferty studies the effect of occupancy rate on case mix, concluding that as a hospital's occupancy rate rises, resources are concentrated on more serious illnesses which happen to have longer average lengths of stay. Length of stay is not a central concern and is implicitly assumed to be exogenous for any given diagnosis. There is no mention in Rafferty of the simultaneity problem. Data for that study were from two hospitals in a single city. Occupancy rates for the individual hospitals were used to explain the hospital's case mix. The Davis and Russell study also falls prey to the identification problem. They find a positive coefficient on occupancy rate in regressions which explain length of stay. They comment on the simultaneity problem to explain the unexpected positive coefficient on occupancy rate. Again, length of stay was not a central concern in their paper.

Two studies which have used hospital beds per capita as explanatory variables are those of Feldstein (1971) and Olowokure (1978). Feldstein's study is an interesting contrast to this one. Feldstein's study predates Pauly and Redisch and treats the hospital's objectives as an amalgam of the objectives of hospital employees. The stylized maximand is "quality of care as perceived by the hospital bureaucracy".[4] Feldstein explicitly assumes that occupancy rates are exogenous in a model which determines, among other things, length of stay. This of course rules out any simultaneous determination of the two variables. With Feldstein's formulation, beds per capita enter directly into demand functions which are estimated using an instrumental variables technique. This contrasts with the present specification in which changes in the number of beds induce movement along the "demand" curve but no movement of the curve itself. Under the assumptions of this paper a demand equation would include occupancy rate rather than beds per capita. If it were assumed here that the constraint on bed days is always binding, the concern with occupancy rate would be unnecessary. First order conditions would define a demand function and an empirical representation more like Feldstein's would be appropriate. In that case, however, it would be questionable why the number of beds should enter into the demand for length of stay. The estimates which follow do include beds per capita, but the equations presented here are interpreted as reduced forms.

---

[4] In principle, the quality-quantity hypothesis is silent on matters of length of stay, so long as this objective is ascribed to hospital administrators alone. If physicians are held to be merely the customers of the hospital, the atomistic model would seem to apply under quantity-quality maximization. Even where physicians are assumed to participate in maximizing quality-quantity, the implications are unclear, since quality and quantity are not well defined (for example, see Newhouse, pp. 248-249).

The study by Olowokure analyzes mean length of stay in acute care hospitals in England. Olowokure employs an extensive list of regressors including, in various equations, some or all of: beds per capita, occupancy rate, hospital size and percent of all patients who are long-stay patients. Because the unit of observation is a hospital, a number of case mix variables are employed. Beds per capita is the sole survivor among the availability measures in an eclectic approach to specification. Beds per capita appears to retain significance in all specifications, although interpretation is complicated by the presence of some duplication among the regressors. One interesting outcome in Olowokure's paper is that hospital size has no significant explanatory power in equations which include case mix and availability measures.

III. Testing the Model

The test of the model consists of regressing the average length of stay on the number of beds and an appropriate list of ceteris paribus variables. A significant positive relationship between the number of hospital beds and the average length of stay would stand as evidence in favor of the cooperative model. Without such a significant relationship we would be unable to reject the null hypothesis that physicians behave atomistically.

The data consist of observations on the 50 states plus Washington, D.C. This spatial aggregation is necessitated by a number of data concerns. Compatible data on doctors, surgeons, and insurance coverage are not readily available for aggregate units smaller than states. Further, it would be difficult, if not impossible, to specify an appropriate level of aggregation that is much smaller than a state. For individual hospitals it would be impossible to identify the populations served. A similar question would arise even at the level of an SMSA, since some urban hospitals will serve a substantial non-urban population. It is likely that people frequently travel across SMSA boundaries in seeking hospital care. This problem is largely avoided by aggregating to the level of states.[5]

The key <u>ceteris paribus</u> variables that are used are: surgeons per capita, primary care physicians per capita, all physicians per capita, and the fraction of the population insured for hospital expense. Additional variables were constructed for special tests, as is noted below. Because of obvious collinearity problems, not more than one variable reflecting the presence of physicians is used in any regression.

Variation in the number of hospital beds per capita is assumed to be exogenous to the model. Variation may come about due to differences in the preferences of licencing boards, in wealth, in the influence of doctors, or by historical accident. We should expect that rapidly growing areas, for example, have fewer beds per person than older, established regions. Alternative hypotheses, which treat beds per capita as exogenous, are considered briefly at the end of this section.

The specification employed reflects differences in the presence of medical professionals and the costs confronting individuals for care (through the insurance variable). There are three maintained hypotheses behind this specification. First, it is assumed that the relative prices of hospital inputs are uniform throughout the United States. Second, it is assumed that

---

[5] The approach taken in Roemer's pioneering study provides a contrast to the one taken here. Roemer considers responses to a change in bed availability in a single hospital. That approach required careful examination of individual patient's hospital records to assure that the region served by the hospital did not change after expansion occurred.

at this level of aggregation, case mix differences average out. Finally it is assumed that the available medical technology is uniform throughout the United States. (Employed techniques may, of course, vary.) The last two assumptions are reasonable only in the cross-section experiment performed here. In time series, both ambient health and available technology would vary throughout the sample. Some testing of these restrictions is possible, as is shown below.

Results are given in Table 1 for data pertaining to all non-federal hospitals. Equation (1) is submitted as the best result in terms of both specification and performance. With the available data, a number of alternative specifications are defendable, and several of these are shown as well. Equations (2) and (3) differ only by the use of surgeons and primary care physicians as alternatives to all physicians to reflect the presence of doctors in the population. Results regarding the sign and significance of the coefficient on beds per capita are consistent across the specifications tested, although the magnitude of that coefficient varies somewhat across specifications.

While log forms appear to perform best among specifications, no striking differences appear when other functional forms are tried. Representative examples of both linear and semi-log forms are shown for comparison.

Among the maintained hypotheses is the assertion that the relative prices of hospitals' inputs do not vary systematically across the sample. The variable RWAGE, total hospital payroll per employee divided by average state income, is entered to test (or relax) this hypothesis. The results shown are broadly consistent with the maintained hypothesis; however, the specification of the variable, limited by data availability, may not be ideal. Use of an undeflated hospital wage produces very similar results. It is assumed that non-labor hospital inputs are sold in a national market, so that relative

Table 1

Ordinary Least Squares Estimates

EQ. NO. / DEP. VAR.	(1) LLOS	(2) LLOS	(3) LLOS	(4) LLOS	(5) LLOS	(6) LOS	(7) LOS
VARIABLE							
LBEDSPER	.796 (5.5)*	.724 (5.1)*	.781 (5.2)*	.797 (5.4)*	.802 (5.6)*		8.71 (5)*
LDOCSPER	.262 (2.7)*			.263 (2.7)*	.261 (2.8)*		3.73 (3.2)*
LPRIMSPER			.229 (1.91)				
LSURGSPER		.326 (3.25)*					
LINSR	-.143 (-.85)	-.161 (-.98)	-.091 (-.52)	-.139 (-.80)	-.151 (.90)		-2.07 (1.00)
LRWAGE				.042 (.132)			
DOCSPER						2.34* (2.1)	
BEDSPER						1234 (4.3)*	
INSR						5.6 (3.8)*	
LNURB65					-.095 (-1.5)		
CONSTANT	6.28 (9.1)*	6.65 (9.9)*	6.43 (9.0)*	6.27 (9.0)*	6.04 (8.6)*	4.18 (3.1)*	53.6 (6.4)*
R^2	.61	.63	.58	61	63	.47	.58
F	(24.4)	(26.7)	(21.6)	(17.8)	(19.3)	(13.8)	(21.4)

t statistics in parentheses.

*Significant at $\alpha = .01$ in two tail test.

Table 1 (cont'd.)

Definitions

All Data for 1974

LOS	-	Average length of stay, total patient days divided by total admissions (AHA Hospital Statistics)
BEDSPER	-	Hospital beds per capita (Hospital Statistics and U.S. Statistical Abstract)
DOCSPER	-	Doctors per capita (U.S. Statistical Abstract 1976)
SURGSPER	-	Surgeons per capita (U.S. Statistical Abstract)
PRIMSPER	-	Primary care physicians per capita (U.S. Statistical Abstract)
INSR	-	Fraction of population with insurance coverage for hospital expense (Sourcebook for Health Insurance Data)
NURB65	-	Nursing home beds per person aged 65 and over. (U.S. Statistical Abstract, 1976). (Data are for 1973, data for 1974 appear not to be available.)
RWAGE	-	Total hospital payroll divided by total hospital personnel, all divided by state per capita income. (Hospital Statistics and U.S. Statistical Abstract.)

labor costs should reflect variations in input costs. (This variable does perform as expected in the two-stage equations reported below.)

Equation (5) was estimated to test for a possible measurement error which might occur if there were differences across states in the way convalescing patients were accommodated. If, for example, some states had few nursing home beds and treated convalescing patients in hospitals, those states would have many hospital beds and long average stays. If such differences were important, the key results shown here might be entirely a consequence of an aggregation problem.[6] With this in mind, the number of nursing home beds per person over the age of 65 was included in a regression. The coefficient has a negative sign (though not significant) as is consistent with the existence of the hypothesized measurement problem. However, inclusion of this variable does not weaken the results on beds per capita; in fact, that coefficient becomes slightly larger. Thus the variation of nursing home beds appears not to introduce an important bias to these results.

As is shown in Section II, atomistic behavior predicts that length of stay will not depend on the availability of beds. These results clearly invalidate the atomistic hypothesis and confirm the cooperative model. The specific values estimated here are of some importance. For equations (1)-(5), the coefficients generated are elasticity estimates. Thus it appears from these equations that variation in length of stay will accommodate between 70 and 80 percent of variation in the per capita endowment of hospital beds.

Alternative Hypotheses

There remain alternative explanations for the associations reported above. While an attempt to anticipate all the possibilities would likely be futile,

---

[6] I acknowledge a referee for gently pointing out this possible problem.

two fairly natural explanations are compelling enough to warrant further exploration. These are considered below.

In the foregoing it is maintained that there are no systematic differences across states in the prevailing levels of health. All that is really required for the tests and interpretations made here is that differences in health are not systematically associated with differences in the endowment of hospital beds. An alternative hypothesis which comes readily to mind is that states do vary in terms of the potential (or demand) for medical intervention, and that availability of hospital resources are consciously adjusted to these different circumstances. If, for example, residents were especially sickly in a particular state, then we should expect both more hospital beds and longer stays. In this model, hospital beds and hospital stays are mutually affected by the same variable, so they are correlated, yet one does not cause the other. This hypothesis would appear to explain away the relationship between hospital beds and length of stay without requiring any particular behavior on the part of physicians.

Fortunately the two models are not observationally equivalent. If the connection between availability of beds and length of stay were the consequence of their mutual dependence upon the health of the local population, then we should expect a significant positive correlation between admissions per capita and length of stay. This implication is quite clearly contradicted. The simple correlation between average length of stay and admissions per capita is -.15.

As a further test, both beds per capita and admissions per capita (ADSPER) were used in a regression with length of stay as the dependent variable. The results are shown here with t statistics in parentheses.

$$\underset{}{LOS} = 14.75 + \underset{(25.2)}{2054} \cdot BEDSPER - \underset{(19.3)}{103775} \cdot ADSPER$$

$$R^2 = .93$$

The "varying sickness" hypothesis, by itself, would require a positive coefficient on ADSPER. The coefficient is negative and significant. Notice also that the equation above is not a correct reduced form under the hypothesis of this paper, since admissions are argued to be interdependent with beds per capita.

A second possibility, closely related to the mutual causality argument, is a reverse causality of simultaneity argument. That is, perhaps there are cross sectional differences in choices regarding lengths of stay, determined by differences in income, insurance coverage, taste, etc. Further, it might be that the number of beds was actively chosen to accommodate these differences. If this were the case, the observed association of length of stay and beds per capita would arise quite apart from the institutional-behavioral assumptions posited here. To check this possibility, two-stage least squares estimates were computed treating beds per capita as an endogenous variable.[7]

The two-stage estimates are shown in Table 2. The most interesting results here are that the coefficients of the log of beds capita become slightly larger, which conflicts with the simultaneity hypothesis. Otherwise, these results are fairly closely consistent with the OLS estimates. A notable change is that the coefficient on the wage variable is now more consistent with standard expectations. The only discordant note is that the coefficients, which are fairly insensitive to specification change in the OLS estimates, are fairly sensitive to specification changes in 2SLS. This problem is known to be

---

[7]It is important to maintain the distinction between this <u>possible</u> simultaneity (of beds and length of stay), which is argued to be peripheral, from the simultaneity of occupancy rates and length of stay, which is argued to be central. In the case of the latter, problems of simultaneity are avoided by estimating the reduced form.

Table 2

Two Stage Estimates

	1	2	3	4
Dep Var	LLOS	LLOS	LLOS	LLOS
Variable				
LBEDSPER	1.01 (4.0)	.815 (3.4)	1.03 (4.0)	.810 (3.3)
LDOCSPER	.258 (2.6)		.256 (2.6)	
LSURGSPER		.315 (3.0)		3.11 (2.8)
LINSR	-.310 (-1.3	-.222 (-1.1)	-.330 (-1.4)	-.255 (-1.2)
LRWAGE			-.039 (-1.1)	-.27 (-8.4)
LNURB65			-.10 (-1.5)	-.070 (-1.0)
CONSTANT	7.32 (6.0)	7.06 (4.4)	7.12 (5.8)	6.91 (6.30)

Variable definitions follow Table 1.

Computed t statistics are in parentheses (only asymptotic distribution of the estimates are known).

LBEDSPER is endogenous. Exogenous variables, in addition to those listed in the equations, are: LRWAGE, LNURB65 (see Table 1); an east-west regional dummy variable, the natural log of population growth since 1960 (U.S. Statistical Abstract 1976); and the log of the fraction of population which is in metropolitan areas (U.S. Statistical Abstract 1976). To avoid problems with log (0), percent urbanized for Wyoming and Vermont set at 0.1.

a characteristic of 2SLS in general. The increased sensitivity to specification together with the absence of any evidence of simultaneity, argue in favor of the OLS estimates over the 2SLS results. However, the results of the two sets of equations are equivalent for hypothesis testing and provide point estimates which are reasonably consistent.

Conclusion

As one of the most important and best measured of non-profit institutions, hospitals have attracted the attention of social scientists who seek to build a theory of non-profit organizations. This paper provides fairly strong evidence in favor of the physicians' cooperative model. Length of stay appears to adjust in ways which are consistent with collective profit maximization for physicians.

Because of the closeness of this work to a very active policy issue, the regulation of hospitals, it is perhaps warranted to exercise more than the usual modesty in presenting the empirical results. While these results are quite satisfying for hypothesis testing, direct use of these point estimates for policy purposes may be premature. First, it should be noted that the confidence intervals are fairly wide. Second, the results presented offer no explanation of how length of stay is shortened as bed availability is reduced. Shortened stays might reflect shorter stays for given treatment or selection of simpler, quicker treatments. Finally, the level of aggregation, over space, time and types of hospitals, may not be ideal. The alternatives, however, are also flawed.

In spite of these concerns, the results serve as an important demonstration of the sensible proposition that the effect of hospital regulation will depend upon the institutional environment in which these regulations operate. Regulators of hospitals have sought to reduce the number of hospital beds in order to reduce the amount of medical intervention consumed by a supposedly over-insured population. If this constraint operates primarily on length of stay, which is shown here to be a possibility, then the effects of such regulation will differ markedly from its intent.

Bibliography

American Hospital Association, Hospital Statistics, Chicago 1975.

K. Davis and L. B. Russell, "The Substitution of Hospital Outpatient Care for Inpatient Care," Rev. Econ. Statist., May 1972, 54, 109-20.

M. S. Feldstein, "Hospital Cost Inflation: A Study of Nonprofit Price Dynamics," Amer. Econ. Rev., Dec. 1971, 61, 853-71.

J. Harris, "The Internal Organization of Hospitals: Some Economic Implications," Bell J. Econ., Autumn 1977, 8, 467-82.

Health Insurance Institute, Source Book for Health Insurance Data, New York, 1975.

J. L. Migué and G. Bélanger, The Price of Health, Toronto, 1974.

J. P. Newhouse, "Toward a Theory of Non Profit Institutions: An Economic Model of a Hospital," Amer. Econ. Rev., March 1970, 60, 59-70.

T. O. Olowokure, "Variations in Average Length of Stay Among Acute Hospitals," Applied Econ., March 1978, 10, 1-10.

M. Pauly and M. Redisch, "The Not-for-Profit Hospital as a Physicians Cooperative," Amer. Econ. Rev., March 1973, 63, 87-99.

J. A. Rafferty, "Patterns of Hospital Use: An Analysis of Short Run Variations," J. Polit. Econ., Jan.-Feb. 1971, 79, 154-65.

M. I. Roemer, "Bed Supply and Hospital Utilization: A Natural Experiment," Hospitals, Nov. 1, 1961, 35, 36-42.

F. I. Sloan and B. Steinwald, "Effects of Regulation on Hospital Cost and Input Use," J. Law and Econ., April 1980, 23, 81-109.

U.S. Bureau of the Census, Statistical Abstract of the United States: 1976, Washington, D.C., 1976.