Competition and Market Failure in the Hospital Industry: A Review of the Evidence

Richard J. Arnould
Lawrence M. DeBrock

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign
Competition and Market Failure in the Hospital Industry: A Review of the Evidence

Richard J. Arnould, Professor
Department of Economics

Lawrence M. DeBrock, Assistant Professor
Department of Economics

The authors are grateful to David Lee, University of Illinois, for his substantial input and assistance on this article.
ABSTRACT

The task of this paper is to synthesize certain of the published research findings in health economics pertaining to hospitals. In particular the theme of the survey centers around the use of concepts of market efficiency common to economists in the field of industrial organization to explain performance of hospital markets. A substantial amount of research has been conducted in recent years related to various aspects of performance of producers and consumers of health care services. These studies have taken a number of directions. Recent actions by regulators and third party payers that provide more incentive for efficient health care delivery systems have generated interest in the nature, causes, and remedies of market failure in this industry. Therefore, the thrust of this paper places emphasis on the delivery of health care services emphasizing hospital services, within a market context and attempts to analyze the existence and extent of market failure within health care markets. The general conclusions of the paper are that the cost structure of hospitals does not warrant there being monopoly producers of hospital services. Therefore, market failure is the result of externalities in the form of moral hazard, most of which is generated through the organization and operation of risk shifting to third party payers. Additional market failure is generated by a variety of attempts by governments to regulate the health care sector. The paper concludes with a discussion of changes that are taking place in the risk sharing formulas that define the methods through which health care services are purchased, the effects these changes are having on market efficiency, and, finally, with a description of empty boxes for which future research is needed.
1. INTRODUCTION

The hospital industry is a large and complex segment of the economy of the United States. The more than 6500 acute-care hospitals vary widely in size and services provided, generating significant interest in the efficiency of the industry. However, an analysis of the industry from the perspective of an efficient industry structure has been generally lacking. Various forms of regulation and the behavior of third party payers are alleged to have reduced or eliminated the incentives for efficient markets, resulting in an industry that is inefficiently organized. Recent actions by regulators and third party payers that provide more incentive for efficient health care delivery systems has generated interest in the nature, causes and remedies of market failure in this industry.

The overriding task of this paper will be to synthesize certain of the published research findings in health economics pertaining to hospitals. In particular, the theme of the survey will center around concepts of market efficiency used by economists in the field of industrial organization. An immediate problem confronted by the authors in carrying out this task is that of coming to an acceptable definition of the field of industrial organization. Bain, if not the founder, certainly popularized the field, states that industrial organization deals with the organization and working of markets (Bain, 1968, p. 1). Scherer states that industrial organization "...is concerned with how productive activities are brought into harmony with society's demands for goods and services through some ongoing mechanism such as a free market, and how variations and imperfections in the organizing mechanism affect the degree of success achieved by producers in satisfying society's wants." (Scherer, 1980, p. 1) Stigler, on the other hand, states that "there is no such subject as industrial organization. The courses taught under this heading have for their purpose the understanding of the structure and behavior of industries...of an economy...[and deal] precisely with the content of [micro] economic theory." (Stigler, 1968, p. 1) Clearly, these definitions are very broad, encompassing most of microeconomic theory (including welfare economics), those aspects of macroeconomic theory that pertain to the manner in which macroeconomic activities feed back to influence market and firm performance, and some aspects of marketing.

Traditionally, economists in industrial organization have compared various market outcomes against those that would hold under a competitively structured market. The competitive market is used as the norm because it generates an equilibrium which maximizes various criteria of economic welfare. Deviations from the structural characteristics of competitive markets were used to predict deviations from competitive performance, the latter being considered as deviations from welfare optimization (Bain, 1949, p. 3). Stigler, and to some extent Scherer, stress the weaknesses in placing too much reliability on market structure as a predictor of economic performance. Instead, those authors hold that one should place greater reliance on market behavior or conduct, the latter being generally defined to include policies.
pertinent to price and product determination

More recently, a body of literature drawn together by Baumol, Panzer, and Willig (B-P-W) has introduced a new approach to evaluate market performance. The thrust of this approach to analysing markets is to divorce itself from the structural norms of the competitive model that pertain to numbers and size distribution of firms in an industry by placing increased emphasis on entry conditions as predictors of market performance. Using stronger criteria for free entry than did Bain or Stigler, they introduce the concept of a perfectly contestable market, defined as a market in which entry is free, potential entrants have access to the same market demands and production technology as the incumbent firms, and potential entrants evaluate the desirability of entry on incumbents' pre-entry prices (Baumol, Panzer, and Willig, 1982, pp. 4-8). This approach heavily rests on the cost structures of single- and multi-product firms.

Clearly, these definitions and approaches leave the field too broadly defined to be treated in one paper, even as it may apply to a single industry. Therefore, an attempt is made here to use elements from both the traditional and the B-P-W approaches to synthesize research on the efficiency of hospital markets. The structure of hospital markets will be described in the next section. After considering current hospital market structures, possible causes of the structural patterns are examined. Most prominently, cost functions of hospitals are discussed to determine whether existing market structures conform to or deviate from those predicted by the cost functions. However, a significant caveat must be made. Most research in which structural information has been reported was not oriented to analyses of market behavior. Therefore, structural information on hospital markets relates to geographic or political areas that may or may not correctly delineate economic markets. Section III will examine the issue of market failure, that is, the possibility of some characteristic of the industry that prevents the market from arriving at an efficient outcome. Section IV examines attempts to use regulation to correct market failure. The final section discusses attempts to provide more market incentives for greater economic efficiency. In this section, we discuss some of the ways in which both the public and private sectors are responding to these changes.

The overriding theme of the work reviewed in this paper points to the argument that the regulatory environment in which hospitals operate and the nature of third party payment systems have provided very limited incentives for efficient production and consumption of health care resources. This appears more true for earlier time periods, with more recent years witnessing some evolution. Changes in the regulatory environment are necessary to generate incentives for changes both in market structure, and producer and consumer behavior. Policies that shift increasing amounts of financial risk from the third party payers to the consumers and producers should continue to bring increased efficiency to hospital markets. The significance of this is that the underlying problem does not appear to be inherent in the structure of
the markets. There is little evidence of economies of scale that would necessitate monopoly market structures. Rather, the underlying problems may actually be found in the nature of regulatory and reimbursement policies.

II. THE STRUCTURE OF THE HOSPITAL INDUSTRY

The hospital industry is of particular interest to students of industrial organization because it embodies a number of unique structural and organizational characteristics. However, the structure of hospital markets was not a subject of extensive discussion until recently. Interest grew in this area as researchers began addressing the question of whether market oriented incentives would result in more efficient markets. A thorough understanding of the structure of an industry is fundamental to answering such a query. Most structural data on hospitals have been gathered by the Health Planning authorities and other regulatory bodies, whose interests have naturally been within political as opposed to market boundaries. However, useful descriptive statistics of the hospital component of the health care sector can be drawn from these data.

A. Description of Hospital Structures

Hospitals account for the largest segment of health care expenditures. In 1983, $147.2 billion went to short-term acute care hospitals; this amounted to 43.9 percent of the $335.4 billion spent in the U.S. on health care (Health Care Financing Review, 1985). The hospitals are configured in a variety of sizes, shapes, and ownership forms. The American Hospital Association places under the rubric of community hospitals all non-federal, short-term general and other special hospitals whose facilities and services are available to the public. In 1983, 5,783 of the 6,888 hospitals in the United States were classified as community hospitals. The remaining hospitals consist of federal hospitals, long term care hospitals, hospital units of institutions, psychiatric hospitals, hospitals for tuberculosis and other respiratory diseases, chronic disease hospitals, institutions for the mentally retarded, and alcoholism and chemical dependency hospitals. The community hospital category consists of 3,347 not-for-profit institutions (the bulk of which are affiliated with religious organizations), 757 proprietary hospitals, and 1,679 hospitals owned by state and local governments. As shown in Table 1, the number of hospitals, both community and in total, has decreased between 1975 and 1983. There has been a decline in the number of hospitals in all categories in the last decade with the exception of proprietary hospitals. The number of proprietary hospitals declined during the early years of the decade, but has shown increases since 1980.

Hospitals also show considerable variation in size, a variable usually measured in terms of bed capacity. It is shown in Table 2 that the average number of beds in a community hospital has been increasing.
since 1975: the beds per hospital grew from 160.3 in that year to 175 in 1983. However, further examination indicates that the distribution of hospital size is extensive. Thirty-five percent of all community hospital bed capacity is provided by the top 10 percent of the hospitals. This group has an average of more than 300 beds each. Its workload is also significant, treating 34 percent of all admissions and accounting for 43 percent of total community hospital expenditures. At the other extreme, community hospitals with capacity of less than 100 beds make up 45 percent of the total number of hospitals, yet represent only 14 percent of total bed capacity and handle just 13 percent of total admissions. In fact, there are approximately 225 community hospitals with 25 or fewer beds and 1200 with 50 or fewer beds. The economic effects of this skewed distribution are unclear. These numbers are national aggregates and provide very little guidance as to market competition. Still, it appears that a relatively few hospitals dominate the statistics.

A recent development in hospital structure has been the emergence of multi-hospital systems. Ermann and Cabel (1984) report that the number of multi-hospital systems has grown from 202 in 1975 to 256 in 1982, or at an average annual rate of 3.4 percent. The percent of short-term acute care hospitals in the U.S. belonging to multi-hospital systems grew to 33 percent in 1982 from 25 percent in 1978. The average system consisted of 7.5 hospitals and over 1400 beds. Again, a pattern of skewed distribution arises. The 34 of these systems that are investor owned are much larger than the not-for-profit systems, the former having an average of 23 hospitals and 3100 beds per system.

It is important to consider the economic motive for the rise of these multi-hospital systems. Their primary potential advantages are better access to capital, economies of scale, system diversification, access to personnel and management, and system planning and organization. Contract management, another growing phenomenon among hospitals, may provide many of these benefits without actually requiring a full merger of the hospitals. However, contract management is not likely to provide improved access to capital markets. Such access is increasingly necessary for hospitals in order to maintain their capital facilities and participate in integration into other areas of the health care sector. For example, 55.1 percent of the multi-systems own ambulatory care facilities, 50 percent own nursing homes, 12 percent own HMO's, 72 percent own physician office buildings, and 43 percent own health care management consulting businesses (Ermann and Cabel, 1984, p 55). Comparable data are not available for hospitals not belonging to systems. However, it is generally believed that single facility hospitals are participating in the movement toward integration, the causes of which will be discussed more fully in a later section of this paper.

The varying mixes of services offered by short-term acute care hospitals truly makes them multi-product firms. The American Hospital Association identifies 46 unique services. The Health Care Financing Administration (HCFA) divides the outputs of hospitals into 467...
diagnostic related groups (DRG) for purposes of classifying hospital outputs provided to recipients of Medicare. Therefore, it is difficult to categorize and, thus, measure the outputs of hospitals. However, an example of the disparity of services offered by different hospitals, consider the fact that while all short-term acute care hospitals provide basic nursing services and emergency services, only 10.5 percent of all hospitals in the U.S. maintain open heart surgery facilities. Thirty-four percent operate CT scanners, 15.8 percent provide cardiac catheterization, 66.2 percent operate blood banks, 16.0 percent provide radiation x-ray therapy, and only 3.8 percent of all hospitals provide organ transplants.

Larger hospitals generally offer a larger spectrum of services, including more tertiary care services. The widest range of services is usually offered by university hospitals and regional tertiary care facilities, e.g., the hospitals used by physicians from the Mayo Clinic. This seems to be changing. Joskow points out that the range of medical services available in smaller hospitals has increased in the last decade in response to competition for the increased flow of medical specialists into health care markets (Joskow, 1981).

3. Hospital Market Structure

The above description of the overall economic activity of the hospital industry did not offer any evidence as to the competitiveness of the industry. In order to move toward that goal, data on market statistics are necessary. It is very useful to provide a description of the manner in which hospital markets are delineated prior to discussing hospital market structure. Political boundaries are straightforward; economic boundaries are more difficult. The origin of patients using particular hospitals or groups of hospitals has been widely used in antitrust cases as an attempt to define economic markets; this is an application of the Elzinga and Hogarty principles of geographic market delineation (Elzinga and Hogarty, 1973). Elzinga and Hogarty contend that to appropriately delineate a geographic market, the analyst should conduct a two-step procedure. First, they suggest that in a properly delineated market, only a small proportion of the product (hospital care) should be imported. Known as the "little in from outside" (LIFO) element, this corresponds to having a small number of patients living within the geographic area going outside to purchase hospital services. Second, they suggest that only a small proportion of the product should be exported; i.e., only a small proportion of the hospital services should be provided to patients living outside the geographic bounds of the market (LOFI). The obvious problem with this approach is that it provides for no exact percentage at which each of the two elements is satisfied. Elzinga and Hogarty suggest values of at least 75 percent; other economists have suggested higher values.

One of the authors of this paper has used this technique to delineate a number of hospital markets as a component of antitrust analyses. However, there is no known delineation of a large sample of
hospital markets. Certain generalizations can be made from the limited information. First, the largest segment of patients at most hospitals reside in relatively close proximity to the hospital. Second, a large number of patients that reside close to one hospital utilize other hospitals. This phenomena is greater when the nearby hospital is small. However, it still exists to a significant degree when both hospitals are large. The result of these generalizations is that hospital markets encompass large regional areas. Small hospitals (usually offering fewer services than large community and/or teaching hospitals) compete directly with large hospitals for the services offered by both. There is a significant flow of patients living close to smaller hospitals to the larger hospitals located further from their residences. This may be due to the larger facility's ability to offer additional equipment and personnel support in case of complications. Similarly, there is a significant flow of patients living near one large hospital to a more distant hospital within the nearby metropolitan area.

Rural areas offer a different picture. Often the local hospital is small, with the ability to provide only limited services. Patient flow follows the patterns of larger metropolitan areas even though distances may be greater. The rural hospital is then in direct competition with other small and larger hospitals in the nearest communities. In fact, in sparsely populated states such as Nebraska, Kansas, etc., patients travel significant distances to purchase hospital services, often leapfrogging cities that have hospitals offering the needed services to purchase those services at regional tertiary care centers. This latter factor complicates the problem of market delineation in both rural and metropolitan areas. Major tertiary care centers, such as university hospitals, serve patients from a very large geographic area, possibly encompassing the entire nation. In some cases these centers are willing to bid for patients of HMO's in need of specific services such as open heart surgery. These patients are routinely transported a great distance to the hospital, thus placing the hospital in direct competition with a large number of highly geographically-dispersed hospitals. The extent of patients leaving any particular geographic area to utilize these major tertiary care centers will determine the degree of market overlap and, therefore, whether or not the geographic areas should be included in the same market.

After considering all of the complexities of determining the relevant market, it is necessary to draw some practical conclusions. If viewed from the perspective of the hospital-patient market (as we have been discussing), hospital markets for most services probably encompass single standard metropolitan statistical areas. They also include communities with no nearer SMSA's, even though some of the latter of which also contain hospitals (Lynk. 1984). For the more exotic services offered by tertiary care centers, the market may encompass multiple metropolitan areas as well as lower populated regions around those areas. Therefore, patients within most communities having a single hospital are within reasonable proximity, as exhibited by patient origin analysis, to other alternative hospitals. This holds even for people in more sparsely populated sections of the country. In fact, Farley
reported that only 265 of the 6000 community hospitals in the U.S. meet the criteria for classification as sole community hospitals set forth by the Tax Equity and Fiscal Responsibility Act of 1982 (TEFRA) and the Social Security Amendments of 1983 (Farley, 1985). However, by the strict Elzinga and Hogarty LIFO and LOFI tests, some of the hospitals would not qualify as having monopolies.

This general description of the problems of delineating the boundaries of market in which hospitals compete provides a caveat to the structural information available. That information, generally available on the basis of hospital bed concentration within SMSA's, will tend to underestimate the true boundaries of the market and, therefore, overstate the levels of market concentration. The average two-firm concentration ratio for 211 SMSA's is 0.584. There is a large amount of variation in this value; large metropolitan areas tend to have very low two-firm concentration ratios. For example, that ratio is 8 percent for Chicago, 7 percent for New York, and 10 percent for San Francisco. At the other extreme, smaller metropolitan areas have high two-firm concentration ratios. That ratio is 75 percent for Pensacola, Florida, 72 percent for Topeka, Kansas, and 73 percent for Sioux Falls, South Dakota.

Certain tentative conclusions can be drawn from this structural information. First, there are very few, if any, hospitals that operate in monopoly markets. Second, hospitals in large metropolitan markets approach an atomistic structure with a large number of hospitals with very small market shares. Third, smaller metropolitan areas tend to be somewhat more concentrated, approaching oligopolistic structure. However, hospitals in these areas usually compete with regional tertiary care centers in larger metropolitan areas. In general, these conditions indicate that hospital markets should have no structural reason to exhibit anti-competitive outcomes.

C The Cost Structure of Hospitals

The structure of hospital markets provides some evidence of the nature and extent of competitive performance that could exist in those markets. Before moving on, it is important to examine various possible causes of market structure. The nature of the cost structure of the firms often dictates how an efficient market is structured and, therefore, the outcome of other performance dimensions.

Estimates of the cost functions of hospitals have been thoroughly reviewed elsewhere (Feldstein, 1974; Jacobs, 1974; Berki, 1972; Hetty, 1969; Cowing, et al., 1983a). Therefore, only general conclusions will be drawn from these studies. Early cost studies concentrated on estimating the degree of economies of scale associated with hospitals. These investigations used very aggregate indices of hospital output, e.g., admissions, bed capacity, bed capacity adjusted for occupancy rates, and bed days of services provided. These early cost studies found evidence of significant economies of scale, usually up to a
hospital size of approximately 500 beds (Cowing, et al., 1983a). Feldstein (1974) concluded that there are neither significant economies or diseconomies of scale. In fact, refinements of these early studies indicate some economies of scale up to a relatively small size, usually considered to be in the vicinity of 100-150 beds. These studies have been criticized for improper treatment of long run and short run effects. However, Feldstein's (1974) second major conclusion is that the short run marginal cost of additional patient care is much lower than average cost.

The studies have also been criticized for failure to account for product mix and quality. Although quality remains largely untreated, recent cost studies have used various methods to account for differences in product mix. Differences in service mix may result in economies of scope, i.e., cost advantages to multi-product firms. Definitionally, economies of scope occur when the product mix is such that the resource cost of producing the bundle of products jointly are less than the sum of the resource costs of producing them separately (B-P-W, 1980). A scaler index developed from a variety of weighting systems has been used in most studies that account for case mix (Pauly, 1978; Rafferty, 1971; Lave, Lave, and Silverman, 1972; Feldstein and Schuttinga, 1977; Luke, 1979). The consensus of these studies is that case-mix effects explain between one-half and two-thirds of the cost variation among hospitals. Cowing and Holtman (1983b) used a translog functional form to estimate the short run cost function of hospitals. This method assumes no specific functional form for the cost function and so uses the flexible form, the translog. It has an advantage in that it provides for a direct measurement of any degree of economies of scope that may result from the multi-service outputs of hospitals. Cowing, et. al found only limited evidence of economies of scope. They included the stock of capital in their model to capture long run equilibrium effects. The empirical results lead to the conclusion that hospitals have too much capital. However, there are two significant weaknesses in this research. First, the translog estimation process requires a very large sample size to be efficient. Cowing, et al could divide all hospital outputs into only five very broad categories, each of which contain very diverse procedures ranging from those that are common and simple to very sophisticated procedures. Generally, only larger hospitals offer the more sophisticated procedures. Therefore, the results of such an estimation may not reveal much of the true behavior of the multi-product cost functions. Second, they provide only a limited ability to estimate long run conditions, the results of which do no relate directly to the product mix of the hospital.

Many gaps remain in the knowledge this literature has generated concerning hospital cost structures. New econometric techniques have been developed to deal with the multi-service nature of hospitals. Also, data bases are being generated that provide hospital output categorized into 467 diagnostic related groups. The sample sizes of these data are becoming large enough to meet the needs of the new econometric estimation techniques. Therefore, new research findings should be forthcoming in this area that answer important questions. The
more obvious examples include estimates of appropriate configurations of hospitals, the effects of hospital mergers, the effects on efficiency of specialized service centers (e.g., outpatient surgery centers and ob-gyn centers), the differences in capital costs between not-for-profit and proprietary hospitals and the effects of these differences on case mix, to mention a few. However, the evidence of only limited economies of scale suggests that hospital markets, except for those that are very small, need not be monopolistically structured.

III. HOSPITAL MARKET BEHAVIOR AND MARKET FAILURE

Hospital markets have attracted the attention of industrial organization economists, as well as those of other specialized areas of economics, largely because of substantial evidence of market failure in those markets. Expenditures for health care increased at an average annual compound growth rate of 13.9 percent between 1960 and 1979, while the average rate of increase in the overall CPI was only 5 percent over the same period (Joskow, 1981, p. 13). This is not meant to imply that all inflation in the hospital industry is the result of market failure. Some of the price increases are the natural result of increased demand for hospital services brought about by growth in the population, changes in the demographic mix of the population and changes in real income. Even more important may be the continuing expansion of insurance coverage. This coverage may result in substantial waste if insurance contracts do not generate risk sharing mechanisms that provide for efficient consumption and production of health care services. Moving to the production side, some health care cost increases may be the result of changes in input costs. Rapidly changing technology in hospitals contributed greatly to these cost increases. Finally, it is always possible that some component of the cost increases may result from internal inefficiencies in the operation of the hospitals.

Other evidence of market failure is provided by the significant amount of excess capacity in the hospital industry. Occupancy levels in U.S. hospitals ranged from less than 10 percent to over 90 percent in 1983. Over 18 percent of the hospitals had occupancy rates of less than 50 percent while only 22 percent had rates greater than 80 percent in 1983. Those rates have been declining as a result of reductions in admissions and average lengths of stay (ALOS) in recent years. The ALOS varies from 10.0 days in Omaha, Nebraska, to 5.9 days in El Paso, Texas. An AHA sample of states indicates that the ALOS varies from a high of 9.6 days in New York to a low of 5.4 days in Utah. Admissions per thousand population and cost per admission vary similarly. As with the case of price increases, clearly not all excess capacity or variation in ALOS, hospital admissions, or charges result from market failure. However, some levels of occupancy are too low to be explained by peak load problems. Variations in ALOS, hospital admissions, and charges are not all explained by differences in demographic characteristics of the patient population and input costs. In fact, a former Secretary of Health, Education, and Welfare testified in 1979 that almost $7 billion of waste in the health care sector resulted from excess capacity and
unnecessary duplication of facilities (Joskow, 1981).

In this section we will discuss some possible causes of market failure. We do so by examining the environment in which hospitals operate. We will concentrate on the impact on hospital behavior of the relationship between third party payers, physicians and hospitals and the broader regulatory environment in which they operate.

A. General Behavior of Hospitals

Hospitals do not necessarily follow the single objective of profit maximization generally attributed to most firms in the economy. There are a number of reasons for this. As stated earlier in this paper, a large number of hospitals are not-for-profit. In fact, many markets are serviced solely by not-for-profit hospitals. Therefore, due to this phenomenon, and other market imperfections discussed later in this paper, it is possible for hospitals to operate according to objective functions that contain criteria other than profit maximization.

Many of the early attempts to model hospital behavior postulated that hospitals maximize output. These models recognized that to maximize output the hospitals compete for high quality attending physicians. Reder argued that physicians desire hospitals that are fully equipped and permit the physician to treat a wide range of ailments as the physician's competence permits (Reder, 1965). Newhouse formalizes this contention by placing quality in the hospital's objective function and constraining the hospital to a break-even solution (Newhouse, 1970). Assuming that increases in quality will increase demand for the hospital's services at a rate less than the increase in operating costs guarantees an equilibrium at a finite level of output. Feldstein (1971) attempted to explain hospital cost inflation in a model that followed the pattern of Newhouse (1970). This model treated quality as an explicit function of the hospital's labor and capital inputs. Hospital services are considered a free good to attending physicians. Therefore, physicians always demand more hospital inputs (quality).

Lee argues that hospital administrators attempt to maximize their status relative to competitors. This is achieved through defensive actions whereby the objective is to minimize the difference between the actual and desired inputs of the hospital subject to a break-even condition. Therefore, hospitals may increase expenditures on inputs without increasing output or revenues due to the defensive nature of the hospital (Lee, 1971).

All of the models described to this point depict the hospital as being controlled by hospital administrators with the objective of maximizing output. Pauly and Redisch change the assumption of administrator control by assuming that hospitals are controlled de facto by their medical staffs. These staffs are assumed to have the objective of maximizing their joint incomes. Equilibrium is achieved in the
market due to an assumption that there is an optimal number of physicians on the medical staff of each hospital. When the number of physicians exceeds the optimal level, new hospitals will be formed. In such a model, an equilibrium in the number of hospitals is achieved when the supply price has been bid up to the point where marginal and average physicians' incomes and the marginal supply price all are equated (Pauly and Redisch, 1973; Pauly, 1974; Pauly, 1980).

Watts took an intermediate position arguing that the extent of medical staff control of hospital administration depends upon the competitive environment for physicians. Where competition exists for physicians, hospitals will operate as physician cooperatives. On the other hand, she argues that hospitals with monopsony power over the "purchase" of physicians will maximize administrator's utility over such factors as income and prestige (Watts, 1976). Harris expands on this dichotomy within the hospital arguing that the hospital is made up of two internal organizations: the administrative staff that manages the provision of diagnostic, therapeutic, and hotel services, and the medical staff that admits and treats patients. Harris refers to this as a command and control system whereby the medical staff commands hospital services for the patients and the administration attempts to control (administer) the availability of these services with little actual control over who is admitted and how many units of the services will be demanded (Harris, 1977).

Finally, Custer argues that the relationship between the hospital and its medical staff is more complex than the notion of a free workshop implies. Custer assumes that under zero co-insurance, the purchase of a unit of health care will provide a certain quantity of health, the latter being a function of the quality of the physician and the hospital at which the physician practices. Physicians, according to Custer, choose a set of office and hospital characteristics that influences the prices physicians can charge. Unlike in the previous studies, Custer assumed there are costs to affiliating with hospital, e.g., physician time on committees. The optimal choice set is one that maximizes physician income subject to the recognized cost of operating the office. Likewise, Custer assumes that hospitals maximize profits by choosing an affiliation fee/attribute combination. The model predicts that competition among hospitals lowers the number of physicians a hospital can attract at any affiliation fee/attribute combination, i.e., that hospital competition reduces the cost to physicians of hospital characteristics. This prediction, which was supported by empirical analysis, is identical to the postulates of earlier studies that hospitals compete for patients indirectly by enhancing quality (Custer, 1984).

In summary, the behavior of hospitals has not focused on price competition, the usual element that enforces efficiency in competitive markets. Even though price competition was an element in many of the models described, quality competition predominated. The result has been outcomes characterized by excessive quality and capital expenditures, and unnecessary duplication of services. Therefore, it is necessary to
direct our attention to why competition does not enforce efficiency, i.e., to other causes of market failure.

B. Risk Sharing Mechanisms as a Cause of Market Failure

In 1980, over 90 percent of the expenditures for hospital services were paid by public and private insurers. Approximately two-thirds of the expenditures for all types of health services are paid by the third parties (Health Care Financing Review, 1981). The existence of insurance in hospital markets is not the problem. Insurance, by reducing the financial risk to patients and providers, greatly increases economic welfare (Arrow, 1963). Market failure has occurred as a result of the characteristics health insurance has taken in the U.S.

Private insurers were the first to emerge on a broad basis. Blue Cross plans emerged during the depression years to cover hospital costs, along with Blue Shield plans to cover physician costs. These plans still account for almost 50 percent of the private health insurance written nationwide. The plans have certain characteristics that relate to issues of market failure. The Blue Cross plans, as part of their enabling legislation, were to operate as not-for-profit insurers in return for favorable tax treatment in many states (Reed, 1947; Frech, 1974). From the perspective of the insurance markets, lower costs derived from the tax advantages permitted Blue Cross to obtain a dominant position in many insurance markets (Frech, 1979). The not-for-profit status of the plans, when coupled with the tax advantages, led to a surplus which could not be distributed as dividends. The actual distribution of this surplus has resulted in a property rights struggle among the providers, plan administrators, and the insureds. If some of the benefits from the tax advantages were passed on to consumers in forms such as lower premiums for equal coverage or greater coverage for the same premium cost as offered by the for-profit commercials, the latter would be substantially impeded from competing effectively, which does not appear to be the case. There remains competition between plan administrators and providers for any remaining rents that result from the tax advantages. Substantial evidence has been found that supports the hypothesis that these groups share the rents. Administrative slack and greater insurance loading charges (Frech and Ginsberg, 1978; Frech, 1976; Vogel, 1977, Eisenstadt and Kennedy, 1981) and higher provider charges were found to be related to the levels of the tax advantages (Arnould and Eisenstadt, 1981; Arnould and DeBrock, 1984a and 1985; Sloan, 1980).

More critical to hospital markets is the tax treatment of health insurance premiums paid by individuals or for individuals by their employers. Along these lines, it is also important to consider the methods by which hospitals are reimbursed by the insurers. Insurance coverage has both a substitution and an income effect. Insurance effectively reduces the marginal cost of health care by an amount proportional to the level of coverage. Therefore, all other things equal, consumers demand more units of health care. The transfer of
Income from the "wellness" state of nature to the "sickness" state generates an income effect by reducing income in the former state by the amount of the premiums. Medicare and Medicaid, publicly funded health care insurance for the elderly and poor, respectively, have generated similar increases in demand for health care services.

Tax treatment of insurance premiums in the U.S. has eliminated most of the income effect (Gensheifer, 1985). Premiums paid by individuals and corporations are treated as a deductible expense. For many individuals, the premiums are a part of an employee benefit package that is not considered a part of taxable income. Therefore, the impact of any increase in premiums in a coverage period, regardless of consumption of health care in the previous period, is greatly diminished. The result is a significant amount of moral hazard in the form of both excessive insurance coverage and excessive consumption of health care services. Financial risks are shifted from the consumer of health care services to the third party insurers (Feldstein, 1974; Arrow, 1963; Zeckhauser, 1970)

Traditional reimbursement mechanisms developed by public and private insurers permitted similar shifting of risk from providers to the third party payers. Providers were (and in many cases still are) paid for health care on a retrospective fee-for-service basis with fees set by a variety of cost-plus pricing mechanisms. This system has the obvious problem of generating more revenues to providers if more services are provided, which can be done without increasing the financial risk to the provider. Therefore, providers had an incentive to expand output, generating additional moral hazard.

The problem of moral hazard is further complicated by two information problems. First, consumer knowledge of health care production is generally limited and costly to obtain. The primary care physician generally fulfills two roles for the patient: diagnosis and production. The physician functions as an agent for the patient prescribing the treatment necessary and then supplying the patient with his own and other health care inputs (Evans, 1983). However, because insurance greatly reduces the marginal cost of health care services, the second problem emerges—neither the consumer nor the primary care physician have an incentive to search for low cost providers (Feldstein, 1974; Frech, 1984).

Attempts to document cost and expenditure increases attributable to moral hazard are numerous (Feldstein, 1973). Pauly (1969) estimated the welfare cost in 1963 to be $450 million. Feldstein (1973), adding to Pauly’s analysis insurance induced price and quality changes, estimated the welfare costs for 1969 at about $2.4 billion. This represents approximately 20 percent of total hospital expenditures.

More recently, the Rand group, in a controlled experiment, found that 25 and 50 percent copayments reduced average total ambulatory and hospital expenditures by 19 and 30 percent, respectively (Newhouse, 1978, Phelps, 1982, Phelps and Newhouse, 1972). Phelps (1982) further
estimates that taxing half of private insurance premiums would reduce health care expenditures by $12 to $13 billion, over $7 billion of which is spent on hospital services.

The evidence of market failure is abundant and the cost of these market imperfections is large. However, unlike many industries analyzed by those who study industrial organization, market structure has little to do with market failure. Instead, market failure in the hospital markets is the result of behavioral characteristics and information problems.

IV. ATTEMPTS TO REGULATE EFFICIENCY

Attempts to regulate hospital efficiency pertinent to the theme of this paper began with the Hospital Survey and Construction Act of 1946 (Hill-Burton). The purpose of this bill was to provide grants and subsidies to the states for new hospital construction and hospital renovation. Growth in hospital capacity was thought necessary to meet increasing demands for hospital services resulting from population growth, technological change in the treatment of health problems, and public and private insurance. Federal programs to expand medical education and, thus, the supply of physicians began in 1963. Both programs were very successful in achieving their intended purposes.

In the decades of the 1960's and 1970's, regulation turned to more direct control of hospital costs. The Comprehensive Health Planning Act of 1966 (CHP) was intended to fund state agencies to develop regional facilities plans. CHP agencies were given little control over hospital supply decisions except through their limited input on Hill-Burton or Public Health Service funds. The National Health Planning and Resource Development Act of 1974 (HPRDA) established a number of health system agencies (HSA) designed to halt the escalating costs of health care by regulating hospital construction. The most significant feature of the HPRDA is the requirement that states establish certificate of need (CON) programs for expansion of bed space and other types of service where the capital investment needed to offer those services is greater than a threshold level. These regulations have proven to be quite important in the final outcomes of many hospital market configurations.

Two effects can be expected from CON regulation. First, CON establishes a barrier to entry, protecting the status quo. Thus, if existing hospitals are inefficient, they will remain inefficient (Sloan and Steinwald, 1980). Second, entry regulation into an industry that does not have natural barriers to entry often leads to inefficient forms of competition (Douglas and Miller, 1974). The airline industry provides an example. Prohibited from price competition and protected from entry, the airlines turned to quality competition by increasing the frequency of flights, the size of aircraft, and in-flight amenities. The result was substantial excess capacity and higher costs per passenger. Analysis of the effects of CON on hospital costs provide results ranging from no evidence of reduced capital expenditures (Watts
and Updegraff, 1975), reduced growth of licensed acute care beds but increased growth of psychiatric, extended and special care beds (Bicknell and Walsh, 1975), to reduced growth of hospital beds but increased employees and auxiliary services per bed (Salkever and Bice, 1975).

Professional Standards Review Organizations (PSRO's) were established in 1972 to serve as watchdogs over Medicare performance. PSRO's reviewed Medicare patient bills to eliminate services that deviate from accepted norms, namely, that have greater social costs than benefits. However, these systems failed, largely because of the difficulty in developing generally accepted measures of social benefits and costs. The PSRO's then reverted to the use of industry norms. However, industry norms provided inefficient guidelines because the data used to establish the norms were taken from a period when providers had little or no incentives to operate efficiently (Zeckhauser and Zook, 1981; Sloan and Steinwald, 1980).

In the late 1970's and early 1980's numerous state governments in addition to the Federal government turned to various cost caps and the Medicare section 223 limits of the Social Security Act Amendments of 1972. These limits pertained to total reimbursements related to daily service charges. However, these limits were not inclusive of high technology ancillary services. The result was a pattern of cost shifting from covered services to other services and from Medicare and Medicaid patients to private pay patients (Meyer, 1983).

Since 1969, eight states have instituted mandatory hospital rate regulations. Rhode Island has a cooperative program, and four additional states have mandatory advisory hospital rate review programs (Joskow, 1981). These programs have followed a variety of regulatory schemes, usually with some level of budget review. Many aspects of rate regulation are similar to traditional utility regulation. The limited evidence that exists on the effectiveness of rate regulation is mixed. Joskow (1981), e.g., concludes that the rate of increase in total hospital expenditures and expenditures per patient day were lower for the regulated states than for a comparison group of unregulated states. This evidence is subject to criticism when analyzed on a state-by-state basis. Further, it suffers from a failure to take into account other possible distortions caused by the regulation.

There should be little surprise at the inadequacy of these regulatory attempts to improve the efficiency of hospital markets when one recalls the causes of market failure. Evidence points to the fact that the causes of market failure are behavioral and not structural. They result largely from a risk sharing system that provides improper incentives for efficient production and consumption of health services, and information imperfections. None of these problems were addressed by the various types of regulation summarized in this paper.
V. CHANGING THE MARKET INCENTIVES

One economic phenomena is clear from the literature reviewed to this point. Hospitals do respond to economic incentives. When third party payers eliminated the effects of price competition, hospitals responded with quality competition. Similarly, when able to shift financial risks to third party payers, hospitals responded by producing (and patients by consuming) excessive amounts of hospital services. Thus, economists can have some optimism that changes in the incentive structure will lead to more efficient market performance. Many changes in policies have been proposed. Most attempt to generate greater incentives for competitive behavior on the part of both consumers and providers. Some have been implemented. These changes will be reviewed here with a brief discussion of provider response.11

A variety of proposals have been made to eliminate the tax subsidy responsible for excessive insurance coverage. Many also are concerned with the low levels of copayments and deductibles. The most extensive plan, put forth by Enthoven (1978), would replace the current tax exclusion of health insurance premiums with a fixed tax credit. Aspects of this plan are contained in a number of bills currently before Congress. The intent of eliminating or placing a cap on the tax exclusion is to provide consumers with a financial incentive to seek efficient health care providers. Although the potential savings from such action are great, the policy has not been implemented. Thus, no empirical evidence exists to support these claims.

Actions have been taken in the public and private sector to shift more financial risk to producers and consumers in an effort to reduce expenditures on health care. There has been a significant increase in the level of deductibles and copayments required by the health care portion of many employee benefit plans. A recent survey of 1200 companies by Hewlett Associates documents the extent of this increase in deductibles and copayments. The number of companies including deductibles in their employee health plans increased to 52 percent in 1984 from 17 percent in 1982; only 50 percent of the companies provided full coverage of "room and board" costs in hospitals in 1984, a decline from 75 percent in 1983. Moreover, those providing full coverage for surgery declined to 27 percent in 1984 from 45 percent in 1979. (Cited in Davis, et al, 1985). Phelps reported that the interim results of the Rand Health Insurance Study provide evidence that a 25 percent copayment reduces utilization and expenditures by 20 percent (Phelps, 1982). If the results of the experiments analyzed by Phelps generalize to the broader population, substantial savings should be forthcoming from this change.

The actions on the part of payers to further utilize market incentives to increase efficiency has led to the rapid development of systems of prospectively determined rates for health care services. These have taken the form of prospective payment systems (PPS), preferred provider arrangements (PPA), and complete prepayment plans. Each form of prospective payment shifts some of the financial risk to
the provider and/or consumer. The effects of these systems have led, in turn, to increases in excess capacity and other pressures on health care providers to supply health care services efficiently.

The most prominent single PPS in the U.S. is the Medicare PPS enacted in 1982 that took effect in 1983. That system delineates units of health care services into 468 diagnostic related groups (DRG's), each of which represents an illness or health problem based on the patient's primary diagnosis, secondary diagnosis, age, and sex. Currently, costs of physician services, ancillary services, outpatient services, and capital costs are not included in the DRG's. Rates are determined by a weighted average of historic costs of the hospital in question, other hospitals in the region, and national hospital costs. The hospital is permitted to retain any surplus if costs are below DRG rates providing an incentive for efficient production of hospital services. The hope is that such a weighting scheme will reduce the interhospital variance in expenditures. Eventually, the hospital-specific and regional weights are to be reduced with more weight being placed on national averages.

The program is too short lived to provide substantial empirical evidence of increased efficiency. While the program will cause some short-term disruptions and burdens, it should provide incentives for more efficient behavior. However, it is clear that the DRG system has provided an improved internal mechanism for hospitals to identify their production functions. Three potential problems exist with the incentives of the system. First, as long as the prospective rates are based on hospital specific and regional hospital costs it is unclear whether it will provide incentives for hospitals to compete with each other for Medicare patients. However, if rates are set appropriately low, while still held above marginal costs, the incentive to compete for Medicare patients may result from hospital efforts to reduce excess capacity (and make some contribution to overhead). Selective price cutting beyond the official DRG levels is being used by hospitals to attract Medicare patients and alleviate some of their excess capacity. There are situations in which hospitals with low occupancy rates are advertising they will treat Medicare patients who meet certain minimum income levels with a complete waiver of copayments and deductibles. Medicare patients above these income levels receive various discounts on hospital and physician deductibles and copayments. Other hospitals are offering fixed maximum charges for all Medicare copayments and deductibles. No documentation is available of the number of health care providers offering these types of discounts; however, popular knowledge seems to be that they are not uncommon.

The second problem is that the PPS does not reduce the incentive to increase revenues by increasing hospital admissions. Finally, where secondary diagnoses are present, it may be possible for hospitals to increase revenues by switching primary and secondary diagnoses in some situations. In this case, the PPS could actually lead to more inefficiencies. No systematic evidence has been found on the extent of these problems with the actual PPS.
Karen Davis, et al (1985), provide a review of the evidence of the effects of Medicare PPS on hospital utilization. They compare pre and post 1983 hospital utilization rates and costs. However, the authors of that paper preface their results with a very important caveat. Specifically, hospital patient days for the elderly increased by an average of 3.4 percent per year from 1975 to 1983, but declined by 8.7 percent from 1983 to 1984 with most of the reduction in patient days resulting from reduced lengths of stay. However, over the same period, patient days for the non-elderly increased by 0.1 percent per year and declined by 7.4 percent respectively, the latter decline being more closely the result of reduced admissions.

Hospital cost increases follow a similar pattern. From 1975 to 1983, hospital costs increased at an average annual rate of 15.7 percent compared to an increase of only 5.4 percent between 1983 and 1984. (Davis, et al, 1985). Cost per patient day has increased by approximately the same amount throughout the entire period. Therefore, the reduction in the rate of increase in hospital costs was the result of fewer patient days, not reduced cost per patient day.(Davis, et al, 1985).

As can be seen from these data, a substantial change occurred in hospital utilization and cost between 1983 and 1984. The reason for the caveat mentioned earlier regarding how much of the credit for this change can be attributed to the Medicare PPS is that changes occurred for non-Medicare as well as Medicare patients. In addition, the Medicare PPS was phased in between October 1983 and September 1984 and certain hospitals in certain states were exempt. Thus, only 25 percent of Medicare payments to hospitals, or less than 5 percent of total hospital reimbursements were based on the PPS. Finally, Davis, et al (1985), indicate that many of the downward trends of hospital utilization commenced prior to the inception of the PPS. There is some evidence that hospitals initiated programs to reduce utilization prior to October, 1983, in anticipation of the Medicare PPS. Needless to say, the Medicare PPS has resulted in the appropriate incentives for increased efficiency in the utilization of health care services. Our ability to separate the quantitative influence of this system from other factors must await the availability of more complete data.

A serious issue that remains to be resolved with the Medicare PPS is the treatment of capital costs. Currently, capital costs are treated as a pass through for which hospitals are reimbursed for actual capital costs based on the proportion of patient days that are attributed to Medicare patients. The U.S. Congress has set October 1, 1986, as a deadline for resolving this issue. A dispute between the Department of Health and Human Services and the Office of Management and Budget involves the appropriate levels of "add-ons" to DRG payments to cover the capital costs and the appropriate treatment of fixed capital in an industry plagued with excess bed capacity. Senators Durenburger and Quayle have introduced a bill that calls for a repeal of state health planning and planning program review requirements and a flat percentage
add on to DRG's. The add on would be phased in over six years with a hospital weight that declines to zero by the sixth year. Another bill introduced by Senators Kennedy and Gephardt would provide a DRG-specific add on adjusted for the capital intensity of each DRG, for regional differences in construction costs, and with an exemption process for hospitals for which capital payments were not sufficient to meet principal, interest, and lease payments on projects undertaken prior to year end 1985.

Clearly, the treatment of capital costs will have a critical impact on the effectiveness of the PPS as a mechanism to provide efficient capital decisions. If the capital cost payment to hospitals is set too low, it penalizes, possibly to the extent of exclusion from the market, hospitals with high capital costs resulting from recent renovation projects designed to increase efficiency and reduce costs while protecting hospitals with older and possibly less efficient physical plants. An example might be a hospital that has expanded outpatient surgery facilities. Similarly, unreasonably low payments could discourage other hospitals with older plants, and, therefore, low capital costs, from undertaking capital improvement projects that would increase efficiency, even if these hospitals have high occupancy levels.

If the payments for capital costs are set too high, an Averch-Johnson effect might result (Averch and Johnson, 1962). The Averch-Johnson model, developed to analyze the nature of capital formation in regulated public utilities, predicts that where the allowed rate of return on capital is greater than the cost of capital, firms have an incentive to overcapitalize. Thus, excessive reimbursements for capital could lead to inefficient expansion of capacity (in the case of hospitals, a lack of contraction) in an industry already plagued by excess capacity.

The determination of the appropriate level of reimbursement for capital costs has been avoided by public and private payers through the use of selective direct contracting for services between the payer and specific providers. These arrangements, referred to as preferred provider arrangements (PPA's), involve direct negotiation of rates between the preferred provider organization (PPO) and the payer (Trauner, 1983). Hospitals are generally reimbursed on a discounted basis of usual charges, with negotiated per diem rates or case rates becoming more common. Physicians may be paid at a discounted rate from usual charges or on the basis of a negotiated fee schedule. Discounts range from 7 to 15 percent for hospitals and from 5 to 30 percent for physicians (Gabel and Ermann, 1985).

PPO's have a number of advantages in addition to the determination of the level of coverage of capital costs. They provide appropriate incentives on prices and inputs and contractual controls on the production of priced services. However, the units of priced services may not be effectively controlled. For example, if hospitalization is priced on a per diem rate, the incentive exists for the PPA to efficiently produce a per diem of hospital care, but no incentive exists
to reduce admissions or lengths of stay of patients admitted. Therefore, to be effective, PPA's must contain stringent conditions for utilization review, the cost of which must be weighed against the PPA discounts.

Finally, PPA's provide a mechanism through which health care providers can offer selective price discounts to attempt to increase the level of capacity utilization. This and other forms of price discounting are not uncommon in industries with excess fixed capital. The U.S. railroads came under regulation in the late 19th century at least partially as a result of wide price fluctuations that occurred during periods of reduced demand. Trucking rates came under regulation during the 1930's for similar reasons. However, trucking firms do not have high fixed costs. Therefore, it is now common belief that regulation to eliminate price fluctuations was ill-conceived and led to welfare losses. When airline rates were deregulated in 1978, one immediate result was selective price discounts to increase load factors. These selective discounts are carefully structured to provide reduced fares and increased utilization among consumers with high elasticity of demand. The discounts are limited to those with substantial planning time usually requiring ticket purchase to be made two to four weeks in advance of travel. Many require the traveler to stay at the destination over Saturday night eliminating their accessibility to business travelers. Similarly, PPA's are offered on a selective basis to payers who have some monopsony power as a result of their ability to guarantee the health care provider a large volume of patients. Therefore, hospitals, which have high fixed costs relative to variable costs, are actively offering discounts in an effort to cover overhead through increased volume.

Probably the most widely publicized case of selective contracting is the California Medicaid system, Medi-Cal. The Medi-Cal program was authorized to accept bids from hospitals for the treatment of Medi-Cal patients in 1982. California law was changed to provide for similar contracting by private insurers and for contracting for physicians services by public and private insurers in 1983. Selective contracting was viewed by the Medi-Cal program as a substitute to HMO contracting needed because of the shortage of available HMO's. Also, as a substitute, the selective contracting provided a mechanism that was reliant upon competitive incentives and the market system.

Four important program policies were established by Medi-Cal that facilitated the contracting: every accredited short-term acute care hospital was invited to negotiate contracts; hospitals and Medi-Cal worked from but were not limited to a "model" contract; payment was on a fixed, per diem basis with only few exceptions; and there were direct negotiations between one state negotiator and the hospital's representative (Johns, 1985). By the end of 1984, 68 percent of the state's eligible hospitals had participated in contract negotiations and 70 percent of those had won contracts. Ninety-two percent of the state's historical expenditures for in-patient Medi-Cal services in areas in which contracts existed were covered by negotiated contracts.
(Johns, 1985).

The effects of selective contracting by Medi-Cal are difficult to assess at this time. Total Medi-Cal savings in the 1983-84 fiscal year were $700 million. The change in payment for in-patient care was estimated to be responsible for $184 million, or 26 percent of this total (Johns, 1985). However, caveats like those discussed above concerning the Medicare PPS are in order because ceteris paribus conditions did not hold for a number of other factors. Other states see enough promise in this form of PPO to have initiated similar programs.

Other PPO's have grown rapidly. Gabel and Ermann (1985) report on surveys that reveal 33 PPO's in 1982 and 143 in 1984. Gabel (1985) reported that PPO enrollments rose four-fold between December, 1984 and May, 1985 to 5.8 million participants up from 1.3 million participants. Hospital and physician groups sponsored nearly one-half of the PPO's with insurance carriers running closely behind (Gabel and Ermann, 1985). The major weakness in PPO's reported by Gabel and Ermann (1985) was the general lack of a rigorous utilization review. Thus, they predicted continued growth in PPO's with increased emphasis by those organizations on utilization review.

Health maintenance organizations (HMO's), the final form of reimbursement mechanism to be discussed, have shown rapid growth as a form of prepaid health care. HMO's provide one contracted source for all health care needs of the consumer group for a specified premium. With revenues fixed in advance, HMO profits are maximized by minimizing plan costs. By providing comprehensive coverage of the health care needs of the group, the problems of defining the limits of coverage are eliminated. Financial risks are passed on to the providers if they are paid a per capita based on the number of enrollees in the HMO. Therefore, moral hazard generated by the incentives of providers to prescribe excessive units of health services is eliminated.

HMO's should be particularly effective in promoting more efficient delivery of health care services as pressures on consumers to search for efficient providers are increased. In 1970, there were 33 HMO's in the U.S. serving approximately three million enrollees. The number of HMO's grew to 290 serving 13.6 million enrollees by 1983, and to 337 HMO's serving 16.7 million enrollees by December, 1984 (Interstudy, 1984). Current increases in annual enrollments average 20 percent. The market penetration of HMO's exceeds 25 percent in many cities. Various cost savings of HMO's have been documented. Most of the documented cost savings are the result of lower hospital admission rates of HMO enrollees than fee-for-service patients. Luft (1981) and Manning, et al (1984) found hospital admission rates for enrollees in a number of HMO plans frequently were as much as 40 percent lower than for fee-for-service patients. Arnould, DeBrook, and Pollard (1984b) found that HMO's produce certain health services more efficiently. One problem with these studies may be selection bias on the part of enrollees. The cost savings found by Luft and others may be diminished as more is discovered about the health and risk characteristics of those
individuals who choose HMO's over traditional third party reimbursement plans. However, Manning, et al (1984) found evidence that the Luft results hold up under control for these factors.

The increased risk borne by providers under prepaid plans is generating structural changes among the traditional provider roles. One notable change is that providers, to minimize total health care costs, are increasing their emphasis on "wellness" programs. More significantly, the structure of providers is changing in response to competitive pressures. These structural changes involve simultaneous moves in different directions all of which are efforts to provide services more efficiently, respond to consumer tastes and utilities, and maintain profitable levels of utilization.

Retrospective fee-for-service reimbursement provided no incentive for providers to promote outpatient surgery even when that was more efficient and as efficacious as inpatient surgery. Hospitals reimbursed on a capitation basis or DRG rate have incentives to require outpatient surgery when possible. This has led to an expansion of outpatient surgery facilities by many hospitals. Along with this development has been an unbundling of hospital services resulting from the growth of free standing surgicenters that offer one day treatment of many types of outpatient surgery. There were 330 surgicenters in 1984 providing over 500,000 operations, compared to 212 centers in 1982 (Davis, et al, 1985). There has been a lower rate of growth in specialty ob-gyn centers.

Similarly, on the physician's side, the number of emergicenters, physicians offices that treat outpatients usually for minor problems and are heavily oriented to service by maintaining long hours of operation and convenient locations, have grown to 2,500 from 260 in 1981 (Davis, et al, 1985). These facilities provide convenience to the patient and permit the physician to capture revenues that otherwise would have gone to the hospital.

Hospitals have responded to this disintegration, or unbundling, of hospital services by expanding their outpatient surgical facilities and/or operating freestanding ambulatory care units. As reported earlier, the American Hospital Association survey of Multihospital systems found that 55 percent of those systems operate ambulatory care facilities. Davis, et al (1985), reports that multihospital systems operate 1,322 freestanding ambulatory care facilities of 1984.

There is some concern that PPO's in certain circumstances could be organized to be anticompetitive rather than procompetitive (Feldstein, 1985). For example, assume that providers, such as hospitals, in the same market area form a PPO. Further, assume that these providers make up a substantial share of the market, are permitted to collude with each other in the prices to be charged by the PPO, are not permitted to offer lower prices to alternative PPOs, and that the PPO has an exclusive arrangement with the enrollees. Under these very rigid circumstances it may be possible for the providers to use the PPO to fix prices at
artificially high levels or to maintain prices by thwarting competition. However, more extensive information about market conditions is necessary to turn this possibility into reality. The Antitrust Division of the U.S. Department of Justice and the Federal Trade Commission have indicated that they had no intent to challenge the PPO's that requested an antitrust review in all but one case. Thus, the preponderance of PPO's are structured to increase competition in health care markets.

Another structural change occurring as a result of competition in health care markets, a type of horizontal integration, is the formation of multihospital systems either through mergers and acquisitions or looser affiliations involving shared services. The reasons for the extent of this activity was discussed earlier in this paper. Many health care planners predict that further horizontal expansion of multihospital systems will diminish in the future. However, the expansion of physician group practices facilitated by single and multihospital systems and in other cases independent from those systems, and various types of vertical integration are occurring at a very rapid pace.

In 1980, 26.2 percent of the physicians in the U.S. were in 10,762 group practices. The average group practice consisted of eight physicians. By 1984, 29.3 percent of the physicians were in group practices with the average group size increasing to nine physicians (Arnould, Pollard, and Van Vorst, 1985). Clearly, most physician service markets remain highly atomistically structured with large numbers of independent physicians.

Many of the older group practices formed under the umbrella of a prepaid health plan or as tertiary care centers. The Kaiser-Permanente Group in California and Group Health of Washington, D.C., are examples of the former. Mayo Clinic, Virginia-Mason, and the Carle Clinic are examples of the latter. In recent years, many of the groups in the latter category have vertically integrated into the insurance market by offering HMO's.

Group practices may result in economies of scale in the provision of physicians' services; however, this is an area relatively untapped by researchers. More probable, if the group practice is integrated into an HMO, economies of scale in the insurance function may dictate the size of the patient base and thus the necessary size of the group practice. Group practices may also have lower contracting and other information costs and superior internal mechanisms to control quality and enforce utilization than IPA's and other contractual arrangements.

Vertical integration also is occurring to provide an adequate patient base and to control under one organizational unit all health care costs and, therefore, all revenues. Hospitals faced with declining occupancy rates are forming physician group practices, as predicted by the hospital models discussed earlier in this paper, to attract larger patient bases. Regional and tertiary care group practices are diversifying into new geographic markets to expand their referral basis.
Kaiser-Permenente has physician staffs in 12 states and hospitals in 22 cities. Mayo Clinic has announced plans to develop satellite facilities to provide increased referrals to the Rochester campus. More recently, that clinic announced plans to acquire hospitals, the purpose being to gain more control over hospital utilization, costs, and revenues.

The smaller hospital and group practice will can be used to provide anecdotal evidence of some of the types of integration occurring and the strategic reasons for each type of integration. The for-profit clinic is a tertiary care physician group practice made up of 135 physicians in 1981. The hospital was a not-for-profit 300 bed hospital at that time. Approximately 50 percent of the patient base of the organizations came from within a 50 mile radius of the main campus. A strategic analysis of the market revealed that (1) there was a large surplus of hospital beds in the market area and (2) as the supply of physician specialists expanded the risk of relying on referrals to the main campus from independent primary care physicians was increasing. Thus, a goal was established to expand the size of the self-contained patient base.

The first action of the organization was to establish a number of primary care physician satellites within a 50- to 60-mile radius of the main campus staffed by employed physicians who would refer patients needing tertiary care to the main campus. Six such satellites have been established with more on the planning board.

The next action was to develop a wholly owned HMO to provide a competitive alternative to other insurance plans and to "lock patients in" to services offered by the organization. As a result of the competitive nature of health care (smaller group practice and independent physicians offered two alternative HMO choices to consumers) and the general acceptance of the group practice mode of delivery in the community, enrollment in HMO's rapidly grew to account for approximately 25 percent of health care services in the market.

With the risks associated with moral hazard now shifted to the provider and with extensive price competition between the HMO's in this market, further actions were taken to control utilization and costs and to enhance services. Some of these actions also were promoted as a strategic response to the Medicare PPS. The organization purchased a skilled nursing facility and developed a home care service. The purpose of both actions was to provide an alternative to the hospital at a reduced cost for patients needing a lower level of care than available in the acute care hospital. The services also were established to provide direct referral of patients not in the HMO to other services provided by the organization—namely, physician and hospital care. Finally, these services were added to enhance total revenues. A major remodeling program was undertaken to enhance outpatient surgery facilities. Outpatient surgery now accounts for over 45 percent of all surgeries conducted by the organization. Retail pharmacies and durable medical equipment sales were established to provide a competitive alternative for HMO and other patients as well as to provide an additional profit center. The organization also contract manages other
health care facilities, negotiates PPA's with hospitals in locations
closer to the satellites than is the main campus hospital, is planning a
retirement center, and has developed a wholly owned malpractice
insurance company. Throughout this same period, the organization has
moved from a low-keyed indirect marketing effort to an extensive
multimedia marketing effort with an extensive marketing department.

Many more recognizable national health care providers, such as
Hospital Corporation of America, Humana, Kaiser-Permanente, could have
been used to provide this anecdotal evidence. The examples would have
been similar but on differing scales of magnitude. The selected
organization was chosen to provide evidence that structural change among
health providers is not limited to a small number of national
proprietary firms.

The purpose of each of these types of integration is to reduce
costs and enhance the patient base and revenues. Clearly, the incentive
for these changes has been brought about by changes in reimbursement
methods that place an increased risk of moral hazard on the producers
and consumers of health care services. Simultaneously, these changes
are placing more reliance of competitive market forces to contain health
car costs. It is not obvious that complete vertical integration, i.e.,
the merger of one type of provider or service firm into another, will
always increase efficiency more than obtainable through alternative
contracting services. However, it is clear that these structural
changes are altering traditional choice patterns between patients and
providers and among various provider types. Empirical evidence of the
full extent and effects of integration is very limited eliminating the
opportunity to pass final judgement at this time. These and other
related questions about the effects of changes in health care markets on
market competition and efficiency are important topics for industrial
organization research.

VI. SUMMARY AND CONCLUSIONS

This paper was written with the intent of providing insight into
problems in the hospital sector by relying heavily on published research
viewed from the perspective of industrial organization. It is
abundantly clear from this review that the problems in the hospital
sector of concern to industrial organization economists are similar, if
not identical, to those of concern to economists in other subspecialties
such as labor, public finance, and applied microeconomic theory
Therefore, it would be presumptuous and egotistical to assert that the
relatively recent (with a few exceptions) interest of economists in the
field of industrial organization will bring dramatic new insights to
this area.

One small, but very important, contribution to be made by
economists specializing in industrial organization is to stress the fact
that hospital services are bought and sold in markets. The important
questions concern how efficiently the market is functioning, i.e., are
resources allocated efficiently, is market behavior competitive, and is the result consistent with efficient market performance. In other words, research questions should address the causes and potential cures of market failure.

Increased emphasis on government policies and private reimbursement mechanisms to provide more reliance on appropriate market incentives by altering the behavior of providers is a recent phenomena that has emerged almost simultaneously with increased state regulation of hospital rates. Thus, it is difficult to empirically separate influences on market performance of "deregulation" from those of "reregulation." Changes in the tax system appear to have the most promise of bringing about broad changes in the incentives of consumers. However, short of those changes the private sector purchasers of health care services already have made significant movements in the direction of efficient management of the health benefits package provided to employees, a movement placing pressures on providers to seek more efficient delivery systems.

Future research should be directed to providing estimates of the efficiency and effectiveness of new health care delivery configurations. This will involve filling in the gaps that exist in hospital cost analyses, projecting appropriate product mixes for hospitals, analyzing the role of proprietary hospitals, and the impact of those hospitals on market performance. Can not-for-profit institutions function like efficient for-profit institutions? Do economies of scale justify multi-hospital systems? What effect will increased emphasis on competition have on hospital-doctor relations and the objectives of hospitals? What effect will increased price competition have on technological change? Can a more price competitive health care system support medical education? Underlying many of these issues is the very basic question that asks "what is the nature of a sustainable price structure for hospital services?" Will that price structure be supported by adequate competition or are cost conditions such that monopolies will exist? Information imperfections have been a significant cause of market failure. Numerous issues concerning the types and effects of information on hospital markets remain unresolved (Pauly and Longwell, 1982). Also, if direct regulation is replaced by reliance on the market, antitrust laws increase in importance as regulators of certain types of market behavior (Frech, 1984). Therefore, numerous research questions arise surrounding the appropriate application of these laws to correct problems of market behavior.

Clearly, the problems and changes occurring in the health care markets provide an almost endless list of theoretical and empirical questions for the future.
<table>
<thead>
<tr>
<th>Year</th>
<th>Total Hospitals</th>
<th>Total Change</th>
<th>Total Community Hospitals</th>
<th>Community Change</th>
<th>For-Profit Community Hospitals</th>
<th>Percent Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>7,156</td>
<td>*</td>
<td>5,875</td>
<td>*</td>
<td>775</td>
<td>*</td>
</tr>
<tr>
<td>1976</td>
<td>7,082</td>
<td>-1.0</td>
<td>5,857</td>
<td>-0.3</td>
<td>752</td>
<td>-3.0</td>
</tr>
<tr>
<td>1977</td>
<td>7,099</td>
<td>0.2</td>
<td>5,881</td>
<td>0.4</td>
<td>751</td>
<td>0.1</td>
</tr>
<tr>
<td>1978</td>
<td>7,015</td>
<td>-1.2</td>
<td>5,851</td>
<td>-0.5</td>
<td>732</td>
<td>-2.5</td>
</tr>
<tr>
<td>1979</td>
<td>6,988</td>
<td>-0.4</td>
<td>5,842</td>
<td>-0.2</td>
<td>727</td>
<td>-0.7</td>
</tr>
<tr>
<td>1980</td>
<td>6,965</td>
<td>-0.3</td>
<td>5,830</td>
<td>-0.2</td>
<td>730</td>
<td>0.4</td>
</tr>
<tr>
<td>1981</td>
<td>6,933</td>
<td>-0.5</td>
<td>5,813</td>
<td>-0.3</td>
<td>729</td>
<td>-0.1</td>
</tr>
<tr>
<td>1982</td>
<td>6,915</td>
<td>-0.3</td>
<td>5,801</td>
<td>-0.2</td>
<td>748</td>
<td>2.6</td>
</tr>
<tr>
<td>1983</td>
<td>6,888</td>
<td>-0.4</td>
<td>5,783</td>
<td>-0.3</td>
<td>757</td>
<td>1.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Hospitals</th>
<th>Beds (000's)</th>
<th>Beds/Hospital</th>
<th>Admissions (000's)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>5,875</td>
<td>942</td>
<td>160.3</td>
<td>33,435</td>
</tr>
<tr>
<td>1976</td>
<td>5,857</td>
<td>956</td>
<td>163.2</td>
<td>33,979</td>
</tr>
<tr>
<td>1977</td>
<td>5,881</td>
<td>969</td>
<td>164.8</td>
<td>34,273</td>
</tr>
<tr>
<td>1978</td>
<td>5,851</td>
<td>975</td>
<td>166.6</td>
<td>34,506</td>
</tr>
<tr>
<td>1979</td>
<td>5,842</td>
<td>984</td>
<td>168.4</td>
<td>35,099</td>
</tr>
<tr>
<td>1980</td>
<td>5,830</td>
<td>988</td>
<td>169.4</td>
<td>36,143</td>
</tr>
<tr>
<td>1981</td>
<td>5,813</td>
<td>1,003</td>
<td>172.5</td>
<td>36,438</td>
</tr>
<tr>
<td>1982</td>
<td>5,801</td>
<td>1,012</td>
<td>174.5</td>
<td>36,379</td>
</tr>
<tr>
<td>1983</td>
<td>5,783</td>
<td>1,018</td>
<td>176.0</td>
<td>36,152</td>
</tr>
</tbody>
</table>

6-24 Beds: 224, 4.3, 19.2, 129.7
25-49: 986, 36.5, 37.0, 1,156.0
50-99: 1,414, 102.9, 72.8, 3,392.4
100-199: 1,382, 195.3, 141.3, 6,942.0
200-299: 726, 175.3, 241.5, 6,516.4
300-399: 438, 148.6, 339.3, 5,569.7
400-499: 274, 121.2, 442.3, 4,378.3
500+: 339, 234.4, 691.4, 8,067.3

Source: *Hospital Statistics*, American Hospital Association, 1984
This paper ultimately turns to a discussion of the nature and causes of inefficiency in the health care sector. These problems have been the subject of an enormous number of research papers. In an attempt to represent this research both fairly and within the confines of a short article, we could not describe all of the research. Therefore, we apologize to those who may feel their research deserved recognition.

Unless otherwise cited, all data presented in this section are referenced from: Hospital Statistics, (Chicago: American Hospital Association, 1984).

Of these specialized hospitals, 85.1 percent have a capacity in excess of 400 beds.

Summary measures in this section were derived from data provided by H. E. Frech, University of California, Santa Barbara, using 1975 data.

Joskow used the Herfindahl index to describe market structure in an earlier paper (Joskow, 1980). This index is the sum of \( S_i^2 \) for the top \( n \) firms. If \( S_i \) represents market share of firm \( i \), the \( n \)-firm concentration ratio is the sum of the \( S_i \) for the top \( n \) firms. The Herfindahl index provides a more encompassing measure of market structure than the 2 or 4 firm concentration ratio (Schmalensee, 1977). However, these data were not available to the authors.

See Cowing et al (1983a) for a summary of these studies.

Information in the following paragraphs also comes from: Hospital Statistics, (Chicago: American Hospital Association, 1984).

Obviously, the validity of the cost functions discussed in the previous section are circumspect if this has not been taken into account in the estimation procedure (see Feldstein, 1974, and Cowing et al, 1983a).

This is determined by a number of exogenous and endogenous conditions.

The Blues were required to use community rating systems, in many areas leaving room for the commercial plans to offer better rates to certain groups within these areas.

A more complete discussion is found in Langwell and Moore, 1982; Longwell et al, 1982; and Pauly, 1982.

BIBLIOGRAPHY


Cowing, T and A Holtman, "Multiproduct Short-run Hospital Cost


Interstudy, National HMO Census, 1985, Minneapolis, MN.


Trauner, J., Preferred Provider Organizations: The California Experiment, (San Francisco: University of California, August, 1983).


