Working Paper

ADDING DEMAND, INCENTIVES, DISEQUILIBRIUM, AND DISAGGREGATION TO HEALTH CARE MODELS

M.V. Pauly

January 1981
WP-81-4

International Institute for Applied Systems Analysis
A-2361 Laxenburg, Austria

NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

ADDING DEMAND, INCENTIVES, DISEQUILIBRIUM, AND DISAGGREGATION TO HEALTH CARE MODELS

M.V. Pauly

January 1981
WP-81-4

Working Papers are interim reports on work of the International Institute for Applied Systems Analysis and have received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute or of its National Member Organizations.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

THE AUTHOR

Mark Pauly is a professor of economics at Northwestern University whose scientific interests include the modeling of patient behavior in a health insurance system, the theory of agency as applied to physicians, and the development of theoretical and econometric models of hospitals. He spent the summer of 1980 with the Health Care Task at IIASA working on microtheoretical behavioral models for health planning.

FOREWORD

The principal aim of health care research at IIASA has been to develop a family of submodels of national health care systems for use by health service planners. The modeling work is proceeding along the lines proposed in the Institute's current Research Plan. It involves the construction of linked submodels dealing with population, disease prevalence, resource need, resource allocation, and resource supply.

This paper analyzes the work completed in the Health Care Systems (HCS) Task and in particular looks at the application of the Disaggregated Resource Allocation Model in an economy where HCS resource allocation is determined by patient demand and is not necessarily in equilibrium with supply.

Related publications in the Health Care Task are listed at the end of this report.

Andrei Rogers
Chairman
Human Settlements
and Services Area

ABSTRACT

This paper investigates the consequences of adding a demand constraint to models of the response of the health care system (HCS) to alternative levels of resources. Disequilibrium econometric techniques are shown to be applicable, and the model is extended to include the possibility of demand creation by physicians and referral by generalists to specialists. It is also shown that incentives affect the response of use of the HCS to resources.

CONTENTS

ADDING DEMAND, INCENTIVES, DISEQUILIBRIUM, AND DISAGGREGATION TO HEALTH CARE MODELS

1. INTRODUCTION

The purpose of this paper is to describe some possible additions to and modifications of the health systems models at IIASA (described in Shigan et al., 1979a) in order to broaden the applicability of those models. In particular, the paper will suggest additional ways of estimating or forecasting resource use and allocation which are especially applicable to countries with decentralized or market oriented health care systems (HCS's). These suggestions indicate how medical care forecasting models such as IIASA's Disaggregated Resource Allocation Model (DRAM) (Hughes and Wierzbicki, 1980) could be modified and extended to systems in which markets are used to allocate some medical resources. They also indicate how demand-based models (e.g., Newhouse, 1974) could be adapted to centrally planned, supply-constrained medical care systems.

The model to be discussed is of a Disaggregated Incentive and Demand Disequilibrium (DIDD) type. It will have the following features that distinguish it from DRAM.

(i) DIDD permits use and resource allocation to be determined by patient demand as well as by provider preferences. It does not assume that demand (at a given quality level) is insatiable, or that excess demand always prevails.

(ii) DIDD proposes an econometric approach to deal with the problem of forecasting when some observations may represent disequilibrium (either excess demand or excess supply), while others represent equilibrium.

(iii) DIDD permits the possibility that observed output and use may depend on incentives to providers, both in the sense of incentives to modify patient demands and in the sense of incentives for increasing productivity given demand.

(iv) DIDD assumes that decisions made by individual enterprises in the system do not necessarily correspond to those desired by a central authority or by any single agent. Rather, the system should be modeled as the interaction of numerous disaggregated individual agents (all assumed to be maximizing utility subject to some constraints), whose objectives, incentives, and constraints may differ.

These suggestions are consistent with the criticism of DRAM by Rutten (in Shigan et al., 1979b, p.141):

> Only in subsectors of the system which are in large part supply determined (in the restricted sense that consumption depends on resources) will this model perform well. But major changes in organization, financing, etc., might affect the model predictions, even in these subsectors, since the model does not take these factors into account. Given the elegant mathematical framework of DRAM and the possibility to disaggregate, it should be investigated if the above-stated disadvantages of this model could be overcome.

In this paper, I will first provide a fairly detailed critique of DRAM in order to motivate the modifications and additions which follow. Then I will present an intuitive explanation of the importance of demand and disequilibrium. Next, I will discuss several more detailed econometric models which try to come to grips with the issue of adding demand and incentives for disaggregated agents. I will also deal with the issues of incentives for production efficiency and incentives to "create" demand. I will close with some observations on empirical applications of the model.

2. THE DRAM MODEL OF HEALTH CARE RESOURCES: AN ECONOMIC CRITIQUE

2.1. Introduction

DRAM represents an attempt to forecast the use of outputs of the medical care system (such as hospital admissions, days of stay, and doctor consultations). These outputs are assumed to be generated by a process that operates as if a single decisionmaker maximized a utility function subject to a resource or budget constraint (or with resource use negatively affecting unconstrained utility). This decisionmaker's utility is thought to depend on the closeness of actual outputs to ideal standards and (sometimes) on money costs, and is maximized subject to a constraint on physical resources used to produce those outputs and (sometimes) on money costs (Gibbs, 1978a, Gibbs, 1978b, Hughes and Wierzbicki, 1980). The model is basically one of rationing, but with choices made as if by a single decisionmaker. When resources are insufficient to achieve <u>either</u> ideal levels (as far as the decisionmaker is concerned) or actual quantities demanded(as far as patients or perhaps physicians are concerned), what kind of outputs in what amounts does the decisionmaker decide to give to what kinds of people?

In a formal sense, the idea of utility maximization subject to a budget constraint is the paradigm of almost all neoclassical economic theory. But one question is rather whether such an "as if" model, and one with this particular set of characteristics, is likely to be the best way of modeling a <u>system</u>.

An example may help to explain the economist's apprehension.

One could describe competitive equilibrium as the result of a single decisionmaker's utility maximization subject to the economy-wide resource constraint. But this characterization would not be as useful as the model of competitive equilibrium because

(a) it fails to model the behavioral response of the multiple agents who are actually in the system to the decentralized incentives they face, and, more importantly,
(b) it is not likely to be able to track or predict very well if the determinants of supply and demand vary over time or across areas or if resources are redistributed.

While one might continually recalibrate the model or add things to the "utility function", such a model could soon become either hopelessly unwieldy or hopelessly inaccurate.

Somewhat the same comment would apply to the health care system, at least in decentralized countries. There is no single decisionmaker; there is demand and supply. Of course, this market is not perfectly competitive, and one of the great controversies in health economics has been concerned with the question of how to model a system if it is neither perfectly competitive nor perfectly monopolistic.

With this word of caution in mind, let us now turn to a more specific consideration of the assumptions underlying DRAM.

2.2. Demand

Surely the most questionable feature of DRAM is the postulate that demand is always and everywhere unsatisfied (and therefore irrelevant); that use is wholly determined by HCS preferences and standards.

> Furthermore, it appears that in <u>none</u> of the places studied (U.S. and U.K.) has the supply of beds reached the level at which in-patient care is given to all individuals who seek it, at the ideal average length of stay. (Gibbs, 1978a, p.5.)

This observation is repeated in the most recent statement:

> It has been widely observed that the demand for health care seems to be insatiable. (Hughes and Wierzbicki, 1980, p.1.)

While such an observation may correctly characterize the situation in the U.K., in which user price is zero and supply has been constrained for decades, one need not undertake sophisticated statistical studies to discover that it is not universally true. With nonzero (though low) hospital user prices and unconstrained supply, the United States finds itself with low average occupancy rates (less than 80 percent for the country as a whole and rates as low as 63 percent for voluntary hospitals with fewer than 100 beds). Surprisingly, some of the lowest occupancy rates occur in the West, where there is also a relatively low number of hospital beds per capita. In a similar way, Canada with a zero money user price but with a supply of beds which, for historical reasons, is fairly generous, also finds itself with significant numbers of empty beds. (Note also that, while low occupancy rates are clearly inconsistent with the assumption of unsatisfied demand, even high rates are possibly consistent with satiation if supply has been set just equal to demand.) As a general proposition then it is simply false to assert that "nowhere have the demands for beds been saturated..." (Rousseau and Gibbs, 1980) and a model built on such an assumption is not going to be universally applicable.

If the proposition is untrue in general for hospital beds, it appears to be even less true for physicians' services. A great concern in the United States is the purported surplus of physicians in certain specialties, especially surgery and medical specialties. Indeed, the "demand creation" literature (discussed further below) depends critically on the assumption that physicians are willing and indeed eager to supply more of their own services at current gross fee levels than consumers would demand if they were given accurate advice. Lest one think that such phenomena only characterize private sector supply, it should be noted that the number of visits per physician for publicly salaried physicians sent to supposed scarcity areas, under the National Health Service Corps programs, is exceedingly low: less than half of the average output levels of a private-practice general practitioner (GP). This either means that public physicians are very unproductive or that, even in supposedly high need areas, there is insufficient demand.

Why did the DRAM modeling effort adopt the postulate of insatiable demand? I suspect that the problem was caused by an uncritical interpretation of the literature; an interpretation to which some loose language by health service researchers has unfortunately contributed, and an interpretation reenforced by the intent to model the British situation. There are actually two propositions in the literature (both theoretical and empirical). One of them is surely true, while the other is plausibly true sometimes and for some types of care, but is subject to considerable dissent and qualification, and is easily misinterpreted.

Surely the true proposition is this: if supply is constrained sufficiently(and sufficiently below any plausible level of demand), then observed or, in Feldstein's terminology, "manifest" demand, will be determined by supply. The distribution of that supply over alternative types of output is then subject to the discretion of the HCS. This was the case that Feldstein (1967) said he was modeling for the U.K. in his classic study of the relationship between hospital beds and use, precisely because it was plausible to argue that the supply of beds was well below the actual demand for them in the United Kingdom. Where supply is less than the actual quantity demanded, it is surely going to be the case that increases in bed supply will be matched by increases in observed quantities used or demanded. This is not, however, "demand creation" in any meaningful sense.

The second argument is that the demand function itself is shifted by changes in supply. That is, the amount of care that consumers demand, or can be persuaded to demand, is somehow influenced by the presence of supply. Figures 1 and 2 show the distinction between these two arguments. D_0 in Figure 1 shows a given demand curve, with quantity demanded as an inverse function of user and/or time price.

Suppose supply is initially at S_0, but total price (user price plus time price) is at P_0. Then there will be excess demand in the amount S_0Q_1, which the system will somehow have to ration. (It should not necessarily be supposed that physicians and hospitals will actually choose to ration out the least valuable inputs S_0Q_1, although that is what DRAM implies). Now let the budget be increased so that supply can be expanded from S_0 to S_1; observed aggregate use or "demand" will increase proportionately from S_0 to S_1. At least up to the resource level needed to supply Q_1 units, demand is unsatiated. It is this situation which DRAM properly (and elegantly) describes.

To see the second, more controversial case, we begin in equilibrium at, say, S_1, D_1, and P_1 in Figure 2. Now let supply be increased to S_2. The argument cited in support of the DRAM assumption (Roemer, 1961, Roemer and Shain, 1959; Harris 1975) is that this increase in supply in itself will somehow cause a shift in the demand schedule from D_1 to, say, D_2. A serious controversy in North America is whether one can conclude that this sort of behavior -- which is quite at variance with any kind of neoclassical economic model -- does in fact occur to an important extent.

It is generally agreed by health economists that the mere finding of a high positive correlation between supply and use (the evidence in the above-cited studies) is not sufficient to establish the proposition that supply is "creating" demand. The problem, as Rosenthal pointed out as early as 1964 (Rosenthal, 1964) is that the observation is equally consistent with demand "creating" supply, since

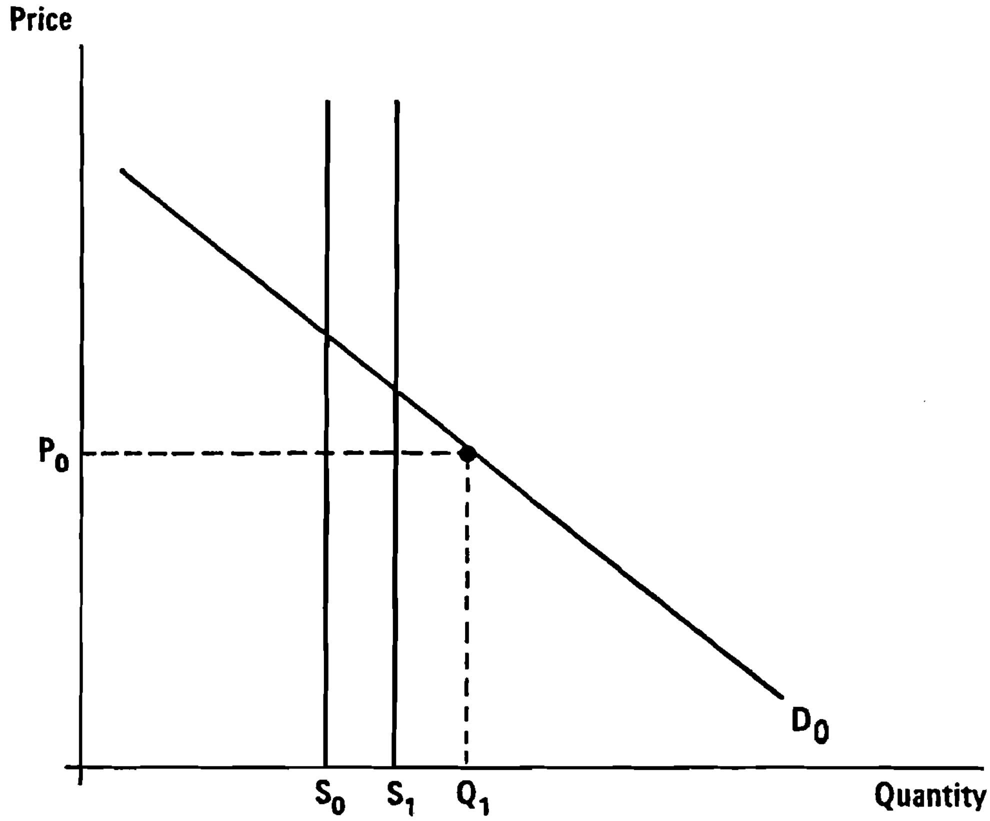

Figure 1. Excess demand and an increase in supply.

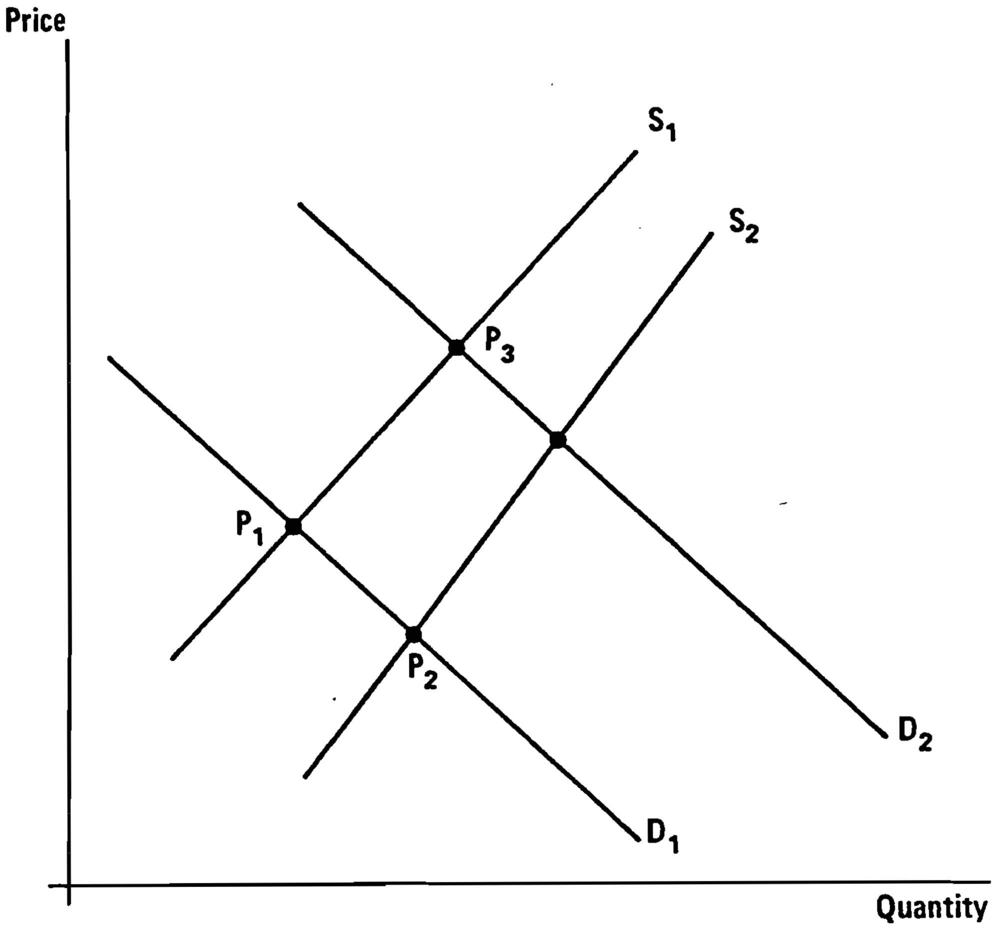

Figure 2. Changes in demand and price and an increase in supply.

in equilibrium, quantity supplied equals quantity demanded. Slightly more formally, as shown in Figure 2, we could be observing not "demand creation", but either of two cases which mimic demand creation. We could, for example, be observing points P_1 and P_2, where a higher quantity demanded is associated with a lower user price and a higher supply. [More generally, we could be observing a change in some characteristic of the good (e.g., quality) other than price.] Or we could be observing points P_1 and P_3, but with demand shifted out by some other influence on demand, such as differences in illness levels, tastes for medical care, or other unobserved influences. If supply responds along the supply function S_1, the actual quantity supplied (and the inputs used to produce that quantity) will be greater when demand is greater.

Even if the possibility of demand creation be granted, and even if one can explain why and how physicians (or others) might create demand for hospital services as well as for their own services, the "insatiable demand" hypothesis still does not follow. The possibility of some demand creation does not imply the certainty of unlimited demand creation, nor does it imply that providers will create demand just because they can do so. In particular, an increase in resources may be granted to shift demand outward, but there is no theoretical or empirical evidence for the proposition that the shift would be proportional to the increase in resources. And yet proportionality would seem to be required by the insatiable demand hypothesis.

To summarize: there is a basic conceptual problem with DRAM's assumption of insatiable demand. It is only legitimate in cases in which the analyst is sure that he is observing a situation of excess demand. My empirical conjecture here is that such situations exist for hospital care in only some countries and for ambulatory visits in almost no countries. A more general model (of which DRAM would be a special case) is one that would include both demand and supply. Such a model would permit (but not require) the possibility of excess demand, and would also permit (but not require) an availability effect in addition to the effect of other demand variables on use. The critical point is that a complete specification of demand needs to be added.

2.3. Provider Behavior

A problem with the specific implementation of DRAM is its failure to model the behavior of the agents in the system. This is especially important with regard to the treatment of costs. There are two things to note here. First, the assumption of cost minimization (or output maximization, even given its division among types of patients) is suspect. There is fairly strong empirical evidence to suggest that, for a number of reasons, decentralized systems may not choose the input combinations which minimize costs (Reinhardt, 1972; Pauly, 1980a, 1980b). The absence of cost minimization shows up as higher costs of lower productivity, but these deviations are not random. Instead, they appear to be related to various incentives to providers, such as those associated with insurance coverage, method and level of physician compensation, etc. It is surely the case that these incentives vary a great deal across systems, and I would conjecture that they probably vary over time even within systems (and even when there are no observable changes in "official" policy). The basic point then is that, by omitting incentives, DRAM both permits the violation of its own assumption of cost minimization and (what is probably more important) may lead to erroneous forecasts.

The thrust of this critique is that DRAM is incomplete as a generalizable positive model of the health care system. That it is incomplete does not, however, mean that it cannot be useful. What is needed is to <u>add</u> some further parts to DRAM, or possibly, for some systems, to either substitute for DRAM or change the way in which DRAM is used. In what follows, techniques for doing so will be discussed in detail.

3. PROBLEMS OF ECONOMETRIC ESTIMATION: GENERAL DISCUSSION

3.1. Introduction

In this section I want to use some simple diagrams to show the kinds of problems that can arise if demand is ignored in a forecasting model. The "problems" are of two sorts: First, and most importantly, ignoring demand can lead to biased estimates of the parameters of DRAM or any other forecasting model. Since the validity of a model depends on the validity of its parameter estimates, such biased estimates could lead to very inaccurate forecasts. Second, in any practical application, it will not be possible to hold all other things equal when predicting the changes over time from changing a particular variable (e.g., resources). In order to control for such changes, some of which may be changes in demand variables, it is necessary to know what such variables are and how they affect the use of care.

3.2. The Nature of the Problem

This discussion will indicate why demand may be an important consideration in predicting the impact of changes in resource availability on the use of medical care -- the primary purpose of DRAM. In most markets, quantity demanded and used is thought to be a function of money-user price. While this relationship appears to hold for medical care as well, money-user price is in fact zero in many health care systems. Even in the United States, third-party insurance payments (public and private) cover about 95 percent of hospital costs, though the extent to which the remaining 5 percent user charge constrains use and cost is not known.

That money-user price is zero, or virtually so, does not, however, mean that the quantity demanded is infinite or equal to the total population. Many, perhaps most, people would not want to live in a hospital even if that were possible, at least given the current form and characteristics of hospital care.

What does seem to be true is that the quantity of hospital care demanded (measured, say,by the desired number of admissions) appears to be affected by what one might call the "quality" or "characteristics" of care. This quality coula mean clinical quality, expected travel, or queuing time. Quality could also be measured, in the case of hospital care by real inputs per admission, as Feldstein (1971) has suggested.

Quantity of care (e.g., hospital admissions) demanded by patients is probably not unlimited for any quality. (It is possible, however, that total expenses (quality times quantity) and total resource consumption are virtually unlimited. A plausible relationship between quality per unit and quantity of care demanded is shown in Figure 3 by the line DD'. The levels of quantity and quality that can actually be received are, however, limited by the resources made available to the system. In the diagram, the combinations of quantity and quality that can be produced from a given amount of resources made available by the government would be shown by such resource lines as R_0.(If quality is available at a constant marginal cost, these lines will have the equation $P \cdot Q = R$, where P is the number of units of quality, measured in dollars. In general, the shape of the resource lines will depend on the production function for quantity and quality.)

Resource lines,R, represent the combinations of quantity and quality the HCS could produce with a given level of resources. As DRAM describes , we can then imagine the HCS as choosing its most preferred point on a resource line by maximizing a utility function in quantity and quality subject to the resource contraint. If resources were R_0, the providers' most preferred point might then be represented by a point such as P_0, the tangency of an indifference curve with the resource constraint. If the resource constraint is then varied,HCS decisions will trace out a "supply" locus such as SS'.

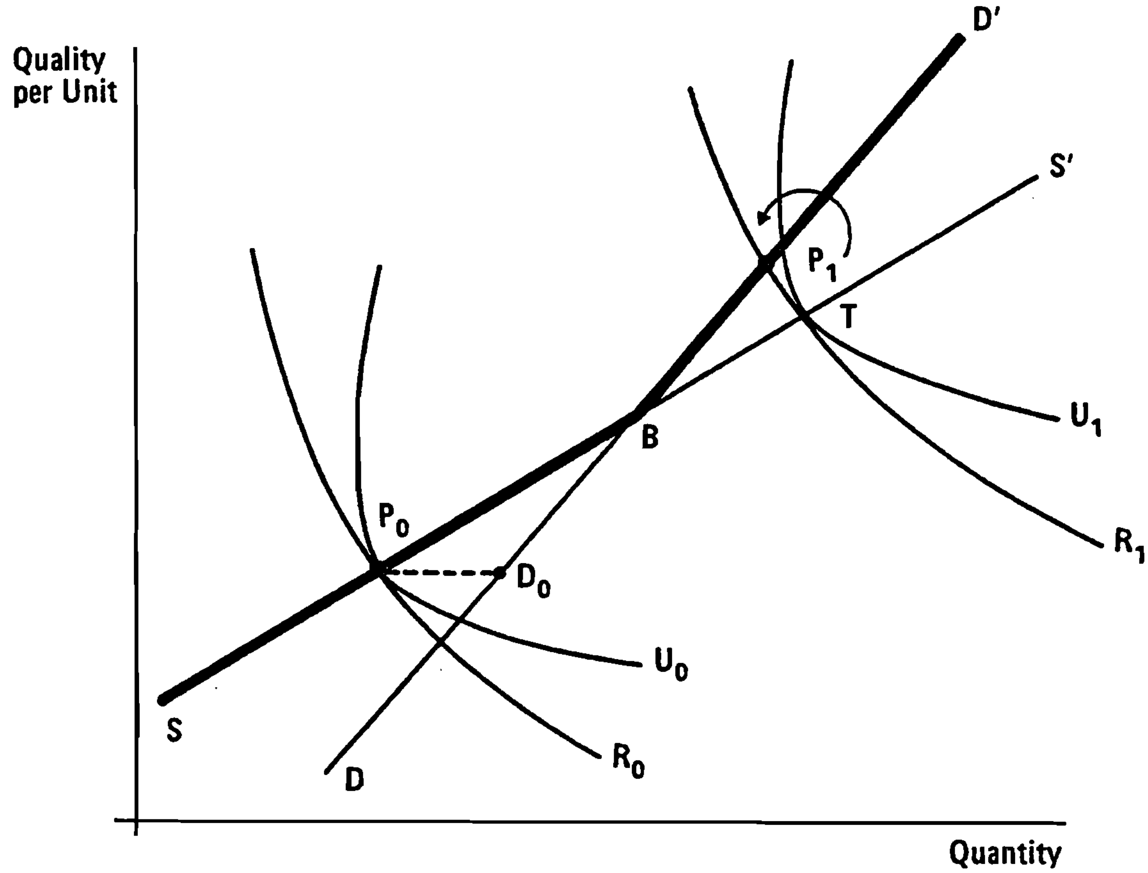

Figure 3. Effect of resources on use with a single demand curve.

As long as consumers are willing to accept what providers prefer, DRAM will explain behavior. But it is possible that consumers are unwilling to demand the output that providers choose. In that case, we suppose that actual quantity (or quality) used is represented by min(Q_D, Q_S), i.e., the minimum quantity at any level of resources that either the HCS is willing to supply or patients are willing to accept.

If then resources are at R_0, actual quantity will be that at P_0, and there will be "excess demand" for quantity of care in the amount P_0D_0. (Conversely, there is an "excess supply" of quality.) This seems to be the sort of situation DRAM is intended to represent. Suppose, however, that resources are at level R_1. Providers would like to be on SS', but now demand (at the provider-preferred level of quality) is insufficient. Instead, the observation will be at point P_1 on line DD'.

If observations are drawn from situations in which resources range from R_0 to R_1, the observed path of equilibria will be represented by the heavy line SD'. This line will reflect neither the preferences of the provider, nor patient demands, but rather some ill-defined hybrid of the two. In the simple linear framework I have used so far, the line SD' will not be well-approximated by any straight line. What is required, obviously, is a modeling framework that permits and identifies the switch between regimes of excess demand and excess supply.

The problem becomes more complicated if demand differs (say, across areas or over time). For example observations might come from the two heavy lines ABC and AB'C' in Figure 4. Then a single line through those points will not fit well, and will not forecast accurately. So it is clear that, at least for high resource levels, accurate forecasting or prediction will require that one take into account possible differences in demand. If resource levels vary widely, and if demand differs, DRAM will not forecast well.

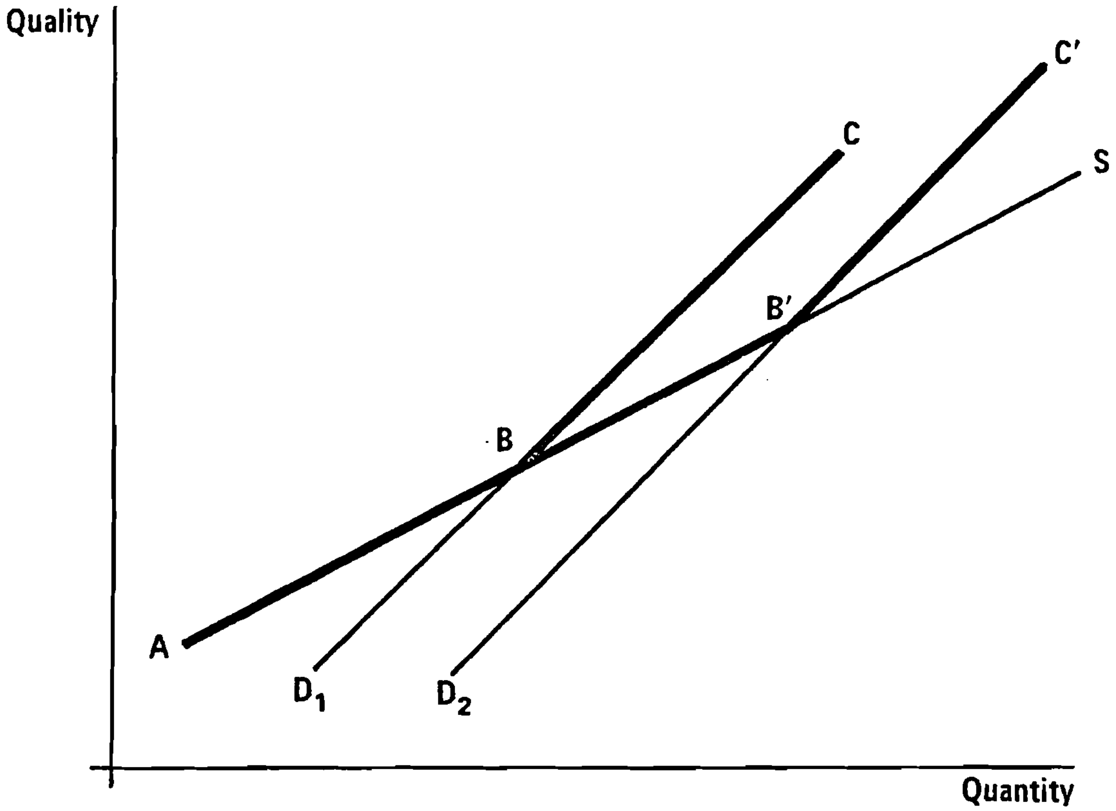

Figure 4. Effect of resources on use with several demand curves.

Of course, if the quantity at which the demand constraint binds is greater than the quantity supplied at available levels of resources, then the DRAM model will still be appropriate. A slightly more realistic version of the provider's objective function removes even this conclusion, however. Suppose that the provider has "standards" for quantity and quality, as represented by the utility function, but suppose the provider also gets disutility from excess demand or the queues its behavior generates. (Alternatively, one may assume that the "standards" change with the level of demand.) The provider is willing to move somewhat away from the standards in order to draw down excess demand. At point P_0 in Figure 5, for example, excess demand would be large. The provider might therefore feel compelled to move to P_0'. If the "response" is proportional to the amount of excess demand, the observed set of points would be represented by the heavy line L_1. This is not a problem if demand is constant, but suppose that demand is sometimes at D_0, sometimes at D_1, and sometimes at levels in between. Then the scatter of observations will be spread between the heavy lines through P_0' and P_0'', and forecasting will be imprecise.

The message conveyed by these examples is that demand must be added to adequately explain or predict use. There is, however, a third possible influence which makes it even more difficult to get an accurate predictive model. This influence arises if the resource budget levels are not random but are selected by a central authority (e.g., a national health authority) based on its perceptions of demand or need, perceptions not perfectly measured by the analyst. Especially in a cross section context, one needs to ask <u>why</u> the central authorities provide the differing levels of resources which generate the data analyzed. Here we are going to assume that all providers have the same behavioral function. (If providers in different areas had different preferences, that would cause even further complications.)

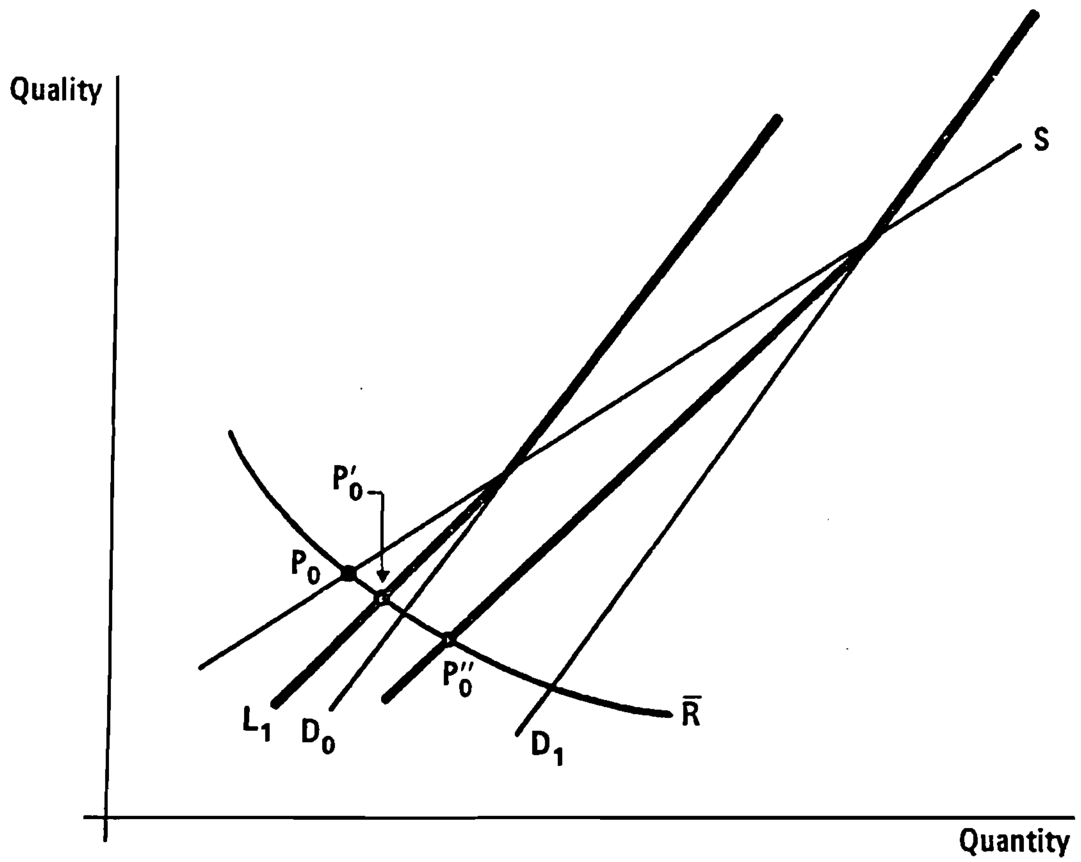

Figure 5. Effect of resources on use with provider response to queues.

A line such as LL' in Figure 6 represents the opportunity locus in one area for the central authority. It must choose among points on that line, balancing quantity, quality, the total budget, and possibly other objectives. As long as the center picks a point on the excess demand segment of the opportunity locus, and as long as all HCSs have the same preferences, its choices will permit a tracing out of the locus implied by those preferences. Problems arise if the center selects points in the demand equilibrium segments in different areas and demand differs across areas. Suppose, for example, that three areas in a country present lines like L', L'', and L''', because of differences in the incidence of illness. The central authorities presumably take that into account (alternatively, they get disutility from sick people going without care), and so choose the resource levels R', R'', and R'''. Observed points would then be P', P'', P'''; it is obvious that a line MM' fitted through these points tells one neither about the utility function of providers nor about demand; and would be worthless in predicting how use would respond if resources were exogenously increased. If illness levels, or other demand parameters, are either not perfectly observed or not inserted into the model, the problem is one of econometric identification. The right-hand variable "resource level" will be correlated with the error term, and so its coefficient will be biased.

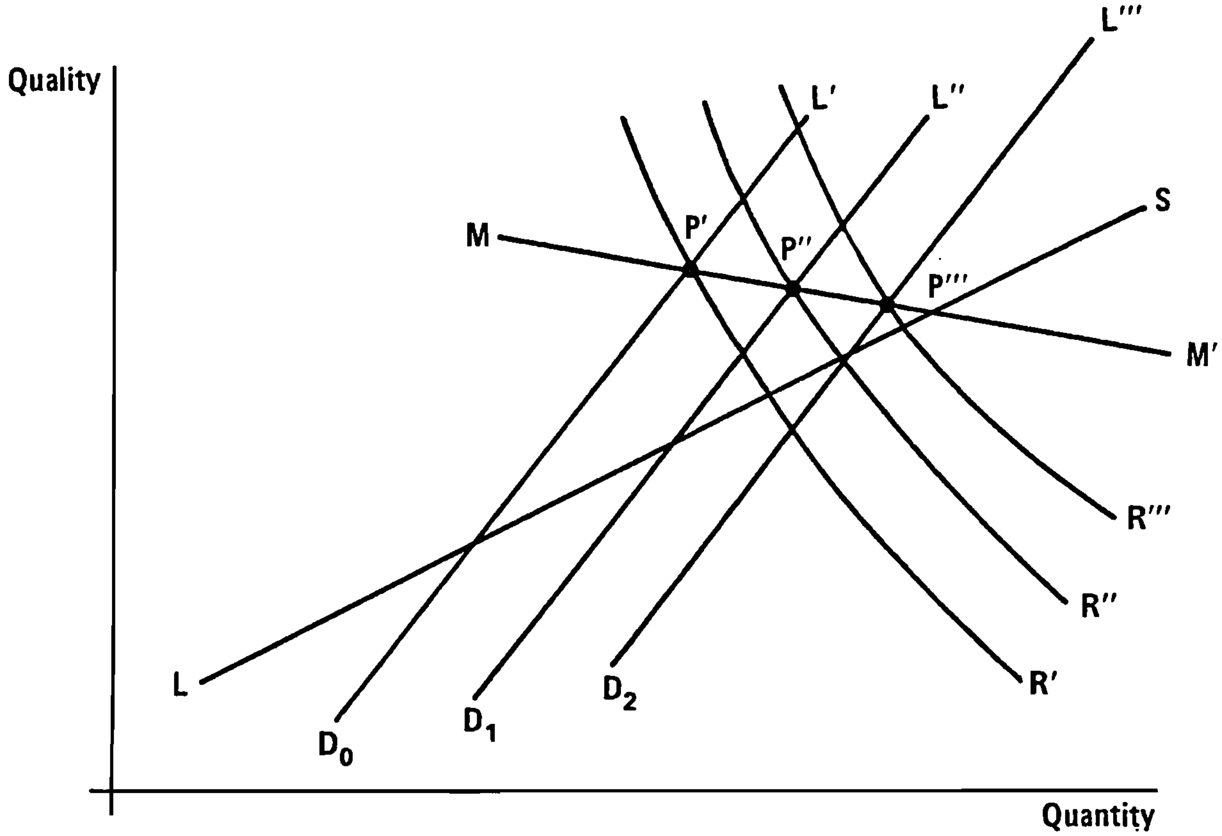

Figure 6. Effect of resources on use with endogenous resource budgets.

4. ECONOMETRIC MODELING OF DEMAND DISEQUILIBRIUM

4.1. Introduction

In order to indicate what these observations imply about modeling the HCS, I first consider a designedly simple model of DIDD, and indicate what a proper econometric specification would be. I then compare that specification with what has actually been used in the literature (including DRAM), and indicate the kinds of biases involved. Subsequent sections will discuss more complicated (and less tractable) versions of the model.

4.2. What is Outside the Model

Before describing the model, I will indicate those parts of the HCS which it omits. DIDD is a model of the "market for health care", to use Rutten's (1979) terminology. This means that it leaves out (or models in only a simple way) the following additional parts of the HCS

- (i) The determination of the level and distribution of morbidity and/or symptoms
- (ii) The market for health manpower
- (iii) The market for non-labor inputs
- (iv) The market for health manpower training

Some of these omissions are more important than others. Topic (i) has already been extensively and adequately treated by models developed at IIASA, and so need not be developed further. Topic (iii) probably can be modeled as simple competitive markets (in market economies), and so is not especially interesting. Topics (ii) and (iv) are related, and have been subject to some analysis at IIASA (Shigan et al., 1979) though not in ways which seem especially applicable to market economies. I will ussually assume for non-physician manpower that training is exogenous, and wage rates are competitively determined. For physicians, there will have to be somewhat more explicit modeling of topic (ii) contained in DIDD, but I will assume again that total supply at any point in time [topic (iv)] is exogenous. Obviously the manpower aspects need further specification than this model will be able to give.

4.3. Econometric Specification: Disequilibrium Modeling

To keep matters simple, let us assume for this section that the demand function is not itself affected by suppliers. The complications raised by the possibility of supplier -- especially physician -- induced demand will be discussed below. Let us also assume that preferences of suppliers are not affected by the magnitude of excess demand.

The general econometric specification for markets in which disequilibrium occurs has received a great deal of attention in recent years (Fair and Jaffee, 1972; Maddala and Nelson, 1974; Fair and Kelejian, 1974). This specification has been applied to the housing market (Fair and Jaffee, 1972), the market for business loans (Laffont and Garcia, 1977), and to centrally planned markets generally (Portes and Winter, 1980). It has not, to my knowledge, been applied to the health care system, even though the possibility of disequilibrium in such markets has been widely noted (Feldstein, 1971; Feldstein, 1977). Here I will only describe the fundamental idea of such methods; more extensive description can be found in the papers cited.

The general econometric specification for situations in which disequilibrium occurs can be written as

$$Q_{Di} = f_1 \quad (X_{Di}) + \varepsilon_i$$

$$Q_{Si} = f_2 \quad (X_{Si}) + \eta_i$$

$$Q_i = \min \ (Q_{Di}, Q_{Si})$$

where Q_{Di} is the quantity demanded in observation area(or time period)i, Q_{Si} is the quantity desired to be supplied in area i, X_{Di} is a vector of values of demand variables for area i, X_{Si} is a vector of values of supply variables for area i, Q_i is the actual quantity in area i, and ε_i and η_i are error terms.

We cannot estimate the demand and supply relationships directly, even assuming that the system is identified, because we observe only Q and not Q_D or Q_S. If no additional information is available, the estimation technique must use a maximum likelihood approach,and then must estimate parameters conditional on the probability of being in one or the other regime.

The unconditional (with regard to regime) density of Q_i is

$$h(Q_i|Q_{Di},Q_{Si}) = g_1(Q_{Di})\int_{Q_{Di}}^{\infty} g_2(Q_{Si})\,dQ_{Si}$$

$$+ g_2(Q_{Si})\int_{Q_{Di}}^{\infty} g_1(Q_{Di})\,dQ_{Di}$$

where h, g_1, and g_2 are the respective frequency distribution, with the latter two reflecting the distribution of ε and η. One then wishes to find parameter estimates to maximize the log-likelihood

$$L = \sum_i \log h \ (Q_i)$$

Maximum likelihood method are available to do this, although it appears that there can be difficulty in finding an algorithm to do so.

This kind of model provides two beneifts. First, it yields consistent estimates of the parameters of both demand and supply. Second, it also permits estimation of the probability that a given observation is in a supply-constrained or a demand-constrained regime.

4.4. Econometric Specification of a Budget-Constrained, Zero User-Price System

I now wish to outline a model applying these general principles to the HCS. It will be assumed that the user price is zero, and that various geographical areas from which observations are to be drawn receive budgets of homogeneous resources. It is also assumed, of course, that each area represents an independent observation; there is no flow of patients across areas.

Suppose then that there is an aggregate structural demand equation for each area of the form

$$Q_D = Q(K, X) \qquad (1)$$

where Q_D is the quantity of care (e.g., hospital admissions) demanded by persons in each area, K is quality, and X is a vector of other demand variables.

There is also a production function of the form

$$h(Q, K, R, Z) = 0 \qquad (2)$$

or

$$Q = \hat{h}(K, R, Z) \qquad (2A)$$

where R is resources and Z is a productivity-shift parameter.

If R is fixed at $\bar{R}$, one possible observation is the solution to

$$Q_D(K, X) = \hat{h}(K, \bar{R}, Z) \qquad (3)$$

That is, given $\bar{R}$, K is adjusted until $Q_D = Q$. So a reduced form equation for Q_D can be written as

$$Q_D = Q_D(X, R, Z) \tag{4}$$

The other possible observation for any area is obtained by following the general procedure that DRAM describes. The utility function of the HCS is

$$V = V(Q, K, T) \tag{5}$$

where T is "tastes", a shift parameter. V is maximized subject to the production function and $R = \bar{R}$. Solution to this problem is a reduced form equation for Q_S, the quantity desired to be supplied, of the form

$$Q_S = Q_S(T, R, Z) \tag{6}$$

(Note that, because the HCS is a monopolist, there is no structural supply equation.) The actual observation Q is then given by

$$Q = \min(Q_D, Q_S) \tag{7}$$

It would appear to be feasible to estimate this model by the methods described in the preceding section. The demand variables and the form of the demand equation could be the generalized linear approximation much used in the literature. The determination of Q_S could (especially in this one-input, two-output case) also be approximated by a linear relationship, and the reduced form itself could be used for forecasting. Alternatively, one could estimate the key parameters for DRAM itself by the ML technique, and then use DRAM to produce forecasts of Q_S.

Once forecasts of Q_D and Q_S are obtained, the actual forecasted Q will be whichever value is smaller. The forecasted K will then be the value that is consistent with this Q, the forecasted values of R, and the other exogenous variables.

One minor qualification. It is possible for the demand equilibrium equation (3) to have multiple solutions. In such a case, the relevant Q_D is the one which is, in some sense, closest to Q_S. How to define "closeness" operationally may sometimes be difficult.

The econometric problem then is to estimate the functions (4) and (6) when it is not known to which function an observation corresponds. An Alternative superior to the ML method is available if some indicator of the probability of being in either regime can be found. For instance, if data on waiting lists is available, it seems plausible that Q is more likely to equal Q_S when the waiting list is long than when it is short. To incorporate this into the model, we can assume

$$pr(Q = Q_S) = g(L) \tag{8}$$

where L is the length of the waiting list.

We must, however, consider one problem with this specification. Suppose that R is not exogenous but is instead chosen by some central authority according to

$$R = R[(Q_D - Q_S), P] \tag{9}$$

It is clear that, without information on P (preferences), the system is not identified, and one cannot get consistent estimates of the impact of R in equations (4) and (6). This is a serious problem but one with no obvious empirical solution short of getting a set of observations where P differs in a measurable way. That is, it may be possible to find some exogenous variables unrelated to demand which affect the authority's preferred level of R. Perhaps some political or historical variables can serve here, or perhaps R is adjusted to serve objectives of interregional equity. Then it would be possible to make R endogenous, in a two-stage procedure, and to proceed with the estimation.

A further extension, but one that seems much more difficult theoretically, is to disaggregate care into various types (e.g., for different diseases). This is, of course, the problem for which DRAM was designed, but there appear to be serious difficulties in developing methods for situations in which there are several closely related outputs (either as substitutes or complements), and for which some markets may clear while others may not. Suppose, for example, the DRAM solution for a given R involves providing quantities which leave some markets with excess demand (for quantity) and others with excess supply. Then DRAM would have to be re-solved with demand constraints inserted. Moreover, since demand may depend on the price or quantity of close substitutes or complements, it will be necessary to specify interactions in the demand system as well as in the HCS preference function. I will discuss a very simple model of this type at the end of the paper, but for the present it should be noted that disaggregation into output types appears to be difficult.

4.5. Excess Demand and Demand Creation in the Literature

To my knowledge, the disequilibrium specification has not been used in the literature on the use of medical care. Instead, the procedure has been to estimate either a function labelled "demand" or "supply-preference", and then often to add, in various ad hoc ways, the influence of the other function.

Since estimation of "demand" functions is more common, we begin with them. The ordinary procedure, used with U.S. (e.g., Feldstein, 1971), Canadian (e.g., Evans, 1974), and Dutch (e.g., Rutten 1979) data is to estimate a function which relates use to a set of demand variables, such as user prices, time price proxies, income, indicators of health status, and other socio-demographic variables. To this set of variables are then added measures of "resource availability", usually hospital beds or physicians per capita in the presumed market area. While sometimes actual values are used in regressions with aggregated use data, the possible endogenity of resources (to area demand) is handled either by two-stage least squares (Fuchs and Kramer 1973; Fuchs 1978) or by using individual rather than aggregated data (Newhouse and Phelps, 1976; Pauly, 1980b).

The explanations for what a significant coefficient on such availability measures is supposed to represent are varied. Sometimes a permanent excess demand story is told (Feldstein, 1971). Sometimes it is argued that, because of the stochastic nature of demand, markets are sometimes in excess demand, so that aggregated by time (e.g., hospital days or admissions per year) will reflect some excess demand (Newhouse, 1974; Newhouse and Phelps, 1976; Pauly, 1980b, Chapter 6). Most often, the argument is that resources somehow create the demand for themselves (Fuchs, 1978; Evans, 1974), though this argument is much stronger for physician stock, physician services, and complementary hospital admissions than for hospital resources and hospital services (Pauly, 1980b).

In virtually every case, resource availability measures are simply entered into a catch-all regression. The only attempt to say anything rigorous about the relationship between theory and specification is by Newhouse (1974) (who argues in the excess demand case that one ought to take the square root of resource measures) and by Pauly (1980b) who attempts, on wholly a priori grounds, to divide his sample into areas in which excess demand is more or less likely to prevail. This procedure is useful for hypothesis testing but not for prediction.

There is not a clear dividing line between demand function estimates and supply-production function estimates. For example, the Evans and Rutten papers actually focus much more on physician preferences and effects on demand generation than on demand variables as such; they might better be called "patient flow" functions than demand functions. But even at the end of the spectrum as represented by production functions, Reinhardt (1972) and Pauly (1980b) at least experimented with inserting per physician demand proxies, again in an ad hoc way.

There are two potential problems with these ad hoc methods:

(1) Parameter estimates will be biased
(2) Standard errors will be inflated

Coefficients on variables that appear only in the demand function (e.g., income or time price) or only in the HCS preference function (e.g., indicators of tastes) will be biased toward zero. The estimated coefficient on resources itself will, in the case shown in the diagrams, tend to be in between the coefficient in the demand function and the coefficient in the provider preference function. Standard errors will be inflated because one is trying to estimate two functions with only one relationship.

In the DRAM model, demand is not explicitly included. It is included implicitly in the sense that the targets might reasonably be affected by at least some of the determinants of demand, such as the prevalence of illness. Other influences -- travel and waiting time, income, sociodemographic variables, and quality -- all of which have been found to influence demand, are ignored, presumably on the ground that such non-medical variables will never be at a level that demand will fall short of supply.

4.6. Econometric Specification in Market Economies

While the preceding analysis was developed for the case of a budget-constrained, zero user-price system, similar observations apply when user price is not zero and the budget is constrained by demand at that nonzero price, as well as by a requirement that the HCS break even. Here two regimes are possible. In one, only demand operates as a constraint, of the form

$$Q_D = Q(K, X, \hat{P}) \qquad (10)$$

where $\hat{P}$ = NINS. P is the user price of care, NINS is the fraction not covered by insurance, and P is the gross price. "Break even" for the HCS implies that

$$PQ = C(Q, K, Z) \qquad (11)$$

where C(.) is a total cost function. That is, the HCS must receive enough revenue to cover its costs. (A subsidy could obviously be added.) If there are constant returns to scale (11) implies

$$P = \overline{AC}(K, Z) \quad \text{or price equals average cost} \qquad (12)$$

Maximizing V(Q, K, T) given the constraints (10) and (12) then yield the equilibrium values of Q and K for the HCS. This model is the same as that suggested by Newhouse (1970) and Feldstein (1971). Here there is no possibility of excess demand, since quality is always adjusted so that $Q_D = Q_S$. Excess demand would mean that the HCS was not obtaining as high a quality or quantity level as it could. (This is not to deny that some outside decisionmaker may want it to have lower quality.) The actual content and form of the utility function, and whether it can depend on some more specific and observable arguments (such as physician income) has been the major unsettled question in the literature on U.S. hospitals; one I shall discuss in another paper. For the present, it is sufficient to work with this general form. One can, however, use this form to ask how use will change as some of the parameters (e.g., NINS) change; the implication is that the change depends primarily on the demand function, though provider preferences also matter.

A picture of such an equilibrium is shown in Figure 7. Here the line DD' shows the combinations of quantity and quality which satisfy the demand constraint (10) and the break-even straint (12). In contrast to the zero user-price case, when the DD' curve was likely to have a positive or vertical slope, here the fact that a user price must be paid causes the curve eventually to have a negative slope. The otherwise unconstrained utility-maximizing hospital then selects a point (such as P*) of tangency of its indifference curve to this opportunity locus. This point must, of course, be in the negatively sloped portion of DD'.

An interesting and relevant case is one in which, in addition to demand and break-even constraints, the provider or system is also subject to a maximum revenue constraint. Such a constraint was proposed by President Carter (though defeated by Congress) in order to contain hospital costs; it also forms part of Sen. Kennedy's National Health Insurance plan. In such a situation, the demand and break-even constraints would still hold, but, in addition, there would be a maximum revenue constraint

$$PQ = R \leq \bar{R} \tag{13}$$

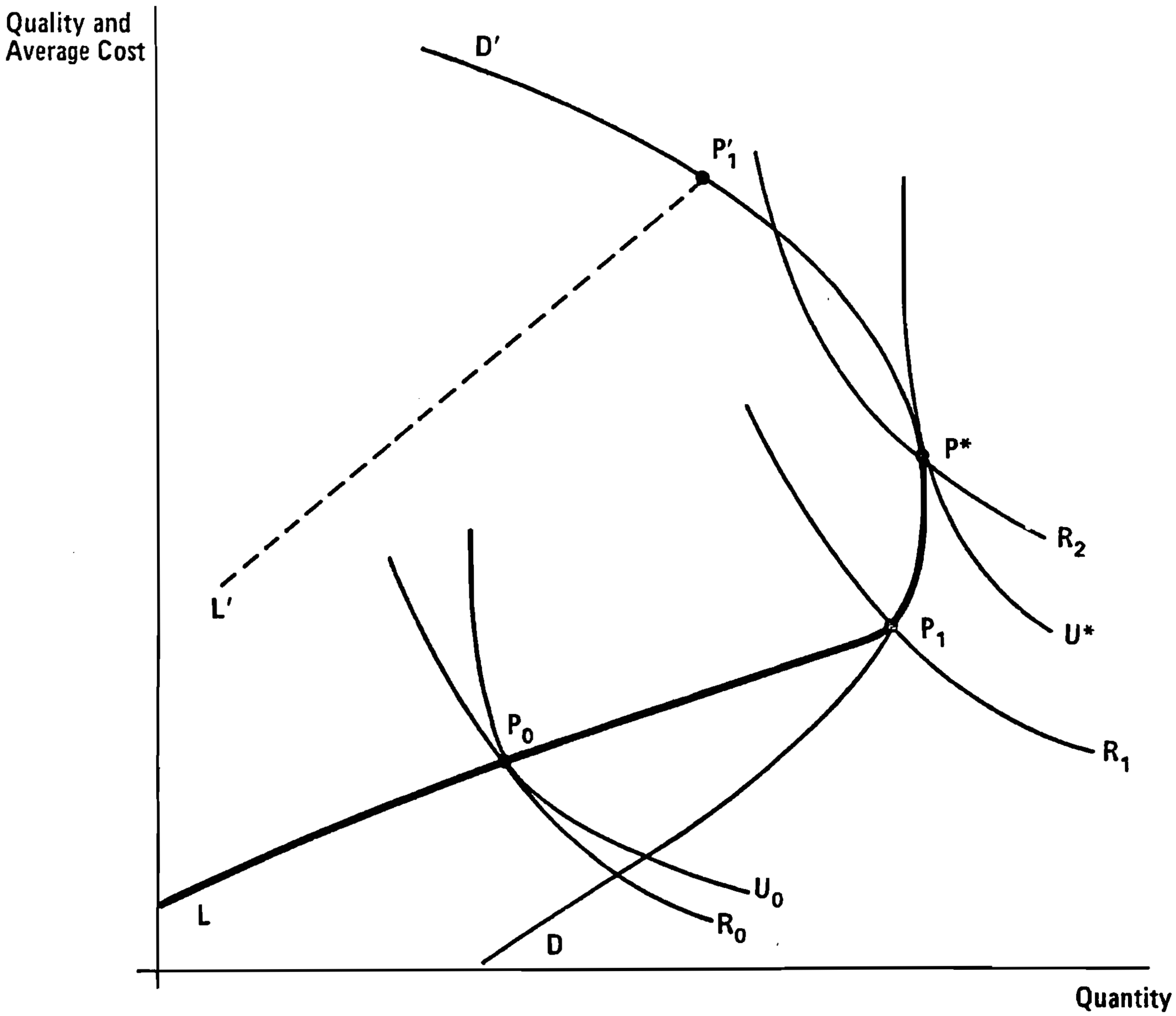

Figure 7. Effect of resources on use with user prices.

Possible outcomes for two alternative configurations of HCS preferences (or relative costs of quantity and quality) are shown in Figure 7. For example, if the resource limit is R_0 and HCS preferences are represented by the indifference curve I_0, then point P_0 would be the observed quantity-quality combination. Up to point P_1 (and resource limit R_1), outcomes would be determined by HCS preferences, as DRAM describes. Beyond R_1, however, outcomes are constrained by the demand-break-even locus DD'. Finally, if the resource limit is increased beyond R_2 (the line through P*), there is no effect on outcome, because at that point the resource limit ceases to be binding. The locus of observed points would then be the heavy line LP_1P^*. Alternatively, for a different configuration of HCS preferences, the path could look like $L'P_1'P^*$.

One way to think of the solution is to think of the HCS as now choosing among

(a) Q_S, the quantity that maximizes V subject only to constraint (13)

(b) Q_D, the quantity that maximizes V given $h(Q_D,K_D,\bar{R},\hat{P})=0$

(c) Q*, the quantity that maximizes V given constraints (10) and (12)

Practically, outcome (c) seems unlikely, since it implies that the revenue constraint is not binding. But the problem of choosing (a) or (b) is then the same as in the zero user cost case, except that

(a) The user price is added to the demand function

(b) There may not be a monotonic relationship between quantity (or quality) and the resource limit

(c) There is some value of the limit beyond which an increase in the limit leaves behavior unaffected

If one thinks then of estimating the effect of changes in the revenue limit on admissions (or on quality), there are in a sense three different regimes:

(1) A regime in which HCS preferences (e.g., as modeled by DRAM) determine outcomes

(2) A regime in which observations trace out the hospitals' break-even constraint. Such a relationship can be nonlinear and nonmonotonic

(3) A regime in which changes in R do not affect hospital behavior

If only regimes (1) and (2) usually hold, a modification of the approach from the previous section could be used. Perhaps there is enough information in the problem to also permit an estimate of the likelihood of being in regime (3). Some econometric problems obviously remain.

4.7. Conclusion

Recent research in the demand for medical care seems to be moving in the direction of including more quality measures or adjusting for quality in some way (Feldstein, 1977). If this trend continues, the proper modeling of the demand constraint would be all the more useful. And as even market-oriented systems move toward budgetary limitations and supply constraints, it will be more essential to distinguish equilibrium from disequilibrium and to characterize disequilibrium. Econometric techniques are available to permit this to be done.

The technique suggested above does more than just permit more accurate parameter estimates. It also permits estimation of the probability that a given observation is in supply-constrained disequilibrium or not. So it might provide a useful tool for estimating the actual impact of regulatory or cost containment devices, and for describing how that impact varies across providers or areas.

5. INCENTIVES AND PRODUCTION EFFICIENCY IN HEALTH PLANNING MODELS

5.1. Introduction

The supply or provider preference side of the model just described is intended to represent how the agents in the HCS convert resources into outputs. DRAM is one way of representing this process, and could in fact be used to forecast the Q_S in equation (6) (once unbiased estimates of the parameters had been obtained).*

But these methods wholly ignore the question of whether it always takes the same amount of resources to produce a given batch of outputs. That is, they assume that production efficiency [measured by the variable Z in equation (2)] was everywhere and always the same. There is fairly strong empirical evidence, however, that

(a) A given batch of measured medical output can be produced with alternative combinations of inputs

(b) A given batch of outputs is sometimes produced with more or less all of the inputs (technical inefficiency)

(c) Inputs are not always chosen to minimize costs (to maximize output for a given money budget)

---

*There are other, possibly simpler, ways of representing this process. For example, one could estimate a set of "HCS demand for output" regressions of the general form

$$Q_i = CONST_i + \alpha_i BEDS + \beta_i DOCS + \gamma_i BEDS \times DOCS$$

where Q_i is the amount of output i (e.g., admissions for upper respiratory infections) and BEDS and DOCS are hospital and physician resources in physical terms. If resources were provided in the form of money budgets, the explanatory variables would be the amount of the budget and the prices of inputs or of outputs.

It would seem that the specification of the relationship between Z and observable institutional structures should be part of a general HCS model.

The empirical importance of such measured inefficiency is not known, nor is the extent to which it may represent a change in some unmeasured quality of output. It is known that costliness and productivity varies in predictible ways with incentives to providers.

For example (Table 1), taken from Reinhardt, Pauly, Held (1979), indicates the relationship between inputs and output in the production function for physician office visits for a sample of U.S. medical groups. Output is measured by the logarithm of the number of patient visits. The critical variable there for purposes of this discussion are two measures of the closeness of the link between physician income and productivity, INCOPROD and PRODDIST. Either measure is statistically significant, with positive coefficients. The coefficients imply that changing the relationship from the bottom (no relationship between productivity and income) to the top (strong relationship between productivity and income) will raise weekly office visits by 20 percent.

A planning model ought to include consideration of such aspects of productive efficiency for two reasons:

(1) Variations in efficiency can lead to errors in forecasting the relationship between resources and output.

(2) Knowing how to increase output by reducing identifiable sources of inefficiency is often highly desirable information for planners, since it permits them to do more with less.

Table 1. Group practice study: two-stage-least-squares estimates of office visit production function.

Description of Variable	Acronym	Estimated Regression Coefficient[a] Equation 1	Equation 2
Practice Inputs			
Number of hours Dr.X spent at the office seeing patients, last 7 calendar days (logarithm)	LBOHRSF	0.5064***[†]	0.5067***
Number of examination rooms in the group per FTE MD (logarithm)	LNEXRPMD	0.1509***	0.1542***
Weekly hours of non-physician medical personnel per FTE ME in the group (includes graduate physician assistants, registered nurses, licensed practical nurses, and technicians)	HRSMED	0.0026**	0.0026**
Weekly hours of administrative personnel, per FTE MD in the group (includes business administrator, secretarial personnel and others)	HRSADM	0.0037***	0.0037***
Total weekly hours of support staff, squared	HRSTOTSQ	$- 0.611 \times 10^{-5}$***	$- 0.618 \times 10^{-5}$***
Physician Characteristics			
Number of years since graduation from medical school	YRSGRAD	0.0203***	0.0204***
Number of years since graduation from medical school, squared	YRGRDSQ	- 0.0004***	- 0.0004**
Medical specialty of physician:			
General practitioner (Yes = 1)	GP	0.4742***	0.4639***
Pediatrician (Yes = 1)	PD	0.4725***	0.4674***
OB/GYN specialist (Yes = 1)	OB	0.3647***	0.3464***
Physician's own assessment of his/her responsiveness to monetary incentives (converted to a dummy variable set to 1 if the physician declares him/herself responsive to monetary reward, 0 otherwise)	OWNRESP	0.0518	0.0560
Characteristics of the Physician's Group Practice			
Type of practice (dummy variable set to 1 if a multispecialty group, 0 otherwise)	MULTSPEC	0.0534	0.0559
Size of group, measured by number of FTE MDs practicing in the group (logarithm)	LNGRPSIZ*	- 0.0687*	- 0.0856**
Is any graduate physician assistant employed by the group (Yes = 1, No = 0)	GPA	- 0.0579	- 0.0581
Percentage of group's income distributed to numbers on the basis of productivity	PRODDIST*	---	0.00277***[c]
Scale from 1 to 10 indicating the closeness of the link between the individual owner-member's income and his/her productivity (1 = not related, 10 = completely related)	INCOPROD*	0.0283**[b]	---
Binary variable set to 1 if 50% or more of the group's revenue came from prepayment, 0 otherwise	PCHMO*	0.0296[d]	- 0.0132[d]
Constant Term (Intercept)			
Represents internists in single specialty groups with PCHMO = 0 and GPA = 0		1.8939***	1.9703***
Statistics			
R^2 (adjusted for degrees of freedom)		0.37	0.38
Regression F-statistic (16,786)		35.57	36.42
Number of observations		803	803

[†]A variable whose acronym is asterisked is treated as an endogenous regressor.

[a]Statistical significance level of coefficients: ***less than 1%; **less than 5%; *less than 10%. Coefficients without an asterisk are not statistically significantly different from zero at a level of 10% or less.

[b]Actual significance level is 1.85%.

[c]Actual significance level is 2.39%.

[d]Significance level exceeds 90%.

Source: Reinhardt, Pauly, and Held (1979).

5.2. Production Efficiency and Incomplete Budgeting

In some countries, medical resources are not all subject to public budgeting or control. For example, a country may control the supply of physicians, but not the level of resources the physician chooses to hire. If this level varies, and if output is affected by the use of ancillary personnel, output may be difficult to forecast by knowing only the publicly determined inputs. If the provider may be assumed to be a cost minimizer-profit-maximizer, it is possible to specify a supply function for his services as depending on output prices and input prices. An aggregate supply function then depends on the number of providers, as well as the prices of inputs and outputs. If the provider deviates from cost minimization in ways which are related to identifiable variables, these variables can also be included in the supply functions. The critical point, however, is that forecasting either total services or their division over types of outputs requires forecasting the vector of output prices, the vector of input prices, and the values of efficiency shift variables. Where output prices are set exogenously, one also must be concerned with the level of demand as well as the level of supply in order to forecast use.

6. PHYSICIAN EFFECTS ON DEMAND IN A DISEQUILIBRIUM SPECIFICATION

6.1. Introduction

The preceding sections have treated the demand for medical care as if it were like the demand for any other commodity; the analysis could equally well have been applied to publicly provided housing, publicly provided transportation, or any other good with qualitative dimensions. The demand for medical care, in the previous discussion, could either be satisfied or not, but its use was independent of the level of excess demand or excess supply; its use depended only on the constrained utility-maximizing behavior of consumers.

One unusual characteristic of the demand for medical care, much discussed in the literature, is that it may be subject to permanent manipulation by providers in an important way. It is usually supposed that the physician, in the rule of advice-giver or agent for the patient, can engage in the activity of "demand creation". It is further supposed that he is likely to do so to an extent which depends upon the excess supply of his own services or of other services.

While the literature on this subject is extensive (and growing), it is not particularly clear or conclusive. One reason for this, I believe, has been the tendency to model this relationship as one of <u>agency</u>. While this approach was perhaps sparked by Feldstein's (1974) suggestion, and while it is useful in some circumstances (cf. Pauly, 1979), my view is that it leads to an empirically inappropriate attempt to divide medical services into those that are patient determined (such as the initial decision to contact a physician) and those that are wholly physician determined (such as decisions on consultations, revisits, hospitalization, or surgery). The second set of decisions involve a much larger share of health resources. It may not be appropriate to view these decisions as entirely physician determined. A better approach is to view the physician as providing advice and certification, but with the ultimate decision being made by the patient (even if he decides to rely wholly on the physician's advice). It is possible, for example, to model the process in a Bayesian framework, and then permit a whole range of patient prior preferences and reactions to physician advice (Pauly, 1980b).

In an informal sense, all decisions after the initial contact reflect some mix of patient and physician preferences, and one should incorporate both the mix and the preferences into the explanation of use. There is sufficient evidence of patient noncompliance with physician "orders", and of patient rationality in the initial selection of the physician who will give advice (Pauly and Satterthwaite, 1980) to reject the model of wholly physician-determined behavior. Of course, there are some consumption decisions, such as receipt of a prescription drug or admission to hospital, which the physician must order, but the point is that physician preferences for a particular type of use are necessary but not sufficient for that use to occur.

Demand creation implies that physicians prefer more use than patients would prefer if they knew what the physician knew. The easiest way to model things is to assume that the physician provides advice to which the patient can react. What the physician can do, however, is alter the content or accuracy of that advice depending on the reward (financial or otherwise) to him from one pattern of patient care use as compared to another, and depending on how patients respond to various levels of accuracy.

A simple way of seeing how this works is presented in Figure 8. Here we assume that the physician who provides advice on recommended services also performs those services himself, and that services and physicians are homogenous. (Circumstances in which the physician prefers to or is required to refer to other physicians will be ignored.) The critical ideal is that the demand for a physician's services may depend on the accuracy of his advice.

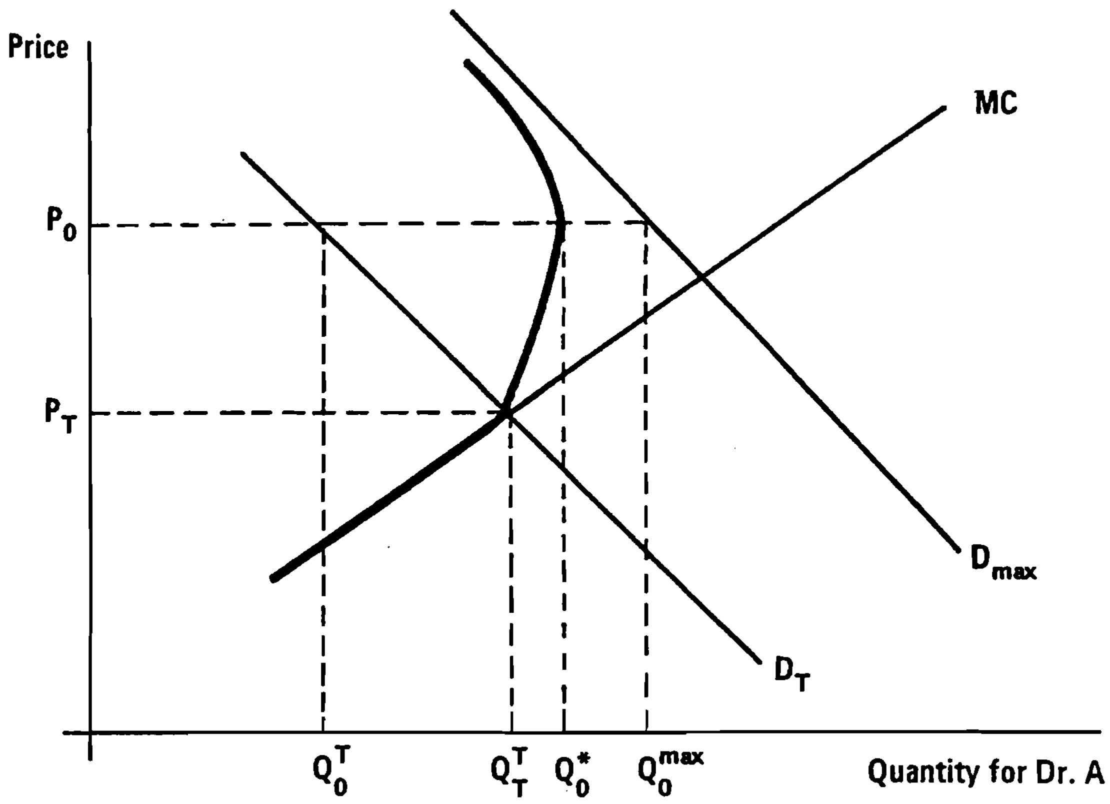

Figure 8. Demand creation in medical markets.

If the physician's advice was as accurate or truthful (given his current knowledge)as it could be, he would experience some demand curve such as D_T. But over some range, the physician can cause the amount demanded from him at various levels of user price (including zero) to change by changing his accuracy.

There is, however, some maximum quantity that people are willing to take from a given physician, no matter what he says; this is represented by D_{max}. The position of D_{max} is determined both by consumer's prior distribution and likelihood functions, and, in some long-run sense, by the individual physicians' reputation with regard to accuracy. That is, D_{max} can be close to D_T because

(1) Consumers are virtually certain about the preferred course of treatment, and so ignore inaccurate advice

(2) Consumers probability distributions are little affected by what physicians say (even if they are quite uncertain)

(3) Inaccurate information will be discovered, exposed, and lead consumers to switch physicians, reducing the long-run pay off to the physician of inaccurate advice

Between D_T and D_{max}, the physician can determine the demand curve he will face. We assume that the marginal cost of his output, including both the explicit costs of inputs he hires (office staff), and a money measure of the opportunity cost of his time, is represented by MC. Assume for simplicity that the physician takes the gross fee he receives as fixed (either by government or by professional consensus). Then we can describe <u>three</u> possible regimes characterizing the observed relationship between gross price and use.

If gross price is below P_T, the situation is unequivocally one of excess demand. The physician might as well give accurate advice (unless he wants to "destroy" demand to avoid the imprecations of unsatisfied demanders), but he is not willing or able to permit that demand to be satisfied.

The difference between this model and one of simple excess demand or supply becomes apparent if the fee exceeds P_T. Suppose, for example, that price is at P_0. Quantity could be Q_0^T, or Q_0^{max}, or anything in between depending on the value of accuracy the physician selects. (If he were only interested in net income, he would move the demand curve out to the D_{max} curve here and for all prices above P_{max}.) In order to explain what level of accuracy the physician will choose, a number of models (Evans, 1974; Sloan and Feldman, 1977; Pauly 1980b) assume that the physician gets disutility from providing inaccurate advice. Then he chooses the level of accuracy by balancing off the increase in his net income from shifting the demand curve against the disutility from providing less and less accurate advice. If the marginal disutility from inaccuracy increases and (eventually) the shift in demand for a marginal reduction in accuracy decreases, there is likely to be an interior maximum between Q_0^T and Q_0^{max}, such as at Q_0^*. The location of this point will depend on the physician's utility function, his income, the reward for creating demand (P-MC), and the impact of accuracy on demand.

The set of observed points will therefore follow the heavy line in Figure 8. At fee levels below P_T, the supply or MC curve is traced out. Above P_T, the curve represents neither a demand curve nor a supply curve, and it is therefore difficult to specify its shape precisely. Because it diverges from D_T but cannot cross D_{max}, the curve above P_T could be thought of as having two segments: a segment with a slope less than either demand curve (possibly positive, as shown in the diagram from P_T to P_R), and eventually a segment roughly paralleling or asymptotic to one or the other demand curve (shown in the diagram above P_R). In this sense, then, there are roughly three regimes:

(1) A regime of excess demand, when the fee is below P_T

(2) A regime when use is mainly determined by provider preferences and utilities, although demand has some influence

(3) A regime in which use is primarily determined by demand

The distinction between (2) and (3) is not as sharp as that between (1) and (2), since both demand and provider preferences will determine use for all prices in excess of P_T, but it does nevertheless seem useful (at least if one is using linear approximations) to try to distinguish the two segments.

This model has been developed for a system in which patients pay some user charges. But even if care is free of money-user charges, the story is much the same, except that, given accuracy, the demand curves now become perfectly inelastic with respect to the price paid to the physician. These can, however, be a second round effect if patient demand for visits is a function of physician service quality. The physician can then obtain patients either by changing the accuracy of his advice or the quality of services he provides. In terms of the figures in section 3, we need only label the "D" curves D_T, redrawn in Figure 9. Up to R_T, the physician will tell the truth and the mix of services will be wholly determined by his preferences. Over the range $R_T R_{max}$, as drawn, demand will be created and use will depend on a combination of provider and patient preferences, along with variables which reflect the desire and the ability of the physician to create demand. Above R_{max}, the scatter of observations traces out the demand curve D_{max}. (The curve could asymptotically approach D_{max}.) "Higher resources" here have a different result depending on whether they are used to pay for more physicians' services (at a given unit price), pay a higher unit price, or provide non-physician resources.

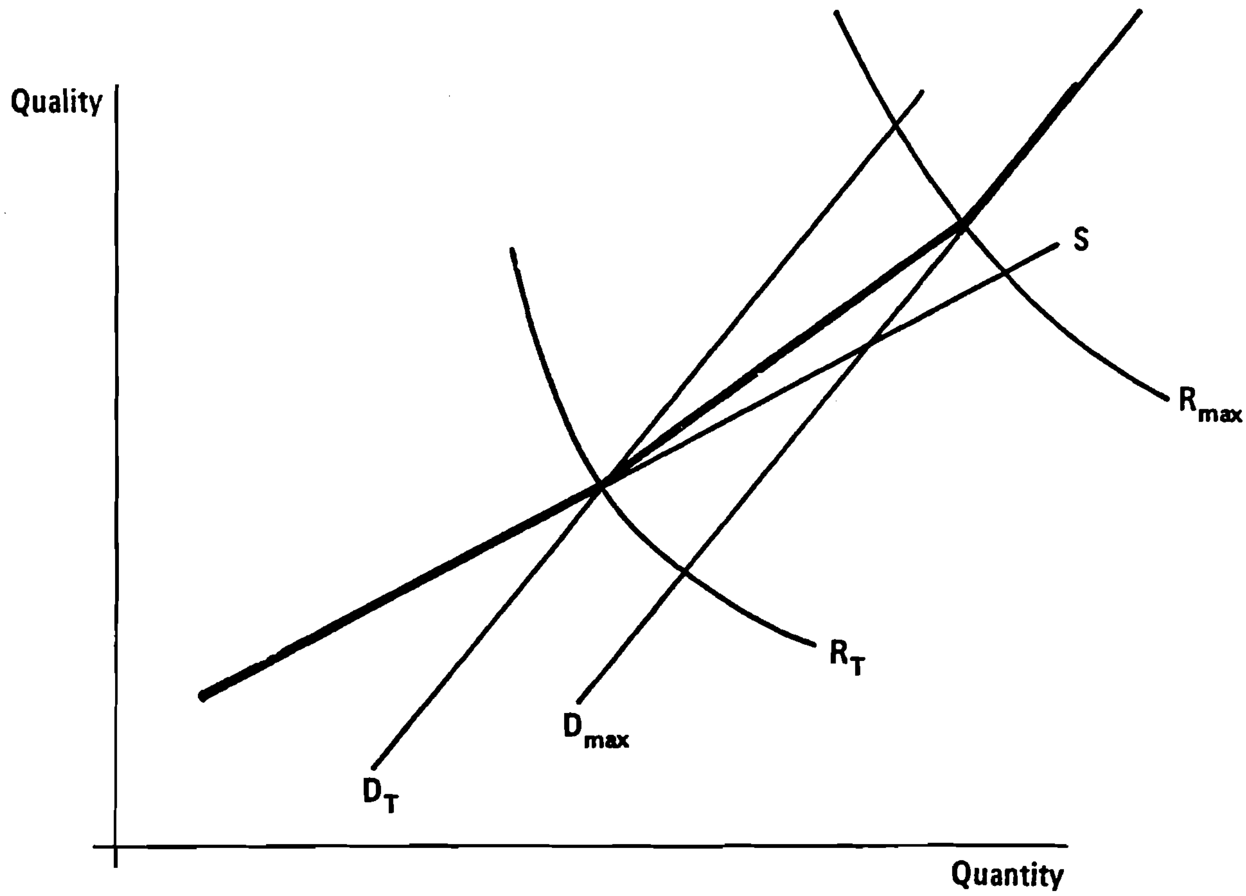

Figure 9. Effect of resources on use with demand creation.

6.2. A Formal Econometric Specification of the Three Regime Model

Let us consider the simple case in which P is fixed at $\overline{P}$. Demand can be written as

$$Q_D = Q(\overline{P}, X, A) \qquad (14)$$

where A is the accuracy of the advice given by the physician. For simplicity it is assumed that there is only one input M, so the production function is just

$$Q = Q(M) \qquad (15)$$

There are four critical levels of output Q:

1. Let Q_S^* be the level given by

$$\max_{Q,M} \pi = \overline{P}Q - W(M) \cdot M \qquad (16)$$

where W(M) is the marginal opportunity cost of physician input, an increasing function of M. Here Q is assumed to be unconstrained by demand, so that Q_S^* corresponds to the amount that a competitive profit-maximizing firm facing price $\overline{P}$ would choose to supply.

2. Let Q_{max} be the level given by

$$\max_{A} Q_D = Q(\overline{P}, X, A) \qquad (17)$$

So Q_{max} correspond to points along the D_{max} curve.

3. Let $\hat{Q}$ be the level given by

$$\max_{A,Q,M} V(\pi, A) \qquad (18)$$

subject to the production function and demand constraints.

4. Finally, let $Q_T = Q_D(\overline{P}, X, A_T)$ where A_T is "true" advice, i.e., the maximum possible value for A.

The three regimes can now be described.

Regime I: Excess demand. If, at $\overline{P}$, $Q_S^* < Q_T$, then $Q = Q_S^*$.

Here the amount that the physician is willing to supply, even if he is unconcerned about the accuracy of advice, is less than that demanded when fully accurate advice has been given. So he may as well provide accurate information, but there will be excess demand.

Regime II: Discretionary Demand Creation. If, at $\overline{P}$, $Q_S^* > Q_T$ and $\hat{Q}_S < Q_{max}$, then $Q = \hat{Q}_S$.

Demand is being created, and the amount of output depends on the preferences of the provider and the response of the patient to information.

Regime III. Maximal Demand. If, at P, Q_S^* Q_T and $Q_S \approx Q_{max}$ then $Q \approx Q_{max}$.

Here demand has been pushed out as far as it can, and so it tracks the D_{max} curve. The provider may not reach this regime, depending on his preferences and the consumer's response.

In Regime (I), use would be wholly explained by supply variables: output price, input price, provider preferences, and parameters of the production function. In Regime (III), use would be wholly explained by demand variables: output price, income, illness levels, tastes, with possibly some variables reflecting the ability of the typical patient to detect inaccurate advice (Pauly, 1980b). Finally, in Regime (II), use would depend on both demand and supply variables. In addition, any parameters describing the utility function V, which might reflect the willingness of the typical physician to trade off income and accuracy, should be included.

It would seem that maximum likelihood methods could be used to distinguish among these regimes, assuming that the system is otherwise identified.

7. A MODEL OF INTERACTIVE DEMANDS USING DIDD

7.1. Introduction

Now I want to consider a somewhat more complex but more realistic model in which there are two levels or types of providers -- primary providers and secondary providers. There will be some relationship between the demand for the services of one type of provider and the price or supply of services of the other type of provider. If neither market were ever in excess demand, the demand function would be specified in the usual way -- by including user prices of closely related goods in the demand function for each good. Such an approach has been used in the literature (Davis and Russell 1972), but it has also been remarked that some of these markets, especially the one for primary case, may be characterized by excess demand. Then, as Feldstein (1971) has noted, it is not appropriate to put price into the demand function. If these related markets are always characterized by excess demand, then the appropriate measure would be the one Feldstein actually uses -- the supply of such resources.

But if the markets for related goods are sometimes characterized by excess supply, sometimes by equilibrium, and sometimes by excess demand, a disequilibrium approach is called for. Moreover, if excess supply leads to demand manipulation, this possibility would need to be incorporated too. Even in non-fee-for-service, zero user-price markets, interaction effects are likely to to be present. Finally, it would seem that characterizing outputs by type of provider alone may be misleading, since a provider may produce some outputs that are complements to another output and other outputs that are substitutes for it.

For these reasons, it seems useful to outline a model with both stages or types of providers, and types of output. Such a model also permits a new use of DRAM.

7.2. Assumptions and Structure

It is assumed that there are two types of provider: a primary care provider and a secondary care provider. A person can obtain care from the second type of provider only after having received at least some services from the first type.

There are three kinds of output, with each kind being provided to patients of various characteristics. The second type of provider provides output C (consultants' services), a vector with elements representing the amount of C given to patients with each set of characteristics. The first type of provider produces two vectors of outputs: diagnosis and recommendation (D) and treatment (T). In order to receive any T or C, a patient must have at least some D; people usually must contact a primary provider to get secondary-level care. (In reality, a sizeable portion of specialists' business in the U.S. comes from patients who initiate the contact themselves without referral, but that will be ignored here.)

It will not be assumed either that patients always do what doctors recommend or that doctors always recommend what patients want. In this simple first version, I will however ignore the possibility of demand inducement.

To construct such a model, I assume that both primary and secondary providers (GP and S, respectively) receive fixed budgets. Such an assumption would represent salaried providers at both levels. The budgets or salaries specify total inputs (e.g., how many hours the provider must work) but not how inputs should be divided over outputs. Technical efficiency in production is assumed.

<u>The GP's Problem</u>. Let D be a vector if diagnostic services for various illnesses or conditions and T a vector of treatment services. The GP's problem is to maximize V(D,T) subject to a simple production function

$$M = D + T = D + tD$$

(where M is total hours and t is the number of treatments per diagnosis) and demand functions

$$D \leq D_D = f_1(X_1)$$

$$t \leq t_D = f_2(X_2)$$

(where X_2 and X_3 are demand variables) and a resource limit

$$M = \overline{M}$$

Given the chosen levels of D and t, the pattern of referrals to S for C depends on the mix of diagnostic and treatment services provided by the GP. That is

$$C_R = f(D, t)$$

where C_R is the number of referrals to the consultant and

$$\frac{dC_R}{dt} < 0$$

The sign on the total derivative captures two influences:

(a) when t is increased, patients have fewer conditions left untreated to refer to specialists

(b) when t is increased, given M, less time is available for D, so fewer new patients are seen, and fewer conditions discovered.

Thus D is complementary to C, while t (and T) is a substitute for C.

The Specialists' Problem. The specialists' problem is to maximize

$$V = V(C)$$

subject to a production function

$$g(C, S) = 0$$

where S is specialist resources and a demand function

$$C \leq C_D = h(C_R, X_3)$$

(where C_D is the number of speicalists' services demanded and X_3 is a set of variables reflecting the patients' compliance with recommendations) and a resource constraint

$$S = \bar{S}$$

The critical point then is that the queue from which the consultant selects, and the character of patients in that queue (in terms of their "need" for C) depends on decisions made by the GP. If the GP sees many different patients and treats them superficially, C_D will be large. If he sees only a few new patients, and treats them intensively, C_D will be small. It also means that one cannot predict a priori the impact of GP supply (an increase in aggregate $\bar{M}$) on C; that depends on how GP's allocate the initial resources between D and t. C will nevertheless depend on M, the number of GP's. and also on what motivates GP's. Hence, forecasting secondary level care requires consideration of GP numbers, resources, and preferences, as well as patient demands.

If, instead of receiving a fixed budget or salary, the GP receives fee-for-service reimbursement, then the level of M will not be fixed. Rather, it will depend upon the levels of P_D and P_T. Moreover, the mix of services will depend upon the relative fee level P_D/P_T. As P_D falls relative to P_T, the GP will take in fewer new cases, and treat the ones he does take more intensively himself, thus resulting in fewer referrals to specialist, and ultimately less specialist business. Here the pattern of use of specialists will depend on absolute and relative fee levels, as well as on whatever clinical standards physicians may value and the number of GP's.

8. EMPIRICAL SPECIFICATION OF DEMAND FUNCTIONS

In this section I will discuss the kinds of variables that one might want to include in demand functions. This will naturally raise the issue of disaggregation into types of medical care and types of providers, and the relationships across markets for different types of services.

The purpose of this discussion will be to indicate the general classes of variables to be used. The specific measures will obviously depend upon the institutional structure and data availability in various countries.

In broad terms, we can distinguish seven general types of influences on individual's demand for a particular type of medical care. These are

(1) Illness or state of health measures
(2) Money and non-money prices of the type of care
(3) Money and non-money prices of substitutes or complements to the particular type of care
(4) Income
(5) Advice or information from physicians and friends

(6) Demographic variables
(7) Other taste variables

A brief discussion of each type of influence is in order. I will concentrate on measures for predicting average use in a market area. Health status measures have been found to be important influences on individual use of medical care. How important they are in explaining differences across areas or in area means over time is a much more open question. In most developed countries, the most easily available health status measures -- infant mortality rates and death rates -- vary too little and/or are not especially associated with medical care use. But the measures used in individual studies have generally not been analyzed in aggregated data. Such measures include work or school-loss days, self-evaluated measures of health, and numbers and types of chronic conditions.

There is some circumstantial evidence to suggest that such morbidity measures, or even more refined disease-specific ones, may vary across populations. Roos,et al. (1977) found that a good deal of the variation in tonsillectomy rates within Manitoba could be explained by variation in the incidence of upper respiratory infections among children. Newhouse showed that, compared to using health and price measures, a planning method based on sociodemographic proxies gave an inferior explanation. So despite the fact that there is no perfect measure of the state of health, it would seem to be desirable to develop one, even one based on a sample rather than data from an entire population.

One of the strongest findings of research in health economics is that money-user price does affect use. However, the evidence on its magnitude or on the variation in that magnitude by type of disease is not conclusive. The multimillion dollar Health Insurance Experiment being conducted by the RAND Corporation in the U.S. is intended to estimate the magnitude of this response, as is the Health Insurance Survey developed by Wilensky and coworkers.

There is also strong evidence on the importance of time or travel costs to use of ambulatory care, and even some evidence that such variables matter for inpatient care.Variations in the quality of medical care have theoretically indeterminate effects on demand;there is,however,almost no confirmation of this.

The importance of close substitutes or complements has often been recognized, but it has been difficult to get confirmation of a prioi hypotheses: for example, it has been hypothesized that reducing the user price of outpatient care (including routine physician office visits) by broadening insurance coverage for such care would cause a reduction in the use of inpatient care. In almost every empirical example, however, it has been found that inpatient demand <u>increased</u> when coverage was extended to outpatient care, suggesting gross complementarity. (This is not to deny that there may be some specific kinds of outpatient services which are substitutes; it is only to suggest that the entire class of such services is likely to be complementary to inpatient services.)

The role of information from physicians has already been discussed. The role of information from other sources is usually handled by various proxies. Recent studies (Pauly, 1980b; Pauly and Satterhwaite, 1980) have provided empirical evidence that information is important in explaining the use of medical care. Income is probably more relevant when the user price is not zero (though it is normally correlated with time cost). The main point here is that some measure of permanent income, rather than income temporarily depressed by illness, should ideally be used.

Sociodemographic variables appear to be related to use and to health. For example, the mother's education has effects on health, but a less well-defined effect on use of medical care. Marital status and family size also affects the demand for care, though not in a way which is consistent across studies. What is obviously needed is a model of household behavior with regard to both health and medical care,and this has yet to be fully developed. However, the work of Grossman and associates comes closest to providing such a theory and tests of it.

"Taste" is in some sense just a label for our ignorance. There is, however, enough evidence for systematic variation in demand with variables that cannot themselves be causal, and enough unexplained variation, to give one pause. For example, length of hospital stay in the U.S. is lower in the West than elsewhere when all other variables have been controlled for, and this phenomenon has so far resisted explanation.

These sets of variables affect individual demands and, by aggregation, the total demand in the medical care market. The demand perceived to be facing any individual provider is obviously related to this total demand, but is also influenced by a variable which might be called the "degree of competition". All of my discussion to this point, and all of the HCS modeling at IIASA, has proceeded on the assumption that the relevant demand is the <u>total</u> demand for the area. This is probably proper for a fully centrally controlled system. But where individual providers have the ability to choose price, quality, or quantity for themselves, then the relevant demand curve is the <u>firm</u> level demand curve, not the aggregate demand curve.

For example aggregate demand for hospital admissions may be little affected by the level of amenities, but the choice of <u>which</u> hospital to use may be greatly affected. This means that the individual hospital may have very little scope for exercising its preferences as to amenities vs. clinical quality vs. volume of admissions.

The degree of competition among providers and its interaction with their motivation is too complex to be discussed in this paper; it is the subject of another forthcoming paper. The principal point to be made here, however, is that such considerations may be of considerable importance in decentralized systems.

9. CONCLUSION

To what extent could the suggestions in this paper be and in actual empirical studies involving DRAM? It would seem that there is relatively little difficulty in adding demand to the model in some way, although getting the specification exactly right, and determining whether relationships are identified, is a formidable task. In a similar way, adding the possibility of demand creation seems relatively straightforward.

While it is not very difficult to put together a workable analytical-forecasting model, and while it may be possible to even use a regression approach that is simpler than DRAM, accurate and useful forecasts depend on more than just the method or the model. They also depend on the type of data available and the quality of estimates obtainable from that data. There has been almost no concern in the various applications of DRAM with the accuracy of parameter estimates. While this was perhaps defensible when the only object was to illustrate the method, the continued use in the most recent papers of very small data sets [e.g., 10 observations to estimate 4 parameters in each of 3 output categories (Aspden, 1980)] cannot be likely to lead to very accurate forecasts. The danger is that the policymaker may reject the idea of the model because of the inaccuracy of the parameter estimates developed. It would be interesting to compare DRAM as a forecasting model with a different kind of model (e.g., to the ad hoc regression model described in the footnote in section 3 or to a simple trend projection), when both are fitted to the same set of data.

One possible and appropriate application of a DIDD-type approach would be to hospital-use data from Canada. Here the user price is zero, but the likelihood of mixtures of excess demand and equilibrium seems especially high. Moreover, there are sufficient numbers of geographically separated market areas to permit adequate statistical analysis. Another possibility would be to use information from some states in the United States where insurance coverage is virtually complete, where there have been state efforts to restrict hospital expenditures, and where information on hospital case-mix is available. New York, Michigan, Maryland, New Jersey, and some other eastern states could probably provide an adequate data base, although the availability of uniform information on admission by diagnosis in all such states would have to be investigated. In any event, such information is available for persons over 65.

Another interesting use would be to compare the coefficients for the same model estimated across countries. Do the demand parameters differ, or are they similar? What about the HCS utility function implied in the behavior under supply constrained regimes? And even what about identifiable correlates of the propensity of the political system to create shortages?

Adding considerations of production efficiency to DRAM in a very general way will probably be more difficult. The influences of various incentives on productivity are not well known, and the form they will take will vary with the institutional arrangements in various countries. The observable characteristics of providers which are related to productivity are not well known, and there may be considerable unexplainable variation in productivity across countries. Where a specific influence (e.g., the form of physician compensation in one country) is under discussion, it may be possible to develop a model to investigate it. In general, however, the degree of competition in markets is also not known and not easy to measure, nor is it simple to model the wide variety of forms that can be taken by compensation schemes and hospital-GP-consultant relationships.

All of these observations suggest that modification of DRAM will be a complicated and time-consuming task. Moreover, at the end one will only be able to tell the high-level decision-maker what his actions will do, not what they ought to do. Nevertheless, the process of crawling before walking seems to be an essential feature of most systems which turn out ultimately to be useful, and this intellectual system may be no exception.

REFERENCES

Aspden, P. (1980) *The IIASA Health Care Resources Allocation Submodel: DRAM Calibration with Data from the South West Health Region, UK.* WP-80-115. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Davis, K., and L. Russell (1972) *The Substitution of Hospital Outpatient Care for Inpatient Care.* Review of Economics and Statistics 54:1003-1015.

Evans, R. (1974) *Supplied-Induced Demand: Some Empirical Evidence and Implications* in M. Perlman, ed. The Economics of Health and Medical Care, London: Macmillan and Co.

Fair, R.C., and D.M. Jaffee (1972) *Methods of Estimation for Markets in Disequilibrium.* Econometrica 40:497-514.

Fair, R.C., and H.H. Kelejian (1974) *Methods of Estimation for Markets in Disequilibrium - A Further Study.* Econometrica 42:177-190.

Feldstein, M. (1967) *Economic Analysis and Health Services Efficiency.* Amsterdam: North Holland Publishing Company.

Feldstein, M. (1971) *Hospital Cost Inflation: A Study of Nonprofit Price Dynamics.* American Economic Review, December, 1971, 61:853-872.

Feldstein, M. (1974) *Econometric Studies of Health Economics,* in Intriliqator and Kendrick, eds. Frontiers of Quantitative Economics II. Amsterdam: North Holland Publishing Co.

Feldstein, M. (1977) *Quality Change and the Demand for Hospital Care.* Econometrica 45 (7):1681-1702.

Fuchs, V., and M. Kramer (1973) *Determinants of Expenditures for Physicians Services in the U.S., 1948-68*. Washington: National Center for Health Services Research.

Fuchs, V. (1978) *The Supply of Surgeons and the Demand for Surgical Operations*. Journal of Human Resources, Supplement 13:35-56.

Gibbs, R. (1978a) *The IIASA Health Care Resources Allocation Submodel: Mark 1*. RR-78-08. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Gibbs, R. (1978b) *A Disaggregated Health Care Resource Allocation Model*. RM-78-01. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Harris, D.H. (1975) *Effect of Population and Health Care Environment on Hospital Utilization*. Health Services Research 10:229.

Hughes, D.J., and A. Wierzbicki (1980) *DRAM: A Model of Health Care Resource Allocation*. RR-80-22.

Laffont, J.J., and R. Garcia (1977) *Disequilibrium Econometrics for Business Loans*. Econometrica 45:1187-1204.

Maddala, G.S., and F.D. Nelson (1974) *Maximum Likelihood Methods for Models of Markets in Disequilibrium*. Econometrica 42:1013-1030.

Newhouse, J. (1970) *Toward a Theory of Non-profit Institutions: An Economic Model of a Hospital*. American Economic Review 60:211-226.

Newhouse, J. (1974) *Forecasting Demand for Medical Care for the Purpose of Planning Health Services*. RAND R-1635-OEO, December 1974. Also in Thompson and Bailey, eds. Systems Aspects of Health Planning 1975.

Newhouse, J., and C. Phelps (1976) *New Estimates of the Price and Income Elasticities of Medical Care Services*, in R. Rosett, ed. The Role of Health Insurance in the Health Services Sector. New York: National Bureau of Economic Research.

Portes, R., and D. Winter (1980) *Disequilibrium Estimates for Consumption Goods Markets in Centrally Planned Economies*. Review of Economic Studies 60:137-157.

Pauly, M.V. (1979) *The Ethics and Economics of Kickbacks and Fee Splitting*. Bell Journal of Economics 12:133-142.

Pauly, M.V. (1980a) *The Effect of Hospital Resources on Physician Productivity and Prices*. Unpublished paper. February, 1980.

Pauly, M.V. (1980b) *Doctors and Their Workshops.* Chicago: University of Chicago Press.

Pauly, M.V., and M. Satterthwaite (1980) *The Effect of Physician Supply on Physician Prices.* Unpublished paper. Evanston, Illinois: Northwestern University.

Reinhardt, U. (1972) A Production Function for Physicians' Services. *Review of Economics and Statisics* 54(1):55-66.

Reinhardt, U., M. Pauly, and P. Held (1979) *Financial Incentives in Group Medical Practice,* in Held and Reinhardt, eds. Analysis of Economic Performance in Medical Group Practices, Final Report, PR 79-05. Princeton, N.J.: Mathematica, Inc.

Roemer, M.I. (1961) *Bed Supply and Hospital Utilization: A Natural Experiment.* Hospitals 35:36-42.

Roemer, M.I., and M. Shain (1959) *Hospital Utilization Under Insurance.* Hospital Monograph Series No.6. Chicago: American Hospital Association.

Roos, N., L. Roos, and P. Henteleff (1977) *Elective Surgical Rates -- Do High Rates Mean Low Standards.* New England Journal of Medicine 297:880-884.

Rosenthal, G. (1964) *The Demand for General Hospital Facilities.* Monograph 14. Chicago: American Hospital Association.

Rousseau, J.M., and R.J. Gibbs (1980) *A Model to Assist Planning the Provision of Hospital Services.* CP-80-3. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rutten, F.F.H. (1979) *Physician Behaviour: The Key to Modelling Health Care Systems for Government Planning.* WP-79-60. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Shigan, E. et al. (1979a) *Health Care Systems Modeling at IIASA: A Status Report.* SR-79-4. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Shigan, E. et al. (1979b) *Modeling Health Care Systems: June 1979 Workshop Proceedings.* CP-79-15. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Sloan, F., and F. Feldman (1977) *Monopolistic Elements in the Market for Physicians Services.* Paper given at a Federal Trade Commission Conference on Competition in the Health Care Sector, Washington, June, 1977.

RECENT PUBLICATIONS IN THE HEALTH CARE SYSTEMS TASK

Shigan, E.N.,ed. (1978) *Systems Modeling in Health Care.* Proceedings of an IIASA Conference, November 22-24, 1977 (CP-78-12).

Gibbs, R.J. (1978) *The IIASA Health Care Resources Allocation Sub-Models: Mark 1* (RR-78-08).

Gibbs, R.J. (1978) *A Disaggregated Health Care Resource Allocation Model* (RM-78-01).

Kaihara, S., N. Kawamura, K. Atsumi, and I. Fujimasa (1978) *Analysis and Future Estimation of Medical Demands Using A Health Care Simulation Model: A Case Study of Japan* (RM-78-03).

Fujimasa, I., S.,Kaihara, and K. Atsumi (1978) *A Morbidity Subnodel of Infectious Disease* (RM-78-10).

Propoi, A. (1978) *Models for Educational and Manpower Planning: A Dynamic Linear Programming Approach* (RM-78-20).

Klementiev, A.A., and E.N. Shigan (1978) *Aggregate Model for Estimating Health Care System Resource Requirements (AMER)* (RM-78-21).

Hughes, D.J. (1978) *The IIASA Health Care Resource Allocaiton Sub-Model Mark 2: The Allocation of Many Different Resources* (RM-78-50).

Hughes, D.J. (1978) *The IIASA Health Care Resource Allocation Submodel: Estimation of Parameters* (RM-78-67).

Hughes, D.J. (1979) *A Model of the Equilibrium Between Different Levels of Treatment in the Health Care Systems: Pilot Version* (WP-79-15).

Fleissner, P. (1979) *Chronic Illnesses and Socio-Economic Conditions: The Finland Case 1964 and 1968* (WP-79-29).

Shigan, E.N., D.J. Hughes, P. Kitsul (1979) *Health Care Systems Modeling at IIASA: A Status Report* (SR-79-4).

Rutten, F.F.H. (1979) *Physician Behaviour: The Key to Modeling Health Care Systems for Government Planning* (WP-79-60).

A Committee Report (1979) to IIASA by the participants in an Informal Meeting on *Health Delivery Systems in Developing Countries* (CP-79-10).

Shigan, E.N., P. Aspden, and P. Kitsul (1979) *Modeling Health Care Systems: June 1979 Workshop Proceedings* (CP-79-15).

Hughes, D.J., E. Nurminski, and G. Royston (1979) *Nondifferentiable Optimization Promotes Health Care* (WP-79-90).

Rousseau, J.M., R.J. Gibbs (1980) *A Model to Assist Planning the Provision of Hospital Services* (CP-80-3).

Fleissner, P., K. Fuchs-Kittowski, and D.J. Hughes (1980) *A Simple Sick-Leave Model used for International Comparison* (WP-80-42).

Aspden, P., R. Gibbs, and T. Bowen (1980) *DRAM Balances Care* (WP-80-43).

Aspden, P., and M. Rusnak (1980) *The IIASA Health Care Resource Allocation Submodel: Model Calibration for Data from Czechoslovakia* (WP-80-53).

Kitsul, P. (1980) *A Dynamic Approach to the Estimation of Morbidity* (WP-80-71).

Shigan, E.N., and P. Kitsul (1980) *Alternative Approaches to Modeling Health Care Demand and Supply* (WP-80-80).

Hughes, D.J., and A. Wierzbicki (1980) *DRAM: A Model of Health Care Resource Allocation* (RR-80-23).

Aspden, P. (1980) *The IIASA Health Care Resource Allocation Submodel: DRAM Calibration for Data from the South West Health Region, UK* (WP-80-115).

Mayhew, L., and A. Taket (1980) *RAMOS: A Model of Health Care Resource Allocation in Space* (WP-80-125).

Mayhew, L. (1980) *The Regional Planning of Health Care Services: RAMOS and* $RAMOS^{-1}$ (WP-80-166).

Nanjo, Z. (1980) *A Simple Mehtod of Measuring the Increase of Life Expectancy when a Fixed Percent of Deaths from Certain Causes are Eliminated* (CP-80-35).