PHYSICIAN HOSPITAL INTEGRATION AND COST EFFICIENCY IN U.S. PRIVATE HOSPITALS IN 1997

TAN BOON SENG
[B.Sc. (Pharmacy) 90; M.B.A. 95 NUS]

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
THE NATIONAL UNIVERSITY OF SINGAPORE
2004

ACKNOWLEDGEMENT

I thank Professor Lars Andrea Stole of the University of Chicago for inspiring me to think about the budget breaking mechanism in explaining physician behavior in the integrated hospital. I am also extremely grateful to Dr. Clint Cummins from TSP International for patiently guiding me through the econometric programming.

I wish to thank Assoc. Professors Phua Kai Hong, Toh Mun Heng and Rachel Davis who served in my dissertation committee and provided helpful comments in the preparation of this dissertation. I appreciate the financial support form my supervisor Professor Lim Chin through his research grant (Grant Number R-313-000-047-112). I would not have completed this dissertation without the generosity of Dr. Darren Carters of Info-X Inc. in sponsoring the ICD9CM-CPT4 crosswalk file for research purpose. I am extremely grateful for the thoughtful critiques by three anonymous examiners who substantially improve the quality of this piece of work.

Numerous people were generous with their time and assistance. Assoc. Professors Kulwant Singh and Go Mei Ling were my academic referees for my application to the program. Finally, I wish to thank my wife Siew Fong for her sacrifices and support in the years that I was doing this program. I dedicate this dissertation to her.

SUMMARY

Our research question is: How does physician hospital integration affect the quality-adjusted cost efficiency in U.S. hospitals in 1997? We view the physician and the hospital manager as a team of agents. Technologically, the physician resides within the firm because he allocates resources in the production of medical services. Production uncertainty in the sense of Arrow (1963) implies variable quality in medical care. Legally, a physician can be an employee in a fully integrated organization (FIO), a partner in a network, or an independent professional in a segregate hospital. The hospital owner (principal) administers a salary cum bonus scheme, which Holmstrom (1982) defines as a budget breaking scheme, for the production team in the FIO. Holmstrom shows that such a scheme removes moral hazard in team agency (i.e. team members shirk when their effort cannot be observed) and leads to Pareto efficiency (which we proxy with quality-adjusted cost efficiency). Eswaran and Kotwal (1984) argue that a self-interested principal faces a moral hazard problem herself and has the incentive to prevent the team achieving Pareto efficiency. Our result shows empirical evidence for this argument in U.S. hospitals: nonprofit FIOs are more cost efficient than nonprofit network or nonprofit segregate hospitals in our sample. However, the for-profit counterparts have similar (quality-adjusted) cost efficiency.

The principal can monitor the agents if she cannot administer a budget breaking incentive scheme in network and segregate hospitals. When the principal is the residual claimant, monitoring is incentive compatible (Alchian and Demsetz, 1972). The property rights theory predicts that for-profit hospitals are more cost efficient than nonprofit ones because the former have well defined residual claimants. For network and segregate hospitals, we find that for-profit entities are more cost efficient than nonprofit ones. Our results show that for-profit and nonprofit FIOs have similar cost efficiency statistically. We argue that both budget-breaking incentive scheme and monitoring are active in the FIOs because firms that administer bonus scheme also monitor their employees. The mechanisms produce opposing forces and indeterminate end point in this subgroup.

Our findings extend earlier debate on cost efficiency difference between for-profit and nonprofit hospital. Recent empirical research generally finds no cost efficiency difference in recent years, but this finding does not refute the property rights theory. By using the team agency theory, we show how physician incentives modify the effect of capital owner incentives to influence cost efficiency.

Key Words: Integration, Team Agency, Cost Efficiency, Hospital, Stochastic Frontier
JEL Classification: C21, D23, I10, L23.

Table of Content

Table of Equations, Figures and Tables

1. Introduction

Our research question is “How does integrating physicians into hospitals affect hospital cost efficiency under payer-driven competition in the U.S. in 1997?” Understanding the driver of hospital cost efficiency is important for health care cost containment policy for two reasons. First, hospital cost is a major part of U.S. health care expenditure. Second, improving cost efficiency is Pareto efficient. The debate in the 1980s revolves around the issue if for-profit hospitals are empirically more cost efficient than nonprofit ones. The theoretical underpinning is Frech (1976) property rights theory: for-profit hospitals have clearly defined residual claimants that nonprofit ones lack. The residual claimant has strong incentive to monitor cost efficiency to improve profit. However, empirical evidence from the 1990s shows little difference between cost efficiency in nonprofit and for-profit hospitals. Sloan (2000) attributes this result to increased competition in the hospital market forcing capital owners to behave in similar ways. However, theoretical models such as Newhouse (1970), Pauly and Redisch (1973) and Harris (1977), indicate that physicians are more influential than capital owners in allocating resources in the hospitals. Hence, the absence of physician behavior in determining cost efficiency in current studies needs to be addressed. Since 1990s, U.S. hospitals have started hiring physicians as employees (i.e. physician hospital integration) to control cost and mitigate physician effects on profit. This trend is congruent with Pauly and Redisch (1973) model where economic profit accrues to physicians. We formally define physician hospital integration as hospital hiring physician as employee. In this dissertation, we apply the team agency theory to explain how physician hospital integration may affect cost efficiency and produce empirical evidence to support our hypotheses. Introducing physician hospital integration provides new insight using team agency theory to examine the research in hospital cost efficiency.

We approach the dissertation in this way to avoid addressing too many complex issues simultaneously: In this chapter (chapter 1), we provide an overview of the dissertation and some background information about the U.S. hospital industry (in section 1.1) to provide a context for the research question. In chapter 2, we examine the debate on relative efficiency of nonprofit and for-profit hospital. We then formulate a richer theoretical framework by adding the integration dimension using team agency theory. In chapter 3, we assume a general hospital cost function exists and examine the techniques available for examining efficiency. We then summarize the discussion in the context of hospital cost analysis. In chapter 4, we examine the nature of hospital as a firm, its position in the healthcare industry, and review measures of hospital inputs and outputs. In chapter 5, we explain the implementation of our empirical strategy. In chapter 6, we present the result and discussion of our research. Finally we conclude in chapter 7.

1.1 Background to Research

In 2001, the United States had 5,801 hospitals managing 0.987 million beds and consuming 37% of the $1.4 trillion healthcare expenditure (National Center for Health Statistics, 2003). The hospital market is monopolistically competitive because providers are imperfect substitutes in their market segment. A market is monopolistically competitive in the short run when there is no strategic firm interaction [i.e. a firm optimizes its objective function assuming a given set of action of its competitors] and firms produce differentiated products. In the long run, a market is monopolistically competitive when there is no substantial mobility barrier (Chamberlin, 1933; Eaton and Lipsey, 1989).

We observe four market characteristics from the data in Appendix B: First, 85% of all U.S. hospitals in 2000 were community hospitals, owning 84% of the beds. Second, most community hospitals were nonprofit, specifically 52% were nonprofit, 13% were for-profit, 37% belonged to a state or local government, and 4% belonged to the federal government[1]. Third, the occupancy rate and average length of stay steadily declined while the number of outpatient visits and percentage of outpatient surgery significantly

[1] The Federal government owns hospitals belonging to the Armed Forces and Veteran Association.

increased in the 1980-2000 period. This trend reflects both the financial pressures on cost containment and technological advances that reduces surgical trauma. With new technologies, many procedures need shorter post operative hospital stay, or even just outpatient surgery. Lastly, there was a general decrease in capacity (both number of beds and hospitals) for all hospitals except for-profit hospitals. The admission to public hospitals (federal and local) fell, while the admission to private (nonprofit and for-profit) fell and then rose, during these two decades. The reduction in capacity was slower than the fall in demand for beds and created excess capacity and declining occupancy. Together with policy changes in healthcare financing, the excess capacity has led to increased capacity in the hospital market. Many health care analysts link increased competition to the trend in physician hospital integration.

Arrow (1963) defines the medical economy as the complex of services which centers on physicians. The hospital is an institution in this economy. The hospital purchases inputs from factor markets (i.e. pharmaceutical products, medical equipment, non medical goods such as building and food, nursing, administrative and physician time) and transform inputs into output using technology with product uncertainty. Arrow (1963) refers to the uncertainty for medical science to produce consistent outcome as product uncertainty. While physician and capital owners are in separate firms in the traditional institutional structure of hospital, it is hard to conceive a hospital production function without any physician component.

Historically, early U.S. hospitals were philanthropic hospitals providing free care for the poor, or were specialized institutions for psychiatric and infectious diseases. The family provided most of the palliative care from home. Advances in modern surgery in the late nineteenth century created the need for institutionalized care using professionally trained labor and specialized capital equipment. Concurrent urbanization was conducive for hospital expansion for two reasons. First, urbanization created new health problems in a crowded environment. Second, the urban working class had increased opportunity labor cost and better ability to pay because of employer sponsored health insurance. These developments favored substituting institutionalized care for home care. By the mid twentieth century, medical technology could modify the course of many diseases with institutionalized care [see Fuchs (1974) for an interesting discourse of the economic history of medicine]. Institutions today provide definite advantage over home for post operative care and during the critical phases of many diseases. Hospitals provide the economy of scale to acquire capital equipment for physicians.

Physicians and hospitals in almost all private U.S. hospitals before the 1980s were separate legal entities, i.e. segregate hospitals. Arrow (1963) states that medical technology cannot precisely predict the outcome of diseases and calls this property "product uncertainty" (Later in Chapter 3, we shall argue that product uncertainty is an interpretation of care quality). Arrow argues that with product uncertainty, risk becomes non marketable and ideal insurance becomes impossible, i.e. it is not possible to pay care providers (i.e. physicians) based on the benefit the consumers (i.e. patients) receive. The social institution which arises to solve this economic problem is the agency relation between physician and patient (otherwise known as medical ethics, professional relation and so on): The patient entrusts consumption decisions to the physician who knows better about the production and utility of the health states. While agency relations exist in many professional relations (such as between a lawyer and a client), the physician-patient relation is unique in two aspects: the consequence is very severe, and physician has better knowledge about patient's health utility. Arrows' work shows that uncertainty in medical technology gives rise to the expected physician behavior to be the perfect agent for patients. Fuchs (1974) suggests that segregation of capital ownership from physician removes the inherent conflict between profit motive and the fiduciary duties to deliver the highest possible care quality. Holmstrom and Milgrom's (1991) multitask agency model provides the theoretical insight to Fuchs (1974). The authors consider a principal who assign two tasks to an agent where only one task is measured easily. If the principal also implement a performance incentive, the agent will neglect the task which is difficult to measure. Care quality is more difficult to measure (than number of discharge) because of product uncertainty. Removing physicians from financial incentive ensure the delivery of the highest feasible care quality. However, an unwanted effect of this arrangement is the over utilization of resources that may improve care quality and increase cost. As a physician makes clinical decisions without cost consideration, he prefers a "likely useless but harmless" intervention to no intervention.

How does product uncertainty arise in the hospital market? To be precise, we have a clear understanding of some diseases (such as polio), an imperfect understanding of most diseases (such as coronary artery diseases), and vague knowledge of a few diseases. Borrowing the terminology from Thomson (1975), these states are respectively known as high technology, halfway technology and non-technology. The state of technology influences the cost effectiveness of intervention: high technology has low cost because the cause is well known and effective treatments are available. For example, the cost burden of polio is low because we have effective antibiotics to treat and vaccines to prevent the disease. Non-technology has low cost because no treatment is available and care givers can only provide symptomatic relief. This category comprises two extremes: terminal diseases which are given palliative care; and self-limiting idiopathic[2] diseases which are often self-medicated. Weisbrod (1991) argues that halfway technology is the most expensive because partial treatments are available in the hospital. The third party payer system insulates the physician and patient from cost consideration and promotes over use of halfway technology. As information about a patient's disease often unfolds over time, the public health referral system is a social institution to minimize the cost burden of diseases: The primary care level (consisting of family physicians and self medication) treats most of the high technology and self-limiting non-technology cases. This level also serves as a gatekeeper for expensive halfway technology in secondary and tertiary care hospitals. Hence, providers in the medical economy deliver a complete range of care during a disease episode, and hospitals deliver the most expensive halfway technology.

Broadly speaking, three institutional structures dominated the U.S. hospital market during different periods. They are namely cost reimbursement, prospective payment system (PPS), and managed care. The cost reimbursement structure dominated the market just after World War II: independent physicians and nonprofit hospitals were financed by cost

---

[2] Idiopathic means 'of unknown cause'

reimbursement from insurance for the care that the patients received. Hospitals could not compete for patients with direct financial payment to physician because of the physician-patient relation which Arrow (1963) describes. Instead, hospitals competed for referrals by investing in capital equipment to attract physicians (this is called Medical Arms Race theory, or MAR). Ethovan (1978) points out that Medical Arms Race is a non-price competition which drives cost above the social optimum. When there are few hospitals in an area (i.e. high industry concentration), Medical Arms Race is less intense, causing cost and price to decrease. In other words, the Medical Arms Race theory predicts that price and concentration is inversely related. Several studies (Robinson and Luft, 1985; Robinson, Garnick and McPhee, 1987) find this effect from data in the 1970s and early 1980s, which has largely disappeared by the late 1980s.

The primary event that eroded MAR is the introduction of prospective payment system. Healthcare financing authority[3] (HCFA), the largest insurer in the U.S., introduced the Prospective Payment System (PPS) in 1983 to reimburse hospital services. Prices of hospital admissions were fixed *ex ante* using diagnostic related group (DRG). In 1992, HCFA extended the method to cover physician services using resource based relative value system (RB-RVS). Other insurers quickly adopted these PPS schemes. PPS becomes standard practice by the 1990s. Shleifer (1985) points out the theoretical underpinning of PPS and coins the term yardstick competition: a seller has the incentive to select efficient technology since a buyer pays the average cost (i.e. price = average cost). Yardstick competition introduces incentive to minimize the cost per admission. However, since DRG does not capture care quality sufficiently, hospitals have incentive not to admit severely ill patients and discharge them early to reduce cost (Dranove, 1987; Ma, 1994).

The advent of managed care [4] and selective contracting is the third structural change which gradually modifies competition. Despite initial resistance from physicians, the

[3] On July 1, 2001 HCFA became the Centers for Medicare & Medicaid Services (CMS).
[4] Managed care arrangements refer to diverse institutions such as staff and group HMO (Health Maintenance Organization), IPA (Independent Practice Association) and PPO (Preferred Provider Organization). In the late 1990s, IPA dominates while various vertically integrated forms such as hospital or physician sponsored network are growing.

signing of HMO Act (1973) into law provided the required regulatory environment for growth of managed care. By 1993, over 70% of all U.S. health insurance enrolled in some form of managed care [see Glied (2000) for a detailed discourse on managed care]. Competitions among hospital since the 1990s were increasingly payer driven with the objective of becoming a member of the provider panel (Dranove, Shanley and White, 1993). Managed care introduces even greater cost pressure because a hospital needs to minimize cost per patient instead of cost per admission. Contrary to the Medical Arms Race scenario, the price concentration relation is now positive because the hospital can resist pricing pressure better when there are fewer hospitals in an area. Therefore, three decades of cost containment policies have increased competition in the hospital markets. Another important trend in 1990s was the rise of various forms of physician hospital (vertical) integration to compete for managed care contract (Shortell and Hull, 1996). Many researchers, such as Burns and Thorpe (1995), Shortell *et al* (1996), believe there is a causal relationship between managed care penetration and integration. Consultants and practitioners, such as Advisory Board (1993) and Dowling (1995) develop multistage market evolution models to characterize this association.

Physician hospital integration raises some important questions in the context of cost containment. Does physician hospital integration improves cost efficiency? How? Will there be any difference between integration of for-profit and nonprofit entities? Varney (1995) states that pro-competitive benefits can occur through reducing agency cost in integration. An explanation may exist in the economic theory literature, but no one has yet applied it to answer these questions. This is the primary contribution of our research.

1.2 Research Problem and Hypotheses

This section consists of key ideas of the analytical framework discussed in Chapter 2. We wish to formulate and test a theoretical model to explain how hospital cost efficiency can change when physicians become employees. Holmstrom (1982) examines efficiency under team agency, i.e. a team of agents jointly produce the output for a principal. Joint production means each agent produces nothing independently, but the team can jointly produce the output. Holmstrom shows that externality will cause shirking under team agency (a moral hazard problem). The principal can overcome shirking to achieve Pareto efficiency if she administers a reward scheme similar to a salary cum bonus scheme (Holmstrom calls this a budget breaking scheme). However, Eswaran and Kotwal (1984) shows that budget breaking will only work if the principal does not have the incentive to sabotage the team to maximize profit (the principal's own moral hazard problem). Therefore, we can form two hypotheses from the above arguments after classifying hospitals into three groups: fully integrated organizations (FIOs) that hire physicians; networks which form alliances with physicians; segregate hospitals where physicians are independent:

H1: In nonprofit hospitals, FIOs are more efficient than network and segregate hospitals
H2: In for-profit hospitals, we will not observe the cost efficiency difference

When there is no opportunity to administer the salary cum bonus scheme (in network and segregate hospitals), the principal can monitor the agents.[5] Here, we can apply the property rights argument (Alchian and Demsetz, 1972; Frech, 1976) which predicts a for-profit entity is more efficient than a nonprofit one. We obtain the third hypothesis:

H3: For-profit network and segregate hospitals are more efficient than nonprofit ones.

There is no *a priori* reason that for-profit FIOs are more efficient than nonprofit ones because of two opposing forces: first, property rights theory predicts that for-profit is more efficient; second, team agency predicts that nonprofit is more efficient because the principal does not face moral hazard. Therefore, the property rights effect in FIO is attenuated and can go both ways.

[5] It is easier to monitor manager than physician.

1.3 Justification for Research

This research extends the property rights debate (i.e. for-profit hospitals are more cost efficient than nonprofit ones) by adding the physician incentive dimension. As Sloan (2000) points out, recent empirical evidence shows that hospitals with different capital ownerships have similar cost efficiency in an environment of increased competition. However, a mere comparison of cost efficiency of nonprofit and for-profit hospitals omits the powerful moderating effect of physician influence on resources allocation. We propose a richer model when comparing cost efficiency between for-profit and nonprofit hospitals by adding the role of team agency in physician and hospital manager.

Our results have direct implication for firm strategies and antitrust regulations in vertical hospital merger. Burns, Gimm and Nicholson (2005) show that initial investment in hospital merger can adversely affect financial performance. Our results show where the payoffs for hospital vertical integration can arise. Varney (1995) argues that hospital vertical integration can be pro-competitive by reducing agency cost. We apply the result from team agency theory to show how vertical integration can be pro-competitive. Therefore, our primary contribution is a theoretical application with empirical support toward policy regulation in health care antitrust and cost containment policy.

1.4 Methodology

We use quantitative method to examine the production unit for hospital care (i.e. the unit of analysis). In chapter 4, we will argue that the hospital and its affiliated physicians this form a production unit to deliver patient care. The concept of production unit is in line with theoretical models such as Pauly and Redisch (1973). Our constructs are factors related to cost efficiency and classification of hospitals. We can examine our research methodology using three types of validities: construct validity, internal validity and external validity (Trochim, 2000).

The data universe in our research question is the set of all U.S. hospitals in 1997. We obtain our data from six sources: the American Hospital Association's annual survey; the National Inpatient Sample from Agency of Healthcare Research and Quality; the Wage by Area and Occupation survey from Bureau of Labor Statistics; and the CPT4-ICD9CM crosswalk file from Info-X Inc[6]. Our sample consists of a cross section of 313 hospitals, and we are satisfied that these hospitals are similar to the National Inpatient Sample[7] version 6 in terms of casemix and important hospital characteristics.

Trochim (2000) states that the three pre-requisites for inferring causal relation (internal validity) are: correlation, temporal precedence and lack of alternative explanations. Controlled experiment [8] has the highest internal validity follow by quasi-experiment and cross section observation. However, only cross section observation and quasi-experiment of hospital type conversions are feasible for our research question. This is because we cannot assign constructs for capital ownership or organizational structure to hospitals. We can compare pre and post conversion equilibrium cost efficiency in quasi-experiment by observing cases over suitable periods. However, this method suffers from two related limitations. First, the time to reach stable cost efficiency after conversion is unknown, making suitable observation difficult to define. Meanwhile, environmental shocks can cause unpredictable change in cost efficiency. Second, the number of conversion is much smaller than population size. The small size increases the effects of outlier. Our next alternative is to observe a large cross section of firms. This is the most common method in econometric modeling. However, we cannot establish temporal precedence using this method. Without establishing temporal precedence, we need to assume that cost efficiency is stable in the sample to obtain valid result. Outliers may arise because some

---

[6] Info-X Inc. generously supplied the crosswalk file from its commercial computer program. The file maps all possible CPT4 codes to ICD9CM codes and vice versa. CPT4 means current procedure terminology version 4; it is the code physician use to submit billable procedure to insurer. ICD9CM means international classification of disease for clinical management version 9; it is the code hospital use to bill insurer. Crosswalk is an insurance jargon that means mapping one code base to another.

[7] The Agency of Healthcare Research and Quality uses a stratified sampling frame to ensure that the National Inpatient Sample is representative of the hospitals in the participant States.

[8] According to Trochim (2000), controlled experiment is a research method that has five elements in its design: comparative groups (control and treatment groups), random assignment of sample to group, treatment administration (the independent variable), and measurement of pre and post treatment effects.

firms are not at their equilibrium efficiency just after environmental shock or organizational conversion. A large sample size mitigates the effect of outliers. A better way is to accommodate some random shocks using stochastic frontier instead of DEA or deterministic frontier. The choice between quasi-experiment and observation weighs slightly to the latter, and confirmation of the result using quasi-experiment will be fruitful.

1.5 Chapter Summary

This chapter provides an overview of the structure of the dissertation and background information about the U.S. hospital market. The next three chapters (chapters 2 to 4) present the research issues in greater detail before proceeding to its execution.

2. Theory of Hospital Cost Efficiency

All firms are cost efficient in the neoclassical world, and cost function is the result of successful cost minimization behavior for a given technology. Deviation from the neoclassical ideal constitutes cost inefficiency. There are several theories about physician behavior which may cause deviation from cost frontier in hospitals. Newhouse (1970) argues that nonprofit hospital jointly maximize quantity and quality. However, too many resources are allocated to care quality for two reasons: First, managers in nonprofit hospitals are not evaluated on profit performance. Second, trustees and physicians prefer high care quality. Therefore, nonprofit hospitals may invest in conspicuously prestigious but inefficient technology. Newhouse argues that measuring hospital (patient care) output requires quantity and quality as joint proxies. In contrast to the neoclassical production function that requires only quantity to proxy output, production uncertainty in hospital technology gives rise to the need for joint proxies. Arrow (1963) explains the meaning of production uncertainty[9]: Within acceptable medical practices, subtle differences in the treatment produce variations in the patient's health status. Pauly and Redisch (1973) model the hospital as physician's cooperative that maximizes the physician's average revenue. Physicians, not capital owners, are the dominant decision makers in allocating resources in hospitals.

These two models share three similarities. First, both equilibriums are not Pareto efficient. Quality is excessive in Newhouse's model, and excessive profit accrues to physicians in Pauly and Redisch's model. Second, physician allocates resources in the hospital even when physicians and hospitals are separate legal entities. This situation contrasts sharply with neoclassical theory where the capital owner allocates resources to maximize profit. Furthermore, hospital care is jointly produced using physician and hospital resources. Therefore, the physician is technologically part of the hospital under the neoclassical

[9] The term 'product uncertainty' is used in Arrow (1963). Technology means the relation between input and output embodied in a production function.

framework, and firm internalizes technology. Adopting the neoclassical framework allows us to use results in the cost theory literature described in later chapters. Third, the Newhouse (1970) and Pauly and Redisch (1973) models are applicable to nonprofit hospitals where physicians are not hospital employees. The principle (patient)-agent (physician) relation in Arrow (1963) drives these two models and the hospital (administrator) has no role. The difference in the agent's behavior results in excessive profit in Newhouse (1970) and excessive quality in Pauly and Redisch (1973).

Harris (1977) views hospitals as two interdependent firms. A hospital provides input to a physician in a complicated and uncertain sequence of events. The hospital manager solves the rationing problem with non-price related decision rules such as rule of thumb and side bargain. Harris's (1977) model captures two important aspects of hospital operation. First, the production technology is *ex ante* uncertain and bargaining becomes a mechanism to allocate resources. Second, the joint production of hospital services is even clearer in the Harris model than the previous two models. Furthermore, the Harris's model is not restricted to nonprofit hospitals unlike Newhouse's (1970) or Pauly and Redisch's (1973) models. Although theoretical models from Newhouse's (1970), Pauly and Redisch (1973) and Harris (1977) suggest the sources of cost inefficiency, there is no empirical analysis that relies on these models. We present our hospital classification in the next section before discussing our synthesis of hospital cost efficiency research[10].

2.1 Classification of Hospitals

We classify hospitals along two dimensions for the purpose of this dissertation. First, we can classify hospitals as nonprofit and for-profit unambiguously. Second, we classify hospitals into those hiring physicians as employees (i.e. integrated) and those which do not. The 1997 American Hospital Association (AHA) annual survey lists seven[11] mutually exclusive relationships between hospitals and physicians, namely integrated

[10] Introducing the classification scheme at this point is convenient for the reader although the scheme arises from the synthesis later.
[11] The default relationship is independent physicians and hospitals (i.e. segregate hospital).

salary model, equity model, foundation model, independent physician associations, group practice without walls, management service organization, (close and open) physician hospital organization (PHO).

The first three models are integrated models. In the equity model, (senior) physicians form a company to own assets. The for-profit company directly hires physicians (both owner-physicians and additional ones) as employees. In the integrated salary model, a nonprofit hospital (or non physician investor) directly hires physicians as employees. In the foundation model, the nonprofit hospital forms a foundation as a subsidiary. The hospital hires physicians indirectly because the latter are technically the foundation's employees. The physicians in all three cases have employment contracts. For the purpose of this dissertation, we call these arrangements fully integrated organization (FIO).

The physicians in the next four models usually have (managed care) service contracts with the hospitals. Group practices without walls (GPWWs) are the most common type of practice group today. Physicians in most GPWWs maintain independent practice but negotiate managed care contract as a group with the hospital. Financial arrangements vary from group to group. Some GPWW leaders decide to incorporate as a medical group, consolidate support staff and standardize procedures (such as credentialing standards). Key owners become managers, issue stock and set up profit-sharing plan.[12] The independent physician association (IPA) is a group practice formed by physicians and tends to be well financed (often backed with venture capital or corporation). The IPA negotiates with payers for a capitation rate including physician fees, then reimburses the physicians (although not necessarily using capitation). Both IPA and its members share the risk of medical costs if capitation payment is lower than required reimbursement for physician. An IPA can be a pure physician cooperative or a mix alliance of physicians and hospitals. The physician hospital organization (PHO) is a partnership between hospitals and physicians to either co-ordinate the delivery of healthcare services to a defined population, or contract directly with a self-funded employer group and/or

[12] This is effectively forming a medical group consisting of physicians. If the owners are hospitals, then the arrangement will become one of the integrated forms in the previous paragraph.

government program. In its weakest form, the PHO is a messenger who analyzes terms and conditions the payer offers before letting each physician decides individually on participation. Usually, the participating physicians and hospitals have standard contract terms to negotiate with payer within a given time frame. The terms are binding if the PHO succeeds; otherwise the participants can negotiate with payers directly after the time window expires. The PHO is a bargaining vehicle where the service contract is between the provider and the payer. PHO is a primer for vertical integration and is often formed as a reaction to selective contracting in managed care. The open PHO opens its enrollment to all hospital accredited physician and is often specialty dominated. The closed PHO limits physician membership by specialty or practice profiling. The management service organization (MSO) is an independent corporation owned by a hospital or PHO. It provides administrative services for a fee to the affiliated medical practices, and also serves as a vehicle for planning and contracting with payer. These arrangements provide some co-ordination for managed care contracting and non-medical administration for the hospital-physician or physician-physician groups. For the purpose of this dissertation, we refer to these arrangements as network hospital.

The medical staff model is the *de facto* arrangement between hospitals and physicians. Under this arrangement, a specialist (as oppose to primary care) physician applies to the hospital medical staff committee[13] to obtain the privilege to admit his patients (i.e. become a medical staff). Theses physicians are independent professionals and may apply for medical staff in competing hospitals. For the purpose of this dissertation, we refer to this arrangement as segregate hospital.

2.2 Cost Efficiency in Nonprofit and For-profit Hospitals

The model which attracts most empirical work is Frech's (1976) property rights model. Frech argues that owners of for-profit hospital maximize profit because of property rights:

[13] The committee comprises existing medical staff and representatives of the hospital and physicians. The Joint Commission of the Accreditation for Hospitals sets guidelines for the operations of the committees.

When it is possible to clearly define individual (capital) owners, they will have the incentive to ensure that the firm is run efficiently (Alchian, 1961; Alchian and Demsetz, 1972). Property rights empower owners to allocate firm's resources, keep the residue, and capitalize on wealth gained by selling the rights. In an efficient capital market, capital owners will discipline inefficient firms by divesting their capital. This mechanism is absent in nonprofit and public hospitals. The attenuated property rights in these hospitals enable hospital managers to pursue non-pecuniary objective at the owner's expense and therefore reduces efficiency. These hospitals do not maximize profit because of imperfect agent. De Alessi (1983) argues that managers in for-profit hospitals are more likely to introduce cost saving innovations and adopt cost minimizing input combinations to lower costs, than their nonprofit hospital counterparts. The Harris's model suggests that hospital manager plays an important role in determining cost efficiency. The hospital manager bargains with physician to allocate resource (even when physicians are not hospital employees), and the bargained outcome influences cost efficiency. Managers in nonprofit and public hospitals maximize an objective function which includes profit and non-pecuniary benefits. Therefore, the theory of property rights motivates research on the efficiency difference between hospitals with different capital ownerships.

However, empirical analysis of hospital cost efficiency is plagued with methodology problems. Almost all U.S. empirical studies to date use accounting data collected assuming hospital as the unit of analysis. For U.S. hospitals that do not hire physician as employees, the cost data do not contain any physician related component. As a result, analyses using these data are not consistent with theoretical models that include physician in the production of hospital outputs. Some U.S. hospitals hire physicians as employees and the accounting data contain physician costs. This sub-set of data will not be comparable with most hospital data. We shall defer the methodology debate on modeling efficiency, and specific issues on input measurement to Chapter 3.

The results from these empirical studies are not conclusive. Wilson and Jadlow (1982) as well as Herzlinger and Krasker (1987) find higher efficiency in for-profit hospitals; Sloan and Vraicu (1983), Becker and Sloan (1985), Gaumer (1986), Shortell and Hughes

(1988), Patel, Needleman and Zechauser (1994), and Sloan *et al* (1998) find no difference. Zuckerman, Hadley and Iezzoni (1994) even find nonprofit hospitals to be more cost efficient than for-profit hospitals. Sloan (2000) succinctly summarizes the current position of the empirical research on the ownership-efficiency nexus: 'There is probably not much difference in technical and allocative efficiency, and if any existed, the increased competition in the 1990s would narrow it [14]..." Furthermore, the increased reliance for all hospital on debt capital [15] and decreasing donation implies that hospitals are becoming more similar than different."

The property rights theory predicts that for-profit hospitals are more cost efficient than nonprofit ones. On the other hand, Fama (1980) highlights that we should not confuse capital ownership with the control of the firm, specifically: "… the control over a firm's decision is not necessarily the province of security holders". Jensen and Meckling (1976) view the firm simply as legal fiction that serves as a nexus for a set of contracting relationship among individual production factor where capital ownership is only one of them. Our earlier discussion of hospital behavioral models by Newhouse (1970), Pauly and Redisch (1973) and Harris (1977), indicates that we need to consider the role of physicians in hospitals. The link between performance and ownership is strong when physician interests are minimal (Schlesinger, Marmor and Smithey, 1987). Including physician influence in the empirical analysis of hospital cost efficiency is a knowledge gap that needs to be addressed. This is especially true when new organizational arrangements between physicians and hospitals have appeared in the U.S. since 1990s. These new arrangements arise to compete for managed care contracts in the era of payer driven competition. Do these new organizational arrangements have higher cost efficiency? If yes, what is the theoretical reason? If no, these new organizational forms can attract antitrust regulation because Cuellar and Gertler (2006) find they can exercise market power in the hospital service market.

[14] It is increasingly infeasible for nonprofit hospital to fund non-pecuniary objectives, whatever these are, with limited donation income in the hospital revenue stream in the U.S. In this sense, hospital ownership research is a declining industry (Sloan, 2000).

[15] Unlike for-profit hospitals, nonprofits are exempted from tax and do not benefit from debt tax shield.

2.3 Team Agency and Cost Efficiency

We begin our analysis by examining a theoretical model. Holmstrom (1982) considers a model with a principle and two agents jointly producing a product. Each agent takes an unobservable action a with private cost $v(a)$ to produce an output $x(a)$. The principal sells the output and allocates a share s to everyone. A balanced budget sharing scheme means $\Sigma s(x)= x$ and each agent obtains the payoff $s(x(a))-v(a)$. Suppose the sharing rule is differentiable, then Pareto efficient Nash equilibrium means the optimal effort $a^* = argmax[x(a)-\Sigma v(a)]$. Differentiating the payoff obtains $s'x'-v'=0$, while the Pareto optimality implies $x'-v'=0$. These two results jointly imply that $s'=1$. However, differentiating the balanced budget $\Sigma s(x)= x$ obtains $\Sigma s'(x)= 1$, which contradicts the earlier result that $s'=1$. Therefore, the conditions for budget balancing and Pareto efficiency cannot simultaneously coexist. This is because moral hazard (i.e. shirking) can occur in multi-agent production even without uncertainty in technology. The principal cannot identify the agent who cheats even if she can observe the joint output as an indicator of inputs. (Contrast this to the single agent case where the principal can identify shirking under certainty). Holmstrom suggests using a budget breaking sharing rule (i.e. $\Sigma s(x)<x$) to overcome this free rider problem. Specifically, let each agent's share be $s_i(x)=b_i$ if $x>x(a^*)$ and $s_i(x)=0$ if $x<x(a^*)$, where b is an arbitrary real number. The solution for b_i satisfies two conditions $\Sigma b_i= x(a^*)$ and $b_i>v(a^*)>0$, which means total share is the optimal output provided that each share is bigger than the individual's private cost (and is positive). This Nash equilibrium is Pareto efficient because $x(a^*)-\Sigma v(a^*)>0$, i.e. the optimal output is greater than total cost to all agents. The budget breaking scheme therefore neutralizes externalities in joint production. We can implement the scheme as a basic wage plus bonus/punishment. In a dynamic context, such punishment means firing the employee. Holmstrom (1982) also shows that budget breaking scheme holds under uncertainty.

Eswaran and Kotwal (1984) argue that under Holmstrom (1982) scheme, the principal herself faces moral hazard. The authors assumes there is a residue $R(x)= x-\Sigma s(x)$ which can be partially used as a bribe *e.* The self-interested principal can offer a secret scheme to one of the agents to receive *b* if $x>x(a^*)$, $b+e$ if $\bar{x} < x < x(a^*)$ and zero if $x < \bar{x}$ where $\bar{x}, e>0$ and $\bar{R} = \bar{x} - b - e > R(x(a^*))$. The equilibrium output is $\bar{x}$ which is less than $x(a^*)$. Therefore, introducing budget breaking gives the principal the incentive to engage in morally hazardous behavior. In the context of for-profit hospitals which hire physicians (i.e. FIO), there are several ways in which this "bribe" can be offered. For example, the hospital owner can start a project that reduces cost efficiency in a profitable year. The project administrator can receive the reward as a promotion or one-off perk (such as project funding).

The works of Holmstrom (1982) and Eswaran and Kotwal (1984) provide us a theoretical framework to examine cost efficiency in hospitals. To begin with, a production function without either physician or hospital services is infeasible (violate the property that input requirement set *V(.)* is non-empty). Physicians and hospital managers jointly produce patient care even when physicians are not hospital employees. When the physician is an employee, the physician and the manager form a team of agents working for a principal who is the residual claimant. The physician is a double agent [16] because he has two principals (i.e. his employer and the patient). Arrow (1963) explains that agency relation between patients and physicians arises because of product uncertainty. Grossman and Hart (1986) define the firm using assets which it owns: firm ownership is about possessing the residual rights over assets. Residual rights are rights not taken away by contract with other parties; specific rights are rights memorialized in contracts. Residual rights negate the cost to set down all possible specific rights in a contract. The boundaries of the firm are delimited by the control of residual rights. In the U.S. context, hospitals can purchase physician services from the market[17] or internalize them within the firm.

When the physician is an employee, the hospital owner can administer a salary cum bonus scheme to the production team. If the principal does not face moral hazard problem

[16] To be sure, the physician often faces conflict between being a professional and an employee. However, corporate medicine is not impossible as in the British National Health System.
[17] We use Grossman and Hart (1986) to anchor the hospital as a firm. Technologically the physician is still part of the production function regardless of the legal relation between physician and hospital.

as Eswaran and Kotwal (1984) describe, this budget breaking scheme will overcome the externality problem to achieve Pareto efficiency according to Holmstrom (1982). If the physician is not an employee, then the principal cannot administer the scheme to the production team. At most she can administer the scheme for the hospital manager because the physician has a service contract (while the manager has a labor contract)[18]. Therefore, integrated hospitals are more efficient than non-integrated ones if there is no principal moral hazard. This scenario happens in nonprofit hospital because there is no profit incentive. On the other hand, moral hazard in the principal occurs in for-profit hospitals and the difference in efficiency will be nullified. Hence, we hypothesize that:

H1: Nonprofit integrated hospitals are more efficient than non-integrated ones[19]
H2: For-profit integrated hospitals are not more efficient than non-integrated ones.

2.4 Monitoring and Cost Efficiency

What happen in non-integrated hospitals when it is not possible to administer the salary cum bonus scheme? An alternative mechanism to this scheme is for the principal to monitor the agent. Alchian and Demsetz (1972) suggest the monitor is the residual claimant for incentive compatibility. Claiming the residue provides property rights. Frech (1976) argues that for-profit hospitals are more efficient than nonprofit ones because there are clearly identified residual claimants. Even in modern firms where management is independent of residual claimant, market discipline from external security (bond and equity) market can act as monitoring agency on managers. Fama (1980) explains how the security market can discipline manager through the effects of managerial labor market. Unlike security a holder who diversifies his wealth through portfolio, a manager invests

[18] The managed care relations with hospital and physician are contracts for services. Therefore, it is difficult to administer the budget breaking scheme because of the ramifications of the service providers. Even if distinct physician group service each hospital, neither hospital nor physician are within the same firm in the Grossman and Hart (1986) sense. It is extremely difficult for physician to receive financial rewards from outside the firm as explained in Arrow (1963). Technologically, hospital care production and financing are distinct.

[19] We measure efficiency using a real number between zero and one calculated by stochastic frontier model.

substantial amount of his wealth (i.e. the wage income stream) in one firm as human capital. The manager is mobile across firms due to constant recruitment by firms, and hence his future wage stream is revaluated. As long as the security market evaluates the firm's value efficiently, a prospective employer can use the firm's value as proxy for the candidate's competence to determine the candidate's future wage. In this case, the manager has the incentive to ensure efficient operation of the firm by monitoring management level above and below him. As long as there is competition for the top job, the top manager will also be monitored by the next level. Fama and Jensen (1983) further pursue the market factor argument that if everything fails to overcome the agency problem, there is always hostile takeover as the final market disciplining device. The driving force for this argument is an efficient capital market where securities are traded to discover price. This happens in the for-profit economy but not in a nonprofit one. Hence, without the influence of team agency argument in the previous section, for-profit hospitals are more efficient than nonprofit ones. We form our third hypothesis as:

H3: In non-integrated hospital, for-profit hospital is more efficient than nonprofit one.

What happens when we compare for-profit FIOs with nonprofit FIOs? Team agency theory predicts that nonprofit FIOs will be more efficient than for-profit FIOs because of the principal's moral hazard problem; property rights theory predicts that nonprofit FIOs will be less efficient than for-profit FIOs. There is no *a priori* reason which force will dominate, but we are likely to obtain no difference statistically.

2.5 Cost Efficiency in Public Hospitals

The common thread in team agency and property rights models is the behavior of payoff maximization in each economic agent. In the agency team, each agent's payoff is the difference between his profit share and the private cost for his effort. The principal's payoff is the residual claim. The literature shows that incentive in public hospitals is more complex than team agency and property rights models.

Lindsay (1976) argues that the Congress is the principal in public hospitals. The Congress's objective is to please voters, so that providing the maximum output for an allocated budget is a desirable outcome for public firms such as public hospitals. Unlike private firms, public firms do not provide output at market price. In fact, outputs are often provided free of charge to the intended consumers. Public firms exist precisely because of market failure to supply certain goods, such as uncompensated medical care, which is politically unacceptable. The Congress cannot use profit to meter the performance of manager in public hospital since the public hospital's output is not priced at market rate (even if this is a regulated price). The managerial labor market in Fama (1980) cannot discipline a public hospital's manager because the firm's performance is not financial measures. Lindsay argues that the Congress uses visible indicators, such as patient day and per-diem cost, to meter manager's performance in public hospitals. The level of care quality in public hospitals is lower than private hospitals because public hospitals just need to meet minimum standards. Patients can boycott private hospitals to force private hospitals to improve quality. Consumers need to complain to the Congress to force public hospitals to raise quality. This process is costly to consumers and provides less timely information. Lindsay (1976) shows that Veteran Affairs hospitals have longer average lengths of stay and lower per diem costs than private hospitals.

Wilson (1989) argues that government agencies have multiple objectives. Government programs have distributional effects where consideration of equity and accountability are often more important than economic efficiency. Therefore, the budget breaking mechanism will also break down because of equity and accountability considerations. We conclude that our theoretical framework has limited application to public hospitals. Cost efficiency difference between public and private hospitals indicates difference in objectives. Therefore, we exclude public hospitals in this research to increase internal validity.

2.7 Chapter Summary

This chapter provides an analytical framework for examining hospital cost efficiency using theories of team agency and property rights. We have developed three hypotheses and identified a situation where theses two theories provide opposite predictions. Given the context of cost containment from Chapter 1, cost efficiency is a natural measure for Pareto efficiency in the discussion in this chapter. Up to this point, we have said nothing about how we can measure cost efficiency. This is the subject of our discussion in the next chapter.

3. Theory of Cost Efficiency Measure

We assume that we can specify hospital production, cost or profit functions as required in the following discussion. We will discuss the issues of specifying hospital cost function in Chapter 4. In line with conventional treatment, we shall first discuss production efficiency and extend the result to cost efficiency. The neoclassical production theory assumes economic agents are successful in maximizing output subject to technological constraints (Similarly, they are successful in minimizing cost and maximizing profit in the respective cases of cost and profit functions). Depending on whether the function is production, cost or profit, there is a positive or negative residue if the economic agent fails in the constrained optimization. However, typical regression analysis produces an estimate of the function's average level that fits the data. The residue can be positive or negative, but neoclassical economic theory only allows either one to exist depending on the estimated function. This paradox arises because the technique estimates mean rather than the frontier. The frontier is the benchmark for measuring efficiency.

Only technical efficiency is meaningful for production frontier. Koopmans (1951) defines an output-input vector[20] (y,x) is technically efficient, if and only if, (y',x') is not feasible for $(y',-x')>(y,-x)$. From this definition, Debreu (1951) and Farrell (1957) suggest definitions for technical efficiencies. The input-oriented technical efficiency arises from the firm's ability to minimize inputs; the output-oriented technical efficiency arises from the firm's ability to maximize output. We can use a suitable distance function to measure technical efficiency from the single output production frontier, or an isoquant for the single and multiple output cases. If the assumption of cost minimizing behavior is appropriate, the cost frontier provides the standard to measure cost efficiency. Achieving input-oriented technical efficiency is necessary but not sufficient for cost efficiency. This is because a technically efficient producer can use inappropriate input mix for a given set of input prices. We can decompose cost efficiency into technical efficiency and allocative efficiency components. Decomposing profit efficiency is even more exacting than cost

[20] y denotes the output quantity, x denotes the input quantity, see Appendix A for a summary on notation.

efficiency. Profit efficiency requires output-oriented technical efficiency, input-oriented allocative efficiency, and output-oriented allocative efficiency (Kumbhakar and Lovell, 2000. p.60).

3.1 Measuring Technical Efficiency with Production Frontier

There are three ways to measure technical efficiency using a production function: non-statistical frontier, (statistical) deterministic frontier, and (statistical) stochastic frontier (Schmidt, 1986).

For a production function [21] $y_i=f(x_i,b)$, the corresponding production frontier is $y_i=f(x_i,b)TE_i$, where TE_i is i^{th} firm's technical efficiency. Taking log, $lny_i=lnf(x_i,b)+lnTE_i = lnf(x_i,b)-u_i$, where u_i is the technical efficiency because $u_i = -lnTE_i \approx 1-TE_i$. We can estimate u_i either as a slack using mathematical programming, or as an error term using regression. We obtain Data Envelopment Analysis (DEA) using mathematical programming, and deterministic frontier using regression. Data Envelopment Analysis uses best practice observations to trace the production frontier [see Charnes, Cooper and Rhodes (1978) for a review of the technique]. This technique is sensitive to noise. The regression technique gives us the deterministic frontier. [Note: there is only one error term here. The stochastic frontier method that we will discuss later has two error terms. The literature often refers to these two error terms as composed errors or composite errors]. We cannot obtain the one-sided error term u_i directly from regression using ordinary least square (OLS). There are three proposed methods in the literature to do this. First, Winstein (1957) proposes the corrected ordinary least square method (COLS)[22]. This method requires finding the error terms e_i from the regression, followed by shifting the error up by a constant term until none of the error terms is negative, i.e. $u_i = max(e_i)-e_i$. Second, Richmond (1974) proposes the modified least square method (MOLS). Richmond obtains the error term e_i from ordinary least square regression follow by

[21] *f(.)* denotes the production function, the arguments are input quantity x_i and function parameters *b.* See Appendix A for the notation convention.

[22] Gabrielsen (1975) is usually credited for defining the corrected ordinary least square.

estimation of the variance of the error term. The author then assumes the error is a half-normal distribution and uses the variance to estimate the expected value. Finally, he shifts the error term up by the expected value to recover u_i, i.e. $u_i = E(e_i) - e_i$. Other one-sided distributions can also be used. Stevenson (1980) uses the exponential and truncated normal, and Greene (1980) suggests the gamma distribution. Afriat (1972) suggests the third method to estimate u_i using maximum likelihood estimator. These three methods estimate the deviation from the deterministic frontier and require assuming any deviation as inefficiency. Hence, the deterministic frontier and data envelopment analysis treat statistical noise as inefficiency. The difference between these two techniques is: deterministic frontier is a statistical technique that allows for drawing inference outside the sample, but data envelopment analysis only allows us to specify bounds when drawing inference.

The problem of treating statistical noise as inefficiency is the motivation for developing stochastic frontier models. Aigner, Lovell and Schmidt (1977), and Meeusen and van den Broeck (1977) pioneer the stochastic frontier model by introducing two independent error terms. The first error term (u_i) measures inefficiency and the second error term (v_i) measures environmental shocks. For example, a fire that damages a factory and reduces output is not inefficiency, and its effect is captured by v_i. The stochastic production frontier is $y_i = f(x_i, b)TE_i.exp(v_i)$, and approximately $lny_i = lnf(x_i, b) - u_i + v_i$. [Note: by definition $e_i = v_i - u_i$ for production function, where e_i is the composed error]. Aigner, Lovell and Schmidt (1977) formulate the composed error as normal-half normal, and estimate the stochastic frontier model using cross section data under three assumptions: the noise v_i is distributed as normal distribution with mean zero and variance σ_v [i.e. $v \sim iid\ N(0, \sigma_v^2)$]; the inefficiency term u_i is distributed as half normal on the positive side only [i.e. $u \sim iid\ N^+(0, \sigma_u^2)$]; the error terms u_i, v_i and parameters b_i are independent. Weinstein (1964) derives the maximum likelihood estimator for normal-half normal. The normal probability density function is $z(v) = \frac{1}{\sqrt{2\pi}.\sigma_v}\exp(\frac{-v^2}{2\sigma_v^2})$ and half-normal probability density function is $z(u) = \frac{2}{\sqrt{2\pi}.\sigma_u}\exp(\frac{-u^2}{2\sigma_u^2})$. Given that u_i and v_i are independent, the joint density becomes $z(u,v)=\frac{2}{2\pi.\sigma_v\sigma_u}\exp(\frac{-v^2}{2\sigma_v^2}-\frac{u^2}{2\sigma_u^2})$. Re-writing the joint density using *e*= *v-u* obtains $z(u,e)=\frac{2}{2\pi.\sigma_v\sigma_u}\exp(\frac{-u^2}{2\sigma_u^2}-\frac{(e+u)^2}{2\sigma_v^2})$. Integrating out *u* to obtain the marginal density, $z(e)=\int_0^\infty z(u,v)du=\frac{2}{\sqrt{2\pi}.\sigma}[1-\Phi(\frac{e\lambda}{\sigma})].\exp(-\frac{e^2}{2\sigma^2})=\frac{2}{\sigma}.\phi(\frac{e}{\sigma}).\Phi(-\frac{e\lambda}{\sigma})$, where $\sigma^2=\sigma_u^2+\sigma_v^2$, $\lambda=\sigma_u/\sigma_v$, $\Phi(.)$ and $\phi(.)$ are respectively the cumulative density function and probability density function for the standard normal[23]. The normal error dominates the half-normal error when λ approaches zero. We get the case where most firms are efficient, and the parameter estimates from ordinary least square are approximately correct. The half-normal error dominates the normal error when λ approaches infinity. We get the deterministic frontier without statistical noise. The marginal density function *z(e)* is asymmetrically distributed with expected value $E(e) = -E(u) = -\sigma_u.\sqrt{2/\pi}$ and variance $var(e)=\frac{\pi-2}{\pi}\sigma_u^2+\sigma_v^2$.

The next step after obtaining the marginal density of the composed error, *z*(*e*), is to obtain its log likelihood function $lnL = constant-I.ln\sigma+\Sigma ln\Phi(e\lambda/\sigma)-1/2\sigma^2\Sigma e_i^2$ for *I* producers. We can maximize the log likelihood function for the respective parameters b_i to obtain its maximum likelihood estimators, and then recover the composed error e_i for each firm. The last step is to decompose e_i into the two error terms. Jondrow, Lovell, Materov and Schmidt (1982) show that if the efficiency term is half-normal, i.e. $u_i \sim N^+(0,\sigma_u^2)$, the distribution of the efficiency term u_i conditional on the composed error e_i, i.e. $(u \mid e)$ is: $z(u\mid e)=\frac{z(u,e)}{z(e)}=\frac{1}{\sqrt{2\pi}\sigma_*}\exp(\frac{-(u-\mu_*)^2}{2\sigma_*^2})/[1-\Phi(\frac{-\mu_*}{\sigma_*})]$ where, $\mu_*=-e\sigma_u^2/\sigma^2$ and $\sigma_*=\sigma_u\sigma_v/\sigma$; We can use either the mean $E(u_i \mid e_i)$ or the mode $M(u_i \mid e_i)$ to estimate u_i as follows:

$$E(u_i \mid e_i)=\mu_{*_i}+\sigma_{*_i}[\frac{\phi(-\mu_{*_i}/\sigma_*)}{1-\Phi(-\mu_{*_i}/\sigma_*)}]=\sigma_*[\frac{\phi(e_i\lambda/\sigma)}{1-\Phi(e_i\lambda/\sigma)}-(\frac{e_i\lambda}{\sigma})] \textbf{ OR}$$

$$M(u_i \mid e_i) = -e_i(\sigma_u^2/\sigma^2) \textit{ if } e_i\leq 0\textit{, and 0 otherwise}$$

Battese and Coelli (1988) propose a more accurate point estimator for TE_i:

$$TE_i=E(\exp\{-u_i \mid e_i\})=[\frac{1-\Phi(\sigma_*-\mu_{*_i}/\sigma_*)}{1-\Phi(-\mu_{*_i}/\sigma_*)}].\exp\{-\mu_{*_i}+\tfrac{1}{2}\sigma_*\}$$

[23] Standard normal refers to the normal distribution with mean zero and variance of one.

This estimator does not rely on the $u_i = -lnTE \approx 1\text{-}TE_i$ approximation and produces more accurate result when u_i is not close to zero. However, both estimators will not converge towards population mean over large number of observation (econometrically inconsistent) because the variation associated with $(u_i \mid e_i)$ is independent of i. This is the limitation of cross-section data. Greene (1997) argues that using different distributions (such as normal-exponential, normal-truncated normal, normal-gamma) will not have as much value as moving towards panel data estimation.

3.2 Production Frontier and Panel Data

Schmidt and Sickles (1984) state that the stochastic production frontier model using cross section data suffers from three related problems. First, we must specify the distribution of the error terms before using the maximum likelihood method. Second, we need to assume the error terms (u_i and v_i) and the parameters (b_i) are independent. However, it is likely that firms which are aware of their inefficiency (u_i) will change input choices, and hence the regression parameters. Therefore, the error terms and the parameters are unlikely to be independent. Third, technical inefficiency estimated with Jondrow, Lovell, Materov and Schmidt (1982) method is not asymptotically efficient since the variance of the mean $E(u_i \mid e_i)$ or the mode $M(u_i \mid e_i)$ does not converge when the sample sizes increase.

A panel data is a set of observations of I firms over T periods, although there is no need for all firms to be observed in each period. When panel data are available, many new techniques become feasible. First, repeated observations can substitute for distributional assumption in panel data. Second, some of these new techniques do not require the efficiency and the parameters to be independent. Finally, the estimated technical efficiency converges when the number of periods increases. However, this benefit is small since panels must be short for technical efficiency to be time invariant. Dor (1994) gives three advantages of panel data over cross-section data: First, panel data are less likely to introduce omitted variable bias. Second, panel data techniques need less distribution assumption on the efficiency term unless we use maximum likelihood method for estimation. Third, these techniques allow direct test for output endogeneity.

The basis of the panel data approach is the association of firm effect in the panel data literature with one-sided inefficiency term from the efficiency frontier literature. If technical efficiency is time invariant, i.e. $lny_i = lnf(x_{it}, b_i) - u_i + v_{it}$, we can use panel data techniques for fixed and random effects to estimate inefficiency. In the fixed effect model, we assume $u_i \geq 0$, the noise v_{it} is normally distributed, and v_{it} and parameters b_i are independent. We need not make any assumption on the distribution of inefficiency term u_i and its independence with parameters b_i or noise v_{it}. We can recover the inefficiency term from the variable intercept a_i for each firm by defining[24] $\hat{u}_i = \max(\hat{a}_i) - \hat{a}_i$. The fixed effect model is the least square dummy variable model in the panel data literature. This method requires at least one firm to be 100% efficient. The fixed effect model has three drawbacks. First, we need to assume there is no selective environmental shock (e.g. a new law affecting only some firms) because u_i captures all the time-invariant effects across firms. Effects of selective environmental shocks will be incorrectly captured as inefficiency. Second, the fixed effect model consumes one degree of freedom for each firm effect. Third, the parameter estimates do not converge to the population mean in the fixed effect model for short time series with period T, although it is still $\sqrt{T}$ times better than cross-section data. Increasing the period decreases validity of the assumption that u_i is time invariant.

The problem of confounding environmental shock with inefficiency in the fixed effect model motivates researcher to formulate the random effect model. In the random effect model, we assume the inefficiency term u_i is a positive random variable (i.e. $u_i \geq 0$) with constant variance and independent of the noise v_{it} and the parameters b_i. [Note: Unlike the random effect model, we need not assume u_i and b_i are independent in the fixed effect model]. We assume noise v_{it} is normal as usual. Starting from the production frontier

[24] Following the convention in econometric literature, the 'hat' (or circumflex) terms indicate estimated terms.

$lny_i=a_0+lnf(x_i,b_i)-u_i+v_{it}$, we rewrite the terms[25] as $lny_i=(b_0-E[u_i])-lnf(x_i,b_i)-(u_i-E[u_i])+v_{it} = a_0^*+lnf(x_i,b_i)-u_i^*+v_{it}$, which fits the one-way error component model in the panel data literature. We can then use two-step generalized least square or maximum likelihood to estimate the function. For the generalized least square method, we begin by using ordinary least square to estimate the parameters b_i. We then re-estimate a_0 and b_i using feasible generalized least square, recover u_i^* and normalize the efficiency term using $\hat{u}_i = \max(\hat{u}_i^*) - \hat{u}_i^*$. The estimates converge to the population mean when the number of firms (I) or number of periods (T) becomes large. The generalized least square is suitable when there is a large number of firms (I is large), and u_i is uncorrelated with the parameters b_i. Hausman and Taylor (1981) develop a test to check if the variance of the inefficiency (σ_u^2) is uncorrelated with the parameters b_i by using the Hausman (1978) test of significant for the fixed effect estimator. The maximum likelihood method for panel data is similar to the cross-section data method. Pitt and Lee (1981) illustrate the maximum likelihood method for the normal-half normal case. The random effect model has two drawbacks. First, we need to assume firm inefficiency is independent of input level (i.e. firm size). Second, we need to assume a distribution when using maximum likelihood method, thereby introducing the risk of specification error.

Recent research in panel data for stochastic frontier focuses on relaxing the assumption on time invariant efficiency (Cornwell, Schmidt and Sickles, 1990; Kumbhakar, 1990; Battese and Coelli, 1992; Lee and Schmidt, 1993). In particular, Lee and Schmidt (1993) re-specify the intercept term as $a_{it}=\theta_t\delta_t$ where δ_t is firm specific effect and θ_t is an estimable parameter. Note that the firm effect varies with time in this specification. Many of these new models are non-linear and complex. The main advantage of these complex models is econometric consistency (i.e. the convergence of parameter estimates towards population mean over a large sample) in fixed and random effect models.

[25] The 'star' terms replace the corresponding terms in the bracket, e.g. $a^* = (b_0 - E[u_i])$

3.3 From Production to Cost Frontier

The concepts in production frontier provide the foundation for discussing cost frontier. There are five important differences between estimating production and cost frontiers. First, the data requirements are different. Estimating production frontier requires only input and output quantities; estimating cost frontier requires output quantities, cost, input expenditures and possibly input quantity or cost share.[26] In general, we can obtain the data for cost frontier more easily than for production frontier. Second, the cost frontier can accommodate multi-product technology directly, while the production frontier requires distance functions. Third, we can accommodate fix inputs in the variable cost frontier but cannot distinguish which inputs are fixed in the production frontier. Fourth, the production frontier has no behavioral assumption and only measures technical efficiency. The cost minimizing assumption is applicable in a competitive environment when the input price (rather than quantity) is exogenous; the output is demand driven and therefore exogenous. In the service industries where output cannot be stored, output-oriented technical efficiency is not meaningful. Lastly, we can decompose cost efficiency into input-oriented technical efficiency and allocative efficiency.

The simplest cost frontier model is the one that uses single equation for cross section data. The Cobb Douglas cost function is $lnc_i = a_0+b_ylny_i+\Sigma b_nlnw_{ni}+v_i+u_i$. Imposing linear homogeneity in input prices[27] to conform to economic theory, and $\Sigma b_i = 1$ for Cobb Douglas form, we obtain $ln(c_i /w_{ki})= a_0+b_ylny +\Sigma b_nln(w_{ni}/w_{ki})+v_i+u_{i.}$ We can estimate this cost frontier by using the methods for production frontier we have described earlier: For example, using Jondrow, Lovell, Materov and Schmdt's (1982) maximum likelihood method, or Battese and Coelli's (1988) exact estimator, for the normal-half normanl composed errors. In fact, the Cobb Douglas production frontier and cost frontier are exactly the same apart from changing a few signs. We can change the cost function to a flexible function, such as translog, to accommodate multiproduct technology. However, the flexible functional form often give rise to multicollinearity in single equation-cross

[26] Cost share means the ratio of the expenditure for an input divided by the total cost.
[27] Recall this means $C(y_{i,}\theta w_i;b_i)= \theta C(y_{i,}w_i;b_i)$

section analysis. This results in insignificant parameter estimates, although the estimates remain unbias. Multicollinearity may be problematic for analyzing production structure which require precise parameter estimates (Harvey, 1977). However, it has no effect on the composed errors, and ultimately the efficiciency measure. Therefore, it is still feasible to use flexible functional form for cost efficiency studies. We can use the variable cost function if we wish to accommodate quasi-fixed inputs (see chapter 4 for details). Using flexible functional form and accomodating fixed input affect only the cost kernel and not the composed errors.

The main problem of using single equation-cross section analysis is its inability to decompose cost efficiency into allocative and technical efficiency components. We need a simultaneous equation system to decompose the efficiency components. Another advantage of using an equation system is the improvement in econometric efficiency when estimating a flexible cost frontier. We can invoke the Sheppard's lemma to implement the simultaneous equation system[28] using two methods: First, we can estimate a system of cost frontier and its cost minimising input demand equations. Second, we can estimate a system of natural logarithim of cost frontier and its cost minimising input share equations. Schmidt and Lovell (1979) use the self-dual[29] Cobb Douglas functional form to estimate the allocative and technical efficiency under four distributional assumptions: the noise is normal (i.e. $v \sim iid\ N(0,\sigma_v^2))$; the inefficieny term is half-normal (i.e. $u \sim iid\ N^+(0,\ \sigma_u^2)$); the error vector of input demands (η_i) is normal with zero mean and variance matrix Σ (i.e. $\eta_i = (\eta_{2i}...\eta_{Ni})' \sim iid\ N(0,\Sigma))$; and these errors $v_i,\ u_i,\ \eta_i$ are independent.

We can use the method in Christensen and Greene (1976) to model multiproduct cost frontier by estimating an equation system comprising the cost frontier and N-1 cost share equations (where N is the number of inputs). After deleting (any) one cost share, we add error terms to each equation in the system, i.e. we estimate the cost frontier $lnc = lnC(y_i, w_i;b_i)+u_i+v_i$ jointly with the cost shares $S_{ni}=S(y_i,w_i;b_i)+\eta_i$, where b_i is the regression parameters and η_i is the vector of cost share error term. Depending on the relation

[28] See section 4.3 for this result.

[29] The Sheppard lemma provides duality between cost and production functions. The dual of the Cobb Douglas production is the Cobb Douglas cost function, hence it is self-dual.

between η_i and u_i, the error term u_i can capture the effects of technical, allocative or cost efficiency. We make the same four distribution assumptions as in the Cobb Douglas case, i.e. $v \sim iid\ N(0,\sigma_v^2)$; $u \sim iid\ N^+(0, \sigma_u^2)$); $\eta_i = (\eta_{2i} \ldots \eta_{Ni})' \sim iid\ N(0,\Sigma)$); and the errors v_i, u_i, η_i are independent. Greene (1980) notice an inherent incosistency in these four assumptions. If we assume that η_i represents allocative efficiency, then u_i captures the effects of technical and allocative efficiencies. This means the distribution for u_i depends on η_i because the cost of allocative inefficiency must vary with the extent of allocative inefficiency. While the equation system provides more efficient estimates of the regression parameters b_i, the failure of independence in error terms (i.e. u_i and η_i) lead to econometrically inconsistent estimates (meaning the estimates do not converge to population mean over a large sample). Alternatively, we can assume η_i represents statistical noise just like v_i. We assume no allocative inefficiency in this case and u_i represents only technical inefficiency. Then, the equation system contains no more information than the cost frontier. Including the share equation provides more efficient parameter estimates but introduces bias from assumptions about share equation error term η_i. This dilemma is known as the "Greene's problem" and limits the usefulness of equation system to estimate cost efficiency.

There are several proposals to overcome the Greene's problem. Schmidt (1984) suggests assuming η_i as allocative inefficiency distributed as a normal function, then breaking down the inefficiency term u as the sum of costs of allocative and technical inefficiencies, (i.e. $u=u_T+u_A$), and assuming the technical inefficiency is distributed as half normal, (i.e. $u_T \sim N^+(0,\sigma_T^2)$). Schmidt specifies the cost of allocative inefficiency in terms of the cost share error η_i instead of assuming a distribution for u_A. Specifically, $u_A=\eta' A \eta$, where A is a NxN positive semi-definite matrix. When u_A is zero, η is also zero; η is not zero when u_A is positive, and u_A is positively correlated with the absolute value of η. We can then derive u_A from η and A without making distribution assumption in u_A. Schmidt proposes that $A=D^{1/(N-1)}\Sigma^+$ where D is the product of non-zero eigenvalues of the multivariate covariance matrix Σ, and Σ^+ is the generalized inverse of Σ. With this specification for A, we can use the maximum likelihood method to estimate the cost frontier parameters, the magnitude of allocative efficiency, and the costs of allocative and technical efficiencies. However, maximizing the maximum likelihood estimator is a formidable task because of the sheer complexity. Melfi (1984) and Bauer (1985) simplify A to implement Schmidt's solution. Melfi (1984) assumes A is an identity matrix so that u_A is the sum of square of input errors. However, this specification forces u_A towards zero. Bauer (1985) allows A to be a positive semi-definite matrix whose elements become $N-1$ additional estimable parameters (where N is the number of inputs). In this formulation u_A is the weighted sum of square of the errors from the share equations. We can then estimate the cost frontier system using the maximum likelihood method after making distribution assumptions on v, u_T and η. Kumbhakar (1991) suggests another specification without additional estimable parameters. However, these four modifications are often empirically intractable because of two problems. First, there are many parameters to estimate even if the cost kernel consists of few inputs and outputs. Second, we often estimate the system by imposing additional structure such as restricting A and Σ to be diagonal matrices. However, there is no *a priori* reason to believe that these imperfect models linking allocative efficiency with share equations provide better estimate than ignoring these relationships.

3.4 Cost Frontier, Panel Data and Other Techniques [30]

The disadvantages of cross section data in the production frontier carries over to estimating the cost frontier. We need to impose two types of assumptions when using cross section data. First, we need to impose assumptions about the distributions of error terms to use the maximum likelihood method. Second, we need to assume these errors are independent of the regression parameters. Still, the parameter estimates do not converge towards population means over large sample using the maximum likelihood method in Jondrow, Lovell, Materov and Schmidt (1982). Similar to the case of production frontier, we can overcome these disadvantages by using panel data techniques. We can use the

[30] The first paragraph of this section is basically a summary of section 3.2 because, apart for a few sign changes, the techniques in estimating production and cost frontiers are essentially the same.

fixed effect model to produce econometrically consistent estimates for long time series if cost efficiency is time invariant. We can also overcome the problem of time invariant fixed effects over long time series with more recent techniques, such as Lee and Schmidt (1993), but the model is nonlinear and complex. If we assume cost efficiency is randomly distributed and independent of the regression parameters, we can use the generalized least square or the maximum likelihood method to estimate the random effect model. In fact, the parameter estimates in the random effect model converges to the population means in either long time series or large cross section. However, the assumption that the regression parameters (b_i) are independent of the firm specific random effect is problematic. This is because firms are likely to modify their input choices using knowledge about their inefficiency.

So far, there has been no satisfactory cost frontier specification for simultaneous equation using panel data techniques because of the Greene's problem[31]. There are two possible cases just like in the production frontier. First, we can estimate a self-dual cost funtion using seemingly unrelated regression for the fixed effect model. We can use the maximum likelihood method for the random effect model. Second, we can impose distributional assumptions to estimate a system of cost frontier and its cost shares using the maximum likelihood method (Kumbhakar and Lovell, 2000 p.174). There are two additional approaches to estimate cost efficiency which are increasingly popular in the empirical literature, namely the thick frontier method and the distribution free approach. Berger and Humphrey (1991; 1992) pioneer the former and Berger (1993) pioneers the latter.

3.5 Application to Hospital Efficiency

The stochastic frontier and data envelopment analysis (DEA) techniques are widely used in hospital efficiency analysis[32]. In fact, over half of volume 13(2) of the 1994 issue of

[31] As we have discussed earlier, the Greene's problem relates to the interpretation of the error terms in the cost shares.
[32] Hospital application using stochastic frontier includes Zuckerman, Hadley and Iezzoni (1994), Wagstaff (1989), Linna, Hakkinen and Linnako (1998). Hospital application using DEA include Valdmanis (1990) and Register and Burnings (1987).

the Journal of Health Economics is dedicated to the debate of these techniques in hospital efficiency analysis. Kooreman (1994) highlights the key difference between these two techniques: DEA measures technical efficiency while stochastic frontier measures a combination of technical and allocative efficiencies. Furthermore, we need to treat all the deviations, including any measurement errors, from the frontier as inefficiency in DEA. Newhouse (1994) explains that the variables in hospital cost function are almost always measured with error. A sufficiently large measurement error for a firm near the frontier can shift the frontier and affect the efficiency score of all firms in that segment. The same argument also applies to random shock in the environment. Using the stochastic cost frontier method, Wagstaff (1989) finds that 90% of the cost variation is due to random environmental shock and only 10% due to firm specific cost efficiency. On the other hand, DEA is a nonparametric technique that does not require specifying a distribution for the efficiency term. This avoids the possibility of introducing specification errors. A working paper by Cummins and Zi (1998) compares the efficiency ranking of insurance companies using several variants of DEA and stochastic frontier models. The authors find that efficiency ranking within the varaints of stochastic frontier models are fairly robust, but the robustness decreases rapidly when comparing stochastic frontier-DEA varaints and within the DEA variants.

4. Theory of Hospital Cost Function

Our research question is framed in the context of the U.S. hospital cost containment policy outlined in Chapter 1. We developed our analytical framework in Chapter 2 assuming that we can measure the cost efficiency in hospitals. In Chapter 3, we examined the methods for measuring cost efficiency assuming we can specify a suitable hospital cost function. In this chapter, we focus our attention on cost theory, problems in existing empirical research of hospital cost function, and measuring hospital inputs and outputs.

4.1 A Review of Production and Cost Theory

The historical development of neoclassical production theory precedes cost theory. Ferguson (1969), Fuss and McFadden (1978), Naidiri (1982), Diewert (1974; 1982) and Chambers (1988) have provided comprehensive reviews which show the close links between production and cost theories. In fact, it is conventional to discuss production theory before examining cost theory. The production function exists because we assume a stable relation between inputs and outputs governed by physical laws. Therefore, a production function describes the technology which transforms inputs into outputs. The various strands of production theory are largely scattered until the consolidation into a coherent theory by Hotelling (1932), Hicks (1932; 1939), Carlson (1939), Sameulson (1947) and Frisch (1965). There are several maintained hypotheses in a well-behaved production function. First, we assume individuals are rational and will use more input only if that helps to increase output. Therefore the marginal product of any input x is positive, implying that the production function $f(.)$ is monotonically increasing, i.e. if $x_1 \geq x_2$, then $f(x_1) \geq f(x_2)$. Second, the technology is feasible and convex. Mathematically, feasibility means the set of inputs that can produce an output y, which is called the input requirement set $V(.)$, is closed and non-empty. Convexity means a weighted combination of two feasible inputs is also feasible. These two assumptions on $V(.)$ are critical to guarantee a solution exist for optimization. Third, Cassels (1936) proposes the "law of marginal productivity" in production: the marginal products, and therefore the marginal rates of technical substitutions, diminish as input x increases. This implies that the production function is concave, i.e. for inputs x_1, x_2 and a constant θ, $f(\theta x_1+(1-\theta)x_2) \geq \theta f(x_1)+(1-\theta)f(x_2)$, and the isoquant is convex to the origin in the input space. For a twice-continuously differentiable production function $f(.)$, concavity mathematically means the Hessian[33] is negative semi-definite. Given the earlier assumption that $V(y)$ is a convex set, the production function becomes quasi-concave, i.e. a concave function with a convex upper level set. Finally, there are assumptions about essentiality, meaning inputs are required to produce output [i.e. $f(0\ldots0)=0$ for weak essentiality, and $f(x_1\ldots0\ldots x_i)=0$ for strict essentiality]; range [i.e. f is single valued, positive, finite and real valued], and differentiability [i.e. f is everywhere twice continuous differentiable].

The production function is a useful tool to study elasticity of scale and input substitutions using cross section data, and technical change using time series. Researchers usually impose two more restrictions on homotheticity and separability. A production function is homothetic when it can be represented by the transformation $F(f^*(x))=f(x)$, where $f(x)$ has the usual properties (i.e. monotonicity, feasibility, essentiality and differentiability) and is homogeneous, i.e. $f^*(\theta x)=\theta f^*(x)$. Homothetic production functions have parallel isoquants, i.e. marginal rates of technical substitution are constant along any ray from the origin. The family of homogeneous production functions, which has constant elasticity of scale, is a subset of homothetic functions. We define separability based on how the marginal rate of technical substitutions of two inputs, x_1 and x_2, relates to a third input x_3. Specifically, separability implies that derivative of the marginal rate of technical substitution[34] between x_1 and x_2 on x_3 is zero, i.e. $\partial(f_1^{'} / f_2^{'})/\partial x_3 = 0$. We can distinguish two forms of separability. In weak separability, x_1 and x_2 belong to the same group but x_3 belongs to another group. In strong separability, x_1 and x_2 belong to different groups, and x_3 does not belong to either group.[35] A separable technology is one that can proceed in

[33] The Hessian is the matrix of second order derivatives of the function.

[34] From economic theory, the marginal rate of technical substitution is the ratio of the marginal products, i.e. $MRTS = f'_1/f'_2$.

[35] In set notation, strong separability means $x_1 \in A, x_2 \in B, x_3 \notin A \cup B$; weak separability means $x_1, x_2 \in A, x_3 \notin A$. Therefore, strong separability is a more restrictive condition.

independent stages. Separability allows collecting inputs into similar groups and makes empirical analysis tractable.

The cost function is a mathematical representation of the cost minimizing problem subject to a technological constraint in a price taking multi-market environment. Hicks (1939) lays the foundation of modern cost theory by analyzing the consumer expenditure function. Samuelson (1947) consolidates Hicks' results into the modern cost theory. Shepard (1953) contributes the idea for analyzing well-behaved cost function as a dual of the production function. Hence, it is equally easy to analyze technology (such as scale, inputs substitutions or technical change) by using either production or cost function. However, the econometric implementation of cost and production functions differs in the assumption about exogenous variables. We assume output quantities are endogenous and input prices are exogenous when estimating production function. We assume cost is endogenous while input prices and output quantities are exogenous in the dual cost function. Cost function requires market observable data such as cost, output quantities and input prices. Production function requires input and output quantities. In general, disaggregated input quantities are more difficult to obtain than cost and input prices.

The formal definition for cost function is $C(w,y)= min[w,x:x \in V(y)],\ x \geq 0$, where $C(.)$ is the cost function and w is the vector of input prices. We need two fundamental assumptions for the cost function. The cost function is only meaningful for a feasible technology, i.e. the input requirement set $V(y)$ must be non-empty. In addition, optimized solution exists if and only if $V(y)$ is closed and bounded. Applicable assumptions on production function also apply to cost function. The positive marginal cost assumption means that a cost function is non-decreasing in input price. In other words, increasing input price can never decrease cost. This assumption also implies that the derived demands for inputs are downward sloping and the cost function is concave and continuous (Chambers, 1988 p.53). The diminishing marginal productivity assumption means that the cost function is concave. The essentiality assumption means that the cost to produce positive output is positive under positive input prices. In other words, it is impossible to produce positive output at zero cost. In weak essentiality, it is costless to produce zero output. Given that the input requirement set *V(y)* is strictly convex and the cost function is differentiable, Shepard (1953) shows that there is a unique cost minimizing bundle mapping the product space into the cost space given by the relation $\frac{\partial c(w,y)}{\partial w_i} = x_i^*(w,y)$. This result is known as Shepard's lemma. Uzawa (1962) uses the Shepard's lemma and the input requirement set *V(y)* to reconstruct a linearly homogenous [36] cost function. Assumption of homotheticity and separability may be applied as further restrictions in empirical analysis. Chambers (1983; p.111) shows that imposing separability is equivalent to restricting the Allen elasticity of substitution [37]: if w_1 and w_2 are separable from w_3, then x_1 and x_3 can only be Allen substitute if x_2 and x_3 are Allen substitute. Flexible cost functions relax the restrictions on elasticity of substitution.

4.2 Incorporating Multi-product Technology

Incorporating the multi-product nature of hospital production into empirical analysis is the most debated issue in the hospital cost literature. The multi-product case extends the single output case by adding the concepts of jointness, input-output separability, and multi-product returns to scale. Samuelson (1966) defines a joint production technology as one that produces a higher production possibility frontier than the separate production of commodities. Jointness arises when the cost of producing two or more products is less than the cost of producing the same products separately, i.e. $C(y_1 y_2) < C(y_1, 0)+C(0,y_2)$. This concept is related to the economy of scope concept in Baumol, Panzar and Willig (1982). Baumol *et al* show that jointness occurs in cost complementarities[38] or shared fixed cost, e.g. $C(y_1 y_2)=F+y_1+y_2$. Jointness implies economy of scope but the reverse is not always true. Kolsen (1968) states the condition to distinguish joint and common costs.

[36] Recall that linearly homogenous means $C(\theta w,y)=\theta.C(w,y)$

[37] Allen (1938) p.503 defines this elasticity measure as $\sigma_{ij}^{A}(y,w) = \frac{C(y,w).C_{ij}''(y,w)}{C_i'(y,w).C_j'(y,w)}$ where C is the cost function with arguments y for the output and w for the input price vector; C_i' means the derivative on j input and C_{ij}'' the second derivative. An equivalent definitions in the production function is $\sigma = \partial ln(x_i/x_j)/\partial ln(f_i'/f_j')$ where x is the input, f' is the marginal product.

[38] This means the marginal cost of a product falls as the output of another product increases, specifically $\partial^2 C/\partial y_1 \partial y_2 < 0$

Joint costs arise when we make product jointly in technically necessary proportions, and the two alternatives to produce any one of them are not to produce the joint bundle, or to treat the other product as waste. Common costs arise when we make product jointly, and the alternative to producing more (fewer) of any one of them is producing fewer (more) of the others. In a joint production, joint cost arises from fixed proportion and common cost arises from variable proportion. In the fixed proportion case, the general cost function is $C=C(y,w_{1,}w_2)$ where y is the sum of output weighted by their proportions. In the variable proportion case, the cost function becomes $C=C(y,w)+C_1(y_{1,}w)+C_2(y_{2,}w)$. In either case, some costs are product specific and allocable to the product. If production is non-joint, then the multi-product production is simply the sum of separate production functions for each output, and the dual cost function is also additively separable. Total cost is the sum of the costs of producing each product separately. Since all costs are separable, there is neither joint nor common cost.

To discuss input-output separability, we first define an input aggregator function $h(.)$ for grouping inputs, an output aggregator function $g(.)$ for grouping outputs, and a multi-product production function $t(.)$. Mundlak (1963) defines input-output separability as the lack of interaction between inputs and outputs, i.e. $t(y_1,y_2,x_1,x_2)= -g(y_1,y_2)+h(x_1,x_2)$. Laitinen (1980) argues that under the input-output separability restriction, firm can choose its output allocation independent of its input decisions, i.e. we do not need information on factor allocation to obtain the cost function. Hall (1973) shows that the separability restriction limits the cost function to the form $C(y_1,y_2,w_1,w_2)=C(g(y),w_1,w_2)$. The cost function becomes a function of output aggregation $g(.)$ and input prices, and marginal cost becomes independent of input prices.

Hanoch (1970) defines the multi-product return to scale (or homogeneity of degree k) analogously as the single output case for the multi-output function $t(\theta^k y_{1,}\theta^k y_{2,}\theta x_{1,}\theta x_2) =0$. It is not possible to define the overall return to scale in the variable proportion case because an equal proportionate change in all inputs may not result in proportionate change in all outputs. Baumol, Panzar and Willig (1982) show that a multi-output production function which is homogenous of degree k, has a cost function that is homogeneous of degree *1/k*, i.e. $C(\theta y_1, \theta y_2, w_1, w_2) = \theta^{1/k} C(y_1 y_2, w_1, w_2)$. The authors distinguish between the overall return to scale (S_N) and the product specific return to scale (S_i). Baumol *et al* define the overall return to scale as the ratio of total cost to the sum of product of output and marginal cost, i.e. $S_N = C(.)/\Sigma(y_i C_i')$. The product specific return to scale is the ratio of average incremental cost and marginal cost for i^{th} product. Return to scale and economy of scale are different concepts: return to scale is the homogeneity between input and output in the production function, whereas economy of scale refers to the change in the long run average cost as output increases. In other words, return to scale refers to movement along a ray through the origin in the input space, whereas economy of scale refers to movement along long run cost curve in the cost-output space. Solberg (1982) shows that the two concepts coincide only in homothetic technology where isoquants are parallel. Input substitution is constant along the expansion path, which is a ray through the origin.

If we impose non-jointness and input-output separability on multi-product cost function, the function becomes multiplicative between input price and output aggregator. The output aggregator is the sum of separate output functions, each of which depends only on the output of one product. If we impose input-output separability and overall constant return to scale on the multi-product cost function, the function becomes multiplicative between input price and output aggregator, but the output aggregator is not constrained. If we impose non-jointness and overall constant return to scale on multi-product cost function, the function is a linear sum of each output quantity weighted by marginal costs of producing each output. The marginal cost of each output may vary with output quantity, and allow for product specific economy/diseconomy of scale. If we impose non-jointness, input-output separability and overall constant return to scale, the multi-product cost function becomes a linear sum of each output quantity weighted by constant marginal costs of producing each output. In this case, there will be no product specific economy of scale. Fuss, McFadden and Mundlak (1978) state that it is important to model multi-output technology with flexible production/cost function where assumptions about jointness, input-output separability and returns to scale are testable hypotheses. However, there is a tradeoff between parameter parsimony and flexibility. This is a particularly serious problem in hospital cost analysis due to the large number of outputs. This section provides the foundation to discuss research issues in functional form specification in the next section.

4.3 Functional Forms

The issue of functional form arises when we wish to use the stochastic frontier method to measure efficiency. This is because the frontier requires a cost or production function as pre-requisite. The purpose of a production or cost function is to characterize a production structure (e.g. scale, input substitutions and technical change). Hence, the purpose of applied cost analysis is to measure the function's value, its gradient vector (i.e. derived demand) and the Hessian (i.e. the matrix for substitution elasticity). The purpose of stochastic frontier is to evaluate the composed errors (that capture efficiency), and not about evaluating production structure. However, there is a need for analysis not to use restrictive functional form that result in specification error. This is because specification errors are captured in the composed errors.

It is useful to examine the historical development of the cost function to understand how flexible forms arise. Cobb and Douglas (1928) derive the first closed form production function to test the theory of marginal productivity. Specifically, they propose the production function $lny = lnA+a_K lnK+a_L lnL,$ where A, K, L and a_i are respectively the total factor productivity, the aggregate quantities for capital and labor, and the regression parameters. Assuming constant return to scale means $a_K+a_L= 1$. Taking partial derivative of y with respect to the inputs K and L gives the respective marginal products. Profit maximizing firms in a competitive market will choose inputs such that marginal products equal real prices, for example $f_K'=w_K/p$. The authors show that the cost share for capital is $w_K K/p.y=a_K$ under constant return to scale. Therefore, the cost share for capital, and similarly for labor, is the regression parameter obtained by ordinary least square. The Cobb Douglas function is monotonic, concave, separable and homogeneous of degree one in inputs. The Allen elasticity of substitution, defined as $\sigma=\partial ln(x_i/x_j)/\partial ln(f_i'/f_j')$

where x is the input, f' the marginal product, is restricted to one because of homogeneity and strong separability. We can derive the reduced form Cobb Douglas cost function from first order conditions of the production function.[39]

The Cobb Douglas function is not suitable for studying capital-labor substitution because the elasticity of substitution is restricted to one. Arrow, Chenery, Minhas and Solow (1961) introduce the CES (constant elasticity of substitution) function where the elasticity of substitution (σ) is constant but not restricted to one. The CES production function is $y=A(\delta K^{-\rho}+(1-\delta)L^{-\rho})^{-1/\rho}$ and is obtained from double integration of the elasticity of scale assuming constant return to scale. The authors assume that the elasticity of substitution (σ) varies across industries but is invariant within industry. We can obtain the elasticity by three methods using the CES function, namely by labor productivity, capital productivity and capital-labor ratio. However, the survey of empirical results from Nerlove (1967) and Berndt (1976) show that elasticity (σ) is close to one using labor productivity but significantly less than one for the other two measures. This is because capital adjusts slowly relative to labor and does not achieve long run equilibrium instantly. Hence, some researchers increasingly use variable cost function which distinguishes variable inputs from quasi-fixed inputs. We need to use quantity instead of price for the quasi-fix input in the variable cost function. The CES function is monotonic, concave, quasi-additive and homothetic. The Cobb Douglas function is a special case of the CES where the elasticity of substitution is one and the Leontief function (Leontief 1936) is a limiting case where the elasticity of substitution is zero. Uzawa (1962) generalizes the CES function to include factors other than capital and labor, Sato (1967) models more disaggregated substitutions[40] with the two-level CES function. Nevertheless, the CES function and its variants still impose *a priori* restrictions on homotheticity and separability.

The motivation for the flexible functional form arises from the need to impose as few *a priori* restrictions on the technology as possible to avoid potential specification error. Diewert (1973) proposes the concept of flexible function which Blackorby, Primont and

[39] The reduced form Cobb Douglas cost function is $C=(\alpha+\beta)(y.w_K^{\alpha}w_L^{\beta})^{1/(\alpha+\beta)}(\alpha^{\alpha}\beta^{\beta})^{-1/(\alpha+\beta)}$

[40] Sato modeled substitutions among skilled labor, unskilled labor, structural capital and machines.

Russell (1978) later refine. A functional form is flexible if we can choose parameters of the function to make its first and second derivatives equal to corresponding derivatives of the estimated function at any point in the domain[41]. Diewert (1971) derives the generalized Leontief production function using the second order Taylor expansion of the production function $y=\sqrt{x}$. The generalized Leontief function is the first flexible functional form.

Christensen, Jorgenson and Lau (1973) introduce the translog function that allows empirically testing the homotheticity and separability restrictions as hypotheses. The quadratic nature of the translog function ensures regularities (i.e. monotonicity and convexity) at least locally. Blackorby and Diewert (1979) demonstrate the duality of the translog cost and production functions, and Chambers (1988) p.169 demonstrates the duality of the profit and production functions. The translog function has become the most popular flexible function in empirical analysis because it has the least number of estimable parameters. The translog cost function with i inputs and m outputs is: [42]

$$lnC = a_0+\Sigma a_i lnw_i+1/2\Sigma\Sigma a_{ij} lnw_i lnw_j+\Sigma b_m lny_m+1/2\Sigma\Sigma b_{mn} lny_m lny_n+\Sigma g_{im} lnw_i lny_m$$

We need to impose two restrictions on the cost function to conform to economic theory. First, the symmetry condition requires $a_{ij}=a_{ji}$, $b_{mn}=b_{nm}$ and $g_{im}=g_{mi}$. Second, the cost function needs to be homogenous of degree one in input prices, i.e. $\Sigma a_i=1$, $\Sigma a_{ij} lnw_j=0$ and $\Sigma g_{in} lny_n=0$. We can also impose additional restrictions to the underlying technology. For homotheticity, the sufficient and necessary condition is $g_{in}=0$. To be homogeneous of degree n in output, we need to impose homotheticity and $b_{mn}=0$ (the function becomes homogenous of degree $1/\Sigma b_n$). To have constant return to scale, we need homogeneity and $\Sigma b_n=1$. If we have constant return to scale and $a_{ij}=0$, we are back to the Cobb Douglas function.

[41] For the n-dimension input vector, there are n marginal products and $0.5n(n-1)$ symmetrical Hessian terms. Adding the value of the function, there are $0.5(n+1)(n+2)$ effects to estimate.

[42] Note that we use i and j for inputs and m and n for outputs, the parameters are a_0 for the intercept, a_i for the input parameters, b_i for the output parameters, g_i for the parameters containing both inputs and outputs. This is explained in Appendix A.

However, there are two main disadvantages of the translog function. First, Brown, Caves and Christensen (1979) show that the *m*-output *n*-input translog function contains *(m+n)(m+n+1)/2* estimable parameters. There is a need to tradeoff feasible econometric estimation with aggregation of inputs/outputs. Second, the translog form cannot take zero value that often occurs in disaggregated data. Caves, Christensen and Tretheway (1980) generalize the translog form by applying Box-Cox transformation. The transformation $\lim_{\lambda \to 0}(Y_i^{\lambda} - 1)/\lambda = \ln Y_i^2$ changes zero outputs to non-zero. However, the transformation complicates estimation especially in stochastic frontier.

We can directly estimate the translog cost function, but multicollinearity can easily arise from a large number of parameters. Christensen and Greene (1976) suggest a more (econometrically) efficient way. They use Sheppard's lemma to obtain cost minimizing input demand functions and then derive the cost shares. They then jointly estimate the translog cost function with cost share, $S_i = \frac{\partial \ln C}{\partial \ln w_i} = \frac{w_i}{C}\frac{\partial C}{\partial w_i} = \frac{w_i x_i}{C} = a_i + \sum a_{ij} \ln w_j + \sum g_{mn} \ln y_n$. We need to specify a stochastic framework for the share equations to estimate the system. Typically, we add an error term (η_i) to each cost share assuming that the error vector is multivariate normal with zero mean vector and constant covariance matrix[43]. However, we wish to estimate the cost function jointly with the cost shares. Since the sum of cost share is one by definition, i.e. $\Sigma b_i = 1$, we need to delete one share equation so that all the equations are independent, i.e. deleting one equation will prevent singular covariance matrix. We can use seemingly unrelated regression to estimate the system. This procedure produces the maximum likelihood estimator (Kmenta and Gilbert, 1968). The parameter estimates are the same whichever share equation is deleted (Barten, 1969).

[43] This is the case when we are interested in the regression parameters. Recall that in the stochastic frontier, assumptions related to the error term η_i and the error terms in the cost frontier is the root of the Greene's problem.

4.4 Hospital Cost Function

Feldstein (1967) pioneers the use of econometric techniques in hospital cost analysis. There are usually two types of motivations for hospital cost analysis: the first is to obtain the characteristics of technology such as scale effects and input substitution; the second is to study the relationship between ownership and cost efficiency. We can also view these two streams of research in an econometric sense: the former is interested in the regression parameters, and the latter is interested in the composed errors.[44] Breyer (1987) highlights three approaches in specifying the hospital cost function: the *ad hoc* or behavioral approach, the structural approach, and the hybrid approach.

Earlier studies tend to use the *ad hoc* specification to explain unit cost[45] among hospitals. Total cost is rarely used in these studies because of three disadvantages: heteroscedastic residue, multicollinearity, and the coefficient of determination (R^2) depending largely on the hospital bed size. Evans (1971) argues that there is a systematic difference between observed and minimum costs in hospitals because the behavior of its agents is influenced by incentive. Therefore, statistical analysis of hospital cost in the real world should include material variables even though these variables do not have any clear role in cost theory. The most common regressors are capacity (bed size), activity (case flow rate, occupancy or average length of stay), casemix (proportion of patients in each category), wage level (proxy for input prices), and dummy variables for teaching status, hospital facilities and characteristics of the market. Authors using the *ad hoc* specification often omit variables that are difficult to measure. Given the heterogeneity of these variables, we cannot easily specify a suitable functional form nor interpret the regression parameters. Most of these studies use a linear additive form which has implicit restrictions, e.g. an additional patient day raises cost by a fixed amount independent of the casemix. These restrictions are not realistic. Cowing, Holtmann and Powers (1983) argue that the problem of the *ad hoc* cost function is the lack of theoretical foundation.

[44] Although we can measure cost efficiency using Data Envelopment Analysis (which is a nonparametric method), only the stochastic cost frontier need a cost function.
[45] Unit cost includes ‘cost per case’ or ‘cost per day’.

Conrad and Strauss (1983) and Cowing and Holtmann (1983) provide the earliest studies which specify a cost function using the cost theory we have described in the previous three sections. These cost functions contain only input prices and output quantity as arguments. If we assume economic agent always minimizes cost, the cost function becomes the dual of the production function through Shepard's lemma. While this approach has a strong theoretical foundation, there is a risk of omitting variables that affect cost. These variables include the hospital's teaching status, ownership and care quality.

Granneman, Brown and Pauly (1986) seek a compromise between the *ad hoc* and structural functions by introducing the hybrid function. The authors modify a translog cost function to obtain $lnC=\phi\underline{X}+\Sigma\alpha_i ln\underline{P_i}+f(\underline{Y,D,CM,R,Z})+\varepsilon$. Here, $\underline{X}$ is a vector of variables affecting cost, $\underline{P}$ is a vector of input prices, $\underline{Y}$ and $\underline{D}$ are vectors of outputs (being inpatient day and discharge cases respectively), $\underline{CM}$ is a vector of casemix variables, $\underline{R}$ is a vector of the sources of revenue, $\underline{Z}$ is a vector of other outputs, and ε is the error term. The authors assume the vector $\underline{X}$ affects cost level but not the shape of the cost function. After imposing further restrictions, the cost curve becomes locally concave, continuous, non-decreasing and homogenous of degree one in input prices. The authors estimate the hybrid function by using ordinary least square.

The hybrid cost function is an uncomfortable compromise. It is basically a structural cost function which includes additional variables. These additional variables are restricted so that the estimation of scale and input substitutions elasticity are not affected. We shall adopt the structural cost function approach in our empirical strategy in Chapter 5 and treat these additional variables in two ways. First, we will apply theoretical argument to incorporate care quality as a joint proxy for output. Second, we will eliminate teaching hospital because we cannot disentangle the effects of patient care and clinical teaching. The remaining part of this chapter will focus on specific topics related to the formulation of hospital cost function. First, we will discuss the issues related to hospital patient care output, especially the issue of casemix aggregation and care quality. We shall introduce the idea that care quality should be one of two proxies for hospital care and legitimately enters a structural cost function. Second, we shall discuss the issues related to hospital inputs, paying special attention to the boundary of hospital. Finally, we shall argue that teaching hospitals produce teaching and patient care with joint technology. Therefore, eliminating these hospitals improve the internal validity of this research.

4.5 Issues Relating to Hospital Outputs

In the neoclassical economic theory, a firm produces products (or services) which consumer buys. We use Grossman and Hart (1986) approach of defining the firm boundary based on asset ownership. The consumer is willing to pay a price equal or above the firm's cost (depending on the market structure) because the products increase the consumer's utility. In this sense, it is difficult to view the hospital's products as patient-day and availability of medical treatment such as pharmaceutics and surgery. The consumer has negative marginal utilities for these services. The consumer's utility arises from the protection or recovery of her health capital in the sense of Grossman (1972).[46] The physician coordinates the production of hospital output (protection or restoration of patient's health) by ordering suitable medical services as information unfold as described in Harris (1977). The output is more like a tailored suit than a standardized widget. Hospital services are like the odd shaped clothes that need to be sewn together for each individual customer. Our earlier review of the U.S. hospital industry in section 1.1 has shown that hospitals use halfway technology which is the most expensive part of the episode of disease. Arrow (1963) argues that medical care is characterized by product uncertainty where there is a great scope for "clinical judgments" by physicians.[47] We expect product uncertainty to be especially severe for halfway technology. There are great variations in the outcome of the same disease treated within acceptable medical practice. In the context of nonprofit segregate hospitals, Newhouse (1970) argues that

[46] Health capital is affected by many factors outside the influence of medical service (including hospital services). These other factors include "investments" (like medical care, diet and exercise) and behavioral modifiers (like wage rate, education level and price level of home produced goods).
[47] In fact Arrow explained that the scope of clinical judgments meant that patient delegate the consumption decision to physician.

quality should be part of the objective function of hospital decision makers. The board of trustees has a taste for care quality as philanthropists. The physicians prefer high quality because of professional expectation (Arrow, 1963) and competition for patients. The manager has a taste for prestige which depends on quality, and is also not evaluated on profit. It is therefore sensible to measure hospital care using joint proxies, hospital services and care quality, and not any one alone. The primary issue about hospital cost containment is about controlling the cost of patient care without compromising care quality.

4.5.1 Aggregating Hospital Services

The hospital is a multi-product firm even if we limit ourselves to patient care services. All hospitals deliver patient care as inpatient service,[48] with some hospitals also delivering outpatient care. Hospital labor and capital are used to produce diagnostic, therapeutic and hotel services, which are really intermediate products ordered by the physicians. We can use admission or patient day to measure the hospital's output. Butler (1995) argues that case is a better unit of measure than patient day as follows: The total cost per admission is the sum of the fixed cost (such as administrative cost for admission and discharge), the variable cost (mainly the hotel component) and the treatment cost. Within the limits of medical technology, it is possible to the treat patient with high intensity-short stay strategy or low intensity-long stay strategy. The former has higher cost per patient day but lower total cost due to lower variable cost. If the hospital output is the protection or recovery of health capital, then the patient day measure is a problem. This is because the hospital cost should vary with the treatment and not the hotel component. Using admission avoids this problem.

Using admission alone is not satisfactory because the casemix proportion varies among hospitals. Barer (1982) outlines three approaches to standardize casemix variation among hospitals: (i) grouping cases without weighting, (ii) grouping cases with intra-group weighting, and (iii) grouping cases and total weighting. Many cost studies such as Granneman *et al* (1986) adopt the first strategy to group cases by medical specialty.

[48] The distinction is inpatient service requires hospital stay while outpatient service does not.

These authors use fractions of cases by specialty as outputs. There are two disadvantages in doing this: First, there is an implicit assumption to weigh each specialty group equally. Second, there is a tradeoff between getting homogenous group and tractable number of groups. This is especially true for the flexible cost function as the number of estimable parameters increases rapidly with the number of groups. The second strategy involves deriving weights for each group using a data reduction technique such as factor analysis or principal component analysis. The third strategy assumes equal weights within each group and uses group specific weights to aggregate the proportion of cases falling within each group into a scalar casemix. In all three cases, we weigh heterogeneous inpatient output with the market value of treated case, i.e. using the price patient pays and not the cost. Three prominent casemix measures have dominated the literature: the International Classification of Disease (ICD), diagnostic related group (DRG) and Resource Need Index (RNI).

The ICD codes[49] are derived from the Manual of International Statistical Classification of Diseases, Injuries and Cause of Death published by the World Health Organization. This is an important starting point for classifying hospital output because the codes are mutually exclusive and exhaustive. However, the large number of groups limits the usefulness of ICD because the hospital outputs contain many cells. On the other hand, the lack of severity measure increases the chance of case variation. The ICD originates from the International Statistic Institute's international list of the causes of death in 1893, and is revised by the Institute every decade. In 1948, the Institute extended the list (at the 6th revision) to include non-fatal conditions and recommended its use for mortality and morbidity statistics. In 1979, the ICD9CM codes[50] became the sole classification system for morbidity reporting in the U.S. The Healthcare Financing Authority (HCFA) took over the annual maintenance and update of the codes in 1985. More importantly, the ICD codes (especially ICD9CM) have become the standard tool for data collection, quality monitoring, research and reimbursement in the U.S.

[49] ICD stands for the International Classification of Diseases

[50] ICD9CM stands for the International Classification of Disease 9th Revision – Clinical Modification. It is a variation of the ICD. The U.S. is the main user for ICD9CM, while Australia and Israel also use it for morbidity reporting.

The original codes for the diagnostic related group (DRG) are derived from ICDA8 (ICD adapted for America, version 8). The first version of the DRG for Medicare-Medicaid[51] Prospective Payment System (PPS) in 1983 is developed in the paper by Fetter, Shin, Freeman, Averill and Thompson (1980). The primary objective of the DRG is to group cases which use similar resources and hence should receive the same reimbursement from the Medicare-Medicaid program. Fetter *et al* initially divide all cases into 23 major disease classes (MDC) to simplify the analysis over diverse diseases in the acute care setting. These early divisions ensure the bundling of medically meaningful groups.[52] Medical meaningfulness is important to ensure the homogeneity in the treatment provided in each group. The DRG is economically meaningful in reducing the variance in hospital cost down the hierarchy from the surgical medical dichotomy, to major disease class (MDC), and finally to DRG. In moving down this hierarchy, clinicians propose a list of variables to group patients. Fetter *et al* calculate the (cost) variance reduction that each variable can produce. The authors select the variable producing the greatest variance reduction as the basis for classification of the groups. They repeat the process with the next variable until the variance or group size becomes small.

Grimaldi and Micheleti (1982) point out four problems with Fetter's DRG which may cause cost variation. First, some DRG groups depend on diagnosis alone and are independent of the treatment, number of diagnosis, and co-morbidity. Second, DRG does not recognize disease staging and severity. Third, there are some errors classifying certain surgical procedures (done outside the operating theatre) as medical cases. Forth, the primary diagnosis often consumes fewer resources than the complications. Horn and Sharkey (1983) produce the evidence that disease severity is a source of cost variation within DRG. The heterogeneity problem is especially acute in the low volume DRG for pediatric and psychiatric diseases. The dilemma in the DRG scheme is the tradeoff between having fewer groups containing inherent variations within each group, and

[51] These two public insurance schemes are managed by HCFA, and are the largest insurance in the U.S.
[52] Fetter *et al* illustrated that patients diagnosed with hemorrhoid, tonsil hypertrophy and normal pregnancy may consume similar dollar value resource, but these cases are medically unrelated.

having more groups with enough numbers within each group. The DRG cost weights are population specific.

There is a general agreement that casemix exerts a (statistically) significant influence on inter-hospital cost variations (see for example Evans and Walker, 1972; Tatchell, 1977; Hardwick, 1986; Butler, 1995). This influence increases significantly when the casemix scheme is more disaggregated. Butler (1995) highlights the twin problems of losing degree of freedom and multicollinearity when imposing casemix directly on the cost function. A feasible alternative is the use of a scalar casemix index constructed separately from patient data in the sample hospital. Hornbrook (1982a; 1982b) explains that the method consists of a diagnostic classification scheme, a weighting scheme and an aggregation formula. The ICD is the most common classification scheme. Klastorin and Watts (1980) argue that aggregation formulae are usually linear for simplicity because there is no *a priori* support for more complex forms. There are three common scalar casemix indices in the literature: First, Evans and Walker (1972) develop the Information Index based on the theoretical foundation of Theil (1967). Second, Ament (1976a; 1976b) develops the Resource Need Index from 1.8 million admissions in 50 hospitals from the Professional Activities Studies. Fetter *et al* (1980) develop the DRG which Medicare-Medicaid has adopted for hospital reimbursement. Sloan, Feldman and Steinwald (1983) state that several studies (Klastorin and Watts, 1980; Hardwick, 1981; Barer 1981) comparing the performance of these three indices in explaining cost variation produce no consensus on which index is superior. Using the weights from any of these three indices to calculate casemix-adjusted admission is a reasonable summary statistics for hospital services. Using the scalar casemix conserves degree of freedom and makes the estimable equations tractable (especially when a flexible cost function is used). However, the weights in the scalar index are population specific. We should expect deviation if we do not derive the weights from our own sample. Since the Healthcare Financing Authority updates DRG weights for U.S. hospitals annually,[53] using DRG weights has the advantage that these weights are derived from our target population in 1997.

[53] This requirement is enacted in the Social Security Act.

4.5.2 Measuring Care Quality

We have explained the need to include care quality in section 4.5. Hospital care quality is a multi-dimensional construct and some dimensions are not easily measured (Newhouse, 1980). The crux of delivering high care quality is the balance of benefit and harm in treatment (Donabedian, 1980; 1988). Donabedian (1980) proposes the Structure-Process-Outcome approach to measure care quality. Structural measures meter the hospital's capacity to deliver care. Structural measures in the empirical literature include teaching status, accreditation[54], the availability of expensive medical equipment[55], and labor intensity such as nurse per bed. Process measures refer to the treatment protocol which physicians use. For example, some hospitals implement clinical pathways[56] to manage selected diseases such as diabetes. Outcome measures are the desired states resulting from the care processes. Indicators for outcome measure are best developed in the literature and deserve a separate discussion. These three types of quality measures are linked in an underlying framework. Donabedian (1988) argues that good structure promotes good process, and good process in turn promotes good outcome.

The crux of the cost-quality nexus is whether quality increases, decreases, or is independent of hospital cost. Scott, Forrest and Brown (1976) show that care quality is positively correlated with cost. This result is intuitive: if the marginal cost of quality is not positive (and assuming positive marginal utility of quality), a rational economic agent will employ infinite resource to deliver higher quality. However, Neuhauser (1971), Shortell, Becker and Neuhauser (1976), and Longest (1978) report negative correlation between cost and care quality. This result arises because of product uncertainty in the Arrow (1963) sense. Physicians prescribe treatments which have benefits and side effects. An aggressive treatment plan is more costly and some of the treatments actually cause

[54] For example staff accreditation (e.g. Board certification) and hospital accreditation (e.g. Joint Commission of the Accreditation of Hospitals certification).
[55] Magnetic resonance imaging is an example.
[56] The clinical pathway is a standardized process in the "best practice" of medicine. Here is a link that provides some basic information http://www.openclinical.org/clinicalpathways.html. The motivation for clinical pathway is really a mix of the quest for care quality and cost containment.

harm [57] which further increases cost. A conservative treatment plan can benefit from increasing treatment to increase care quality. It is impossible to determine if a treatment is *ex ante* excessive (even if it is clear *ex post*) because of product uncertainty. Garber, Fuchs and Silverman (1984) argue that the cost-quality nexus depends on the clinical condition. Donabedian, Wheeler and Wyszewianski (1982) propose that the cost-quality relation is non-linear and monotonically increasing. Furthermore, the marginal cost of care quality diminishes over output range. Therefore, care quality clearly affects cost through clinical efficiency and must be part of the hospital output in the structural cost function. Given this background, we now examine two outcome measures for care quality:

(A) Risk Adjusted Mortality Index

Unlike most outcome data, mortality rates are readily available, and consistently and unambiguously recorded. Mortality rates, especially those associated with unexpected mortality, represent an adverse care outcome and therefore measure care quality in hospitals. Mortality rates correlate with less drastic failure in care quality such as delayed recovery, residual disability and increased risk of future illness. However, raw mortality rates are not useful because care quality is only one of many factors for hospital mortality.[58] Pollack, Ruttimann and Geston (1987) emphasize the importance of adjusting mortality rates for disease categories because casemix and severity have substantial impact on mortality.

Desharnais *et al* (1988) develop the risk adjusted mortality index using 6 million discharges[59] and validate its reliability and generalizability. The authors divide the cases by diagnostic related group (DRG) cluster[60] into the high mortality group (with raw

---

[57] This is call iatrogenesis in the medical jargon.
[58] The other factors are the patient's physiological attributes (such as age, gender and co-morbidity), social conditions (such as financial situation affecting nutrition and housing; education affecting health behavior and self care), and the type, stage and severity of diseases.
[59] The database was from the Commission on Professional and Hospital Activities, which is now part of the Cecil G Sheps Center for Health Services Research, University of North Carolina, Chapel Hill.
[60] Some DRG were the same disease with modifiers for different age group and co-morbidity. The authors needed to use DRG cluster instead of DRG because age group and co-morbidity are parameters of the logistic regressions.

mortality rate above 5%) and the low mortality group. There are 64 high mortality clusters that accounted for 17% of the discharge and 72% of the mortality. The authors prepare a contingency table for each cluster in the low mortality group and use logistic regression (with nine patient characteristics including age and co-morbidity) for the high mortality group. Desharnais *et al* then compare the predictive power of mortality rates, defined as mean error per cluster, under four specifications: using the crude mortality rate (i.e. no adjustment for casemix), DRG adjusted mortality rate, only contingency tables (i.e. only adjust for casemix, age and co-morbidity), and combination of contingency table and logistic regressions. They find the corresponding mean errors per cluster are 1.08, 0.416, 0.393 and 0.357. This result shows that the full model is the most powerful, but casemix alone accounts for 92% of the improvement in predictive power.

Desharnais *et al* (1988) use in-hospital mortality to determine the risk adjusted mortality index. This measure can become bias if hospital administration is able and willing to transfer pessimistic cases to tertiary (e.g. teaching hospitals) and palliative care centers (e.g. hospices). In-hospital mortality is sensitive to the length of stay (Jenks, William and Kay, 1988). Specifically, hospitals in New York State have longer length of stay than California and 25% higher in-hospital mortality, but 1.6% lower 30-day post hospitalization mortality. Therefore, including mortality within a reasonable post-discharge time window can improve the validity of the measure. The Healthcare Financing Authority uses an arbitrary 30 days post-discharge window for insurance claim. Fleming *et al* (1991) improves the reliability with a DRG adjusted variable time window. However, three issues arise from adding this post-discharge window. First, the window increases the influence of patient's socio-economic characteristics and can be wrongly attributed as hospital care quality. Second, if patient has re-admission to another hospital during this window period, the hospital to assign the case becomes ambiguous. Third, data management becomes more difficult.

(B) Surgical Complication Rates

We need another care quality measure to supplement the risk-adjusted mortality because of two related reasons. First, mortality is an extreme outcome in medical care and not the full spectrum. Second, care quality is a multi-dimensional construct according to Newhouse (1980). Therefore, it is sensible for us to use a supplementary outcome measure such as surgical complication rates. These indicators measure different aspects of care quality. For example, mortality measures quality delivered in intensive care ward (using more uncertain technology) and surgical complication rates measure care quality in surgical units such as theatre and surgical wards. Re-admission rate is another possible indicator, but it is more sensitive to administrative factors such as discharge policy.

Using the healthcare financing authority's co-morbidity and complication list, Desharnais *et al* (1988) designate 70 ICD9CM codes as those likely due to surgical complication. The authors assume a condition as co-morbidity, and not complication, if either is possible. For example, the authors exclude pneumonia from the list because pneumonia may be a primary medical condition or arise from poor post-operative care. This surgical complication rate is a feasible measure although it understates the prevalence of surgical complications.

4.6 Issues Relating to Hospital Inputs

The main issue related to measuring hospital input is the treatment of physician labor. The hospital uses capital and hospital labor as inputs for producing hospital services. If we view the hospital industry as part of a system to deliver medical care, instead of a producer of intermediate products, then physician labor and cost ought to be part of the hospital cost. A hospital cost function without physician is infeasible, i.e. the input requirement set $V(y)$ will be empty. Most U.S. hospitals and physician practices are separate legal entities, and most (if not all) cost analyses use accounting data. This implies that most hospital cost analyses exclude the physician component, and it is rare to see physician input price in the cost function. The trend in physician hospital integration (i.e. hospital hiring physician as employee) adds to this complication as the accounting data from integrated and segregate hospitals are not comparable.

Bays (1980) is a rare paper that include physician inputs in the cost function. The author exploits the fact that hospital record case diagnosis, treatment and procedure by using the ICD codes. [61] He matches the ICD8A code to the relative weights for insurance payment scheme for physician[62] to estimate the (market) value of physician input. Given this information, Bays compares his estimations, with and without the correction, for the average cost function, *AC=AC(Bed Size, Case Flow, Casemix Proportion)*. He finds that the average cost function is quadratic with physician input adjustment and linear otherwise, showing that physician inputs matter. The healthcare financing authority's decision to pay for physician services using the resource based relative value system[63] (RB-RVS) since 1992 means that we have a more refined database than Bays. Therefore, including physician inputs in hospital cost analysis is theoretically mandatory and technically feasible.

4.7 Issues Relating to Teaching Hospital

There is strong evidence that teaching hospitals have higher average cost after adjusting for casemix. [64] Cameron (1985) finds that after adjusting for casemix, university hospitals, major and minor teaching hospitals are respectively 26%, 10% and 8% more costly than non-teaching hospitals. From our review of the hospital industry in section 1.1, hospitals are part of a system to deliver patient care using inputs from the supporting industries such as pharmaceutical industries and professional schools. In this sense, the teaching component is part of the input[65] to produce skilled clinicians and not the output of the

[61] Specifically, Bays used the ICDA8 (International Classification of Diseases, version 8, adopted for US).
[62] Bays used the weight data from the California Relative Value study.
[63] The billable procedures are coded using a modified form of the Current Procedure Terminology version 4 (commonly called CPT-4) developed and maintained by the American Medical Association.
[64] Some examples of cost studies that found higher cost in teaching hospitals are Pauly (1978), Culyer, Wiseman, Drummond and West (1978), Watts and Klastorin (1980), Sloan and Steinwald (1980), Jones (1985) and Cameron (1985).
[65] The production of skilled clinical labor will need to include educational institutions like the professional schools for nurses, physicians and paramedics.

medical industry *per se*. The debate of whether a patient should pay for the teaching component is beyond the scope of this dissertation. However, it is clear that additional teaching cost is not inefficiency.[66] Separating clinical components and teaching is the crux of the issue.

Sloan, Feldman and Steinwald (1983) argue that because patient care, research and clinical teaching are jointly produced. Hence, there is no definite way to allocate time devoted to these activities. Since clinical teaching and patient care are produced in variable proportions, there is a common cost that is not allocable to each product. Hadley (1983) points out that any accounting approach in assigning the cost between teaching and patient care is essentially arbitrary. Therefore, we exclude teaching hospitals from our sample to increase internal validity of this research.

[66] Frick, Martin and Swartz (1985) offer four possible reasons for teaching hospitals to have higher cost: First, they use more ancillary services that increase cost (empirical evidence from Cameron, 1985); second, they attract more severely ill patients; third, they provide better quality (empirical evidence from Becker and Steinwald, 1981); and lastly they are simply less efficient. There is also a counter point that teaching hospitals have higher cost. Hadley (1983) argues that from Becker (1975) theory of general human capital, trainees pay for the clinical teaching by accepting a lower pay than his productivity. There is substantial empirical controversy on this point. For example, Hosek and Palmer (1983) and Robinson and Luft (1985) find no difference, but Garber, Fuchs and Silverman (1984) finds higher cost in faculty physicians.

5. Empirical Strategy

We have argued in chapter 2 that the team agency framework[67] provides a new perspective to investigate the debate whether for-profit hospitals are more cost efficient than nonprofit ones. Three falsifiable hypotheses are formed using the framework:

H1A: Nonprofit FIOs has higher efficiency score than nonprofit segregate hospital
H1B: Nonprofit FIOs has higher efficiency score than nonprofit network hospital
[The above result is due to Holmstrom (1982)]

H2A: For-profit FIOs has similar efficiency score as for-profit segregate hospital
H2B: For-profit FIOs has similar efficiency score as for-profit network hospital
[The above result is due to Eswaran and Kotwal (1984)]

H3A: For-profit network hospital has higher efficiency score than nonprofit ones.
H3B: For-profit segregate hospital has higher efficiency score than nonprofit ones.
[The above result is due to Alchian and Demsetz (1972)[68]]

Kooreman (1994) points out that data envelopment analysis (DEA) can only measure technical efficiency while the stochastic frontier can measure cost efficiency. We have chosen the stochastic frontier because of the need to measure cost efficiency. We have chosen cross section over panel data because of our resource constraint (data are expensive) and our preference for simpler analytical techniques. Although panel data techniques are clearly superior, there are still tradeoffs in each technique. We have seen in Chapter 3 that some panel data techniques for stochastic cost frontier need not require assumptions about the distribution of the inefficiency term. However, the fixed effect model requires the assumption of time invariant efficiency. Accommodating the time

[67] This is mainly the arguments in Holmstrom (1982), Eswaran and Kotwal (1984).
[68] We exclude the for-profit vs. nonprofit comparison of FIOs because team agency and property rights theories predict opposite result. Therefore, we expect the efficiency score between the two groups to be statistically equal. We will discuss this result but not test it out as a formal hypothesis because there is no *a priori* information of the relative size.

varying efficiency in the fixed effect model needs the tradeoff of making error terms complex and nonlinear. The Achilles' heel of the random effect model is the assumption that efficiency and parameters are independent. We assume the composed errors are normal-half normal and use the Battese and Coelli (1998) exact estimator.

Introducing the simultaneous equation system can improve parameter efficiency (i.e. reduce the parameter variances), but the Greene's problem makes decomposing cost efficiency into the technical and allocative components difficult. Since our theoretical model is silent on the source of the cost inefficiency, decomposing into allocative and technical efficiency is not required. Our cost frontier specification has only two cost share equations, and one of them need to be deleted to prevent singular covariance matrix (see section 5.1 for details). Therefore, expect the reduction in parameter variance is small. So, we estimate cost efficiency by the single equation method.

We need to specify a cost function to estimate the stochastic cost frontier. Since we do not have *a priori* information about the production structure, we choose a flexible cost function to avoid specification error. The translog form has the least parameters, but we expect multicollinearity to remain serious. Harvey (1977) explains that multicollinearity makes the parameter estimates less precise (i.e. high variance producing insignificant estimates) but the estimates remain unbiased. Since characterizing the production structure is not our objective, insignificant parameter estimates do not pose a problem. We only need the parameters to be unbiased because they are used in the second stage to partition the composed errors into inefficiency and white noise. Farrar and Glauber (1967) warn that multicollinearity can impart a bias towards incorrect model specification. Since multicollinearity arises because non experimental data tend to be correlated, correcting the problem by using *a priori* information is extremely difficult.

A variable cost function is chosen because capital does not adjust instantly (see section 4.3 for detail). We have argued in section 4.4 that the hospital output is patient care. Measuring patient care requires quantity and quality as joint proxies. We should aggregate hospital service for parameter parsimony. DRG weighting is used because the annual updates ensure the weights are derived from a population similar to our sample. We need at least two care quality measures to accommodate the multi-dimensional aspect of this measure. Donabedian's (1980) structure-process-outcome model of quality shows that quality measures are linked. Two outcome measures, risk adjusted mortality index (RAMI) and surgical complication rate, are used for two reasons: First, our conceptualization of the hospital outputs is closer to care outcome than hospital structure or clinical process. Second, we chose outcome measures that can be measured by a continuous scale. In section 4.6, we have shown the need to include physician labor, in addition to capital and hospital labor, as hospital inputs. The cost of physician labor should also be included in the variable cost. Given this overview, we will now proceed to the specifics of cost frontier specification, variable construction, data sources and software choice.

5.1 Cost Frontier Specification and Programming

We specify the following translog cost frontier where a_0 is the intercept; a_i are the parameters for inputs; b_i are the parameters for outputs; g_{ij} are the parameters for mix input/output; VC is the variable cost; w_i are the parameters for input price, with subscripts $i=L,M$ for hospital and medical labor respectively; y_i are parameters for quantities, with subscripts $i=K,D,A,Y$ for capital[69], 1/RAMI, 1/(surgical complication rate), and DRG weighted discharge respectively, e_i is the error term that can be decomposed into an inefficiency term u_i and white noise v.

$$lnVC = a_0+\Sigma a_i lnw_i+\Sigma b_i lny_i+1/2\Sigma\Sigma a_{ij} lnw_i lnw_j+1/2\Sigma\Sigma b_{ij} lny_i lny_j+\Sigma g_{ij} lnw_i lny_j+e_i$$

Equation 1: Unrestricted Translog Cost Function

The conditions of symmetry (i.e. $a_{ij} = a_{ji}$ and $b_{ij} = b_{ji}$) and linear homogeneity in input prices (i.e. $\partial lnVC/\partial w_i=1$ meaning $\Sigma a_i+\Sigma b_i lny_i+1/2\Sigma\Sigma a_{ij} lnw_{ij}+\Sigma g_{ij} lny_j=1$) are imposed in the above equation. The homogeneity condition implies that the first term (Σa_i) is one,

[69] Note that we use a quantity measure for capital (i.e. y_K) because we specify a variable cost function. Therefore, we lose one cost share.

and the subsequent two terms are zero for each lnw_i and lny_i. From the first term, we obtained $a_L+a_M=1$. From the second term, $a_{LL}lnw_L+a_{ML}lnw_L+a_{LM}lnw_M+a_{MM}lnw_M=0$, and the condition is true if $a_{LL}+a_{ML}=0$ and $a_{LM}+a_{MM}=0$. For the third term, the condition becomes $g_{LK}lny_K+g_{LD}lny_D+g_{LA}lny_A+g_{LY}lny_Y+g_{MK}lny_K+g_{MD}lny_D+g_{MA}lny_A+g_{MY}lny_Y=0$, i.e. $\Sigma((g_{Li}+g_{Mi})lny_i)=0$. Then, $g_{LK}+g_{MK}=0$, $g_{LD}+g_{MD}=0$, $g_{LA}+g_{MA}=0$, and $g_{LY}+g_{MY}=0$. Substituting these constraints back into the cost frontier gives[70]:

$$
\begin{aligned}
lnVC = {} & a_0+[a_Llnw_L+(1-a_L)lnw_M]+\Sigma b_ilny_i+1/2[a_{LL}(lnw_L)^2+a_{MM}(lnw_M)^2+2a_{ML}lnw_Llnw_M] \\
& +1/2\Sigma\Sigma b_{ij}lny_ilny_j+[g_{LK}lnw_Llny_K+g_{LD}lnw_Llny_D+g_{LA}lnw_Llny_A+g_{LY}lnw_Llny_Y \\
& -g_{LK}lnw_Mlny_K-g_{LD}lnw_Mlny_D-g_{LA}lnw_Mlny_A-g_{LY}lnw_Mlny_Y]+e_i,
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow ln(VC/w_M) = {} & a_0+a_Lln(w_L/w_M)+\Sigma b_ilny_i+1/2a_{LL}[ln(w_L/w_M)]^2+1/2\Sigma\Sigma b_{ij}lny_ilny_j \\
& +[g_{LK}ln(w_L/w_M)lny_K+g_{LD}ln(w_L/w_M)lny_D+g_{LA}ln(w_L/w_M)lny_A+g_{LY}ln(w_L/w_M)lny_Y]+e_i,
\end{aligned}
$$

The substitution for $a_{LL}(lnw_L)^2+a_{MM}(lnw_M)^2+2a_{ML}lnw_Llnw_M$ arises from re-writing the whole term as $a_{LL}(lnw_L)^2-2a_{LL}lnw_Llnw_M+a_{LL}(lnw_M)^2=a_{LL}(lnw_L-lnw_M)^2$. The following translog cost frontier is estimated using ordinary least square as the first step of our analysis:

$$
\begin{aligned}
ln(VC/w_M) = {} & a_0+a_Lln(w_L/w_M)+b_Klny_K+b_Dlny_D+b_Alny_A+b_Ylny_Y+1/2a_{LL}[ln(w_L/w_M)]^2 \\
& +b_{KD}lny_Klny_D+b_{KA}lny_Klny_A+b_{KY}lny_Klny_Y+b_{DA}lny_Dlny_A+b_{DY}lny_Dlny_Y+b_{AY}lny_Alny_A \\
& +1/2[b_{KK}(lny_K)^2+b_{DD}(lny_D)^2+b_{AA}(lny_A)^2+b_{YY}(lny_Y)^2+[g_{LK}ln(w_L/w_M)lny_K \\
& +g_{LD}ln(w_L/w_M)lny_D+g_{LA}ln(w_L/w_M)lny_A+g_{LY}ln(w_L/w_M)lny_Y]+e_i,
\end{aligned}
$$

Equation 2 : Restricted Translog Cost Function (Symmetry and Linear Price Homogeneity)

After estimating this equation, we recovered the error term e_i for each hospital by subtracting the actual variable cost from the fitted variable cost. It is crucial that the estimated parameters are unbiased at this step. There is no requirement for best efficiency (i.e. smallest variance), hence multicollinearity is not a critical problem[71] as long as the parameters are of the correct signs. This is the starting point to decompose the error term.

[70] Ruud (2001) shows the detail of these calculations.

[71] This is unlike studies to evaluate the production structure where the purpose is the precise estimates of the parameters.

The error term $e_i = u_i+v$ is assumed to be normal-half normal, implying the log likelihood function is $lnL=ln2-Nln\sigma+\Sigma ln\Phi(e\lambda/\sigma)-1/(2\sigma^2)\Sigma e^2$ [see footnote for notation[72]] We maximized this function with respect to the parameters to obtain the maximum likelihood estimate for the parameter, using the ordinary least square residue (e_i) from the previous paragraph as the starting point. We then used the estimated parameters to recover the new residue (e_i) and iterated the process until convergence. We used the Battese and Coelli (1988) exact estimator to decompose the residue e_i into its components. The estimator is $E(-u_i \left| e_i \right.) = \ln[\frac{1-\Phi(\sigma_*-\mu_{*_i}/\sigma_*)}{1-\Phi(-\mu_{*_i}/\sigma_*)}].\{-\mu_{*_i}+\frac{1}{2}\sigma_*\}$, where $\sigma_*= \sigma_u^2\sigma_v^2/\sigma_u^2$ and $\mu_{*i}= e_i\sigma_u^2/\sigma^2$. The TSP output contains λ and σ, but not σ_{*i} or μ_{*i} directly. We obtained σ_* and μ_* using our proof in Appendix E, i.e. $\mu_{*i} = e_i\lambda^2/(1+\lambda^2)$ and $\sigma_*=\lambda^2\sigma^2/(1+\lambda^2)$. The point estimator is simplified using the symmetry property of the standard normal 1-$\Phi(z)=\Phi(-z)$. Hence, $E(u_i\,|\,e_i)=ln\Phi((\mu*_i/\sigma_{*i})-\sigma^*)+(-\mu_{*i}+0.5\sigma_*)-ln\Phi(\mu_*/\sigma_*)$. Finally, $CE_i=exp(u_i\,|\,e_i)$ gives us the hospital's efficiency score.

5.2 Data Sources and Software Selection

This research requires two levels of linked data. The construction of three variables (care quality, DRG adjusted admission and physician cost component) require the aggregation of patient level data to hospital level. Two variables for the cost efficiency modeling stage (capital and labor quantities) and two constructs for the hypotheses stage (hospital type and ownership) are available at the hospital level. The variable cost calculation requires hospital and physician components. The hospital component is available directly from hospital level cost data; the physician component requires aggregating patient-level data to the hospital level. These variables are calculated using the following databases:

[72] Recall from section 3.1 that u_i is the inefficiency term, v is white noise, $\Phi(.)$ is the cumulative normal density, $\lambda= \sigma_u/\sigma_v$, $\sigma^2 =\sigma_u^2+\sigma_v^2$.

5.2.1 The American Hospital Association (AHA) Annual Survey 1997

The AHA annual survey is a comprehensive hospital level database covering more than 6,000 hospitals. The scope of the survey covers utilization, organizational structure, personnel, hospital services and financial and accounting data. The 1983-7 AHA data represents 92% of the short-term community hospitals (Gaynor and Anderson, 1995). Given its long history[73] most U.S. hospital databases are designed to link with the AHA survey through an identifier (i.e. the AHAID field).

5.2.2 The National Inpatient Sample (NIS)

The Agency of Healthcare Research and Quality (AHRQ)[74] produces two patient level hospital databases from its healthcare cost and utilization project (HCUP) as at 1997. [75] The National Inpatient Sample (NIS), derived from the State Inpatient Databases (SID), is a stratified random sample of approximately 20% of all U.S. community hospitals. The NIS is more suitable than SID for our analysis because SID will not produce a random sample of U.S. hospitals if HCUP participation is not random. The NIS sampling frame consists of five strata: rural/urban location, number of beds, region, teaching status, and ownership. All discharges are retained for each sampled hospital. The NIS database is updated annually by using a recursive procedure to account for changes in strata sizes, composition and sampling rates. The goal of this procedure is to maximize the year-to-year overlap while keeping a constant sampling rate for all hospitals within each stratum. Consequently, time series analysis may be bias because the data is a subset of hospitals with continuous membership in the stratum. Specifically, the subset contains fewer hospitals that opened, closed, split, merged or changed strata.

The 1997 NIS includes 1,046 hospitals with about 7.1 million discharges from 22 states. However, AHA hospital identifiers are coded as missing to comply with the regulations

[73] The first AHA survey started in 1946.
[74] AHRQ is the research arm for the Department of Health and Human Services.
[75] Subsequently, three more databases are added for the statistics of emergency room, ambulatory surgery and pediatric cases.

from five states (Kansas, Georgia, Hawaii, Tennessee and South Carolina) and are rendered unusable. We therefore have usable data from 17 states: Arizona, California, Colorado, Connecticut, Florida, Illinois, Iowa, Maryland, Massachusetts, Missouri, New Jersey, New York, Oregon, Pennsylvania, Utah, Washington, and Wisconsin. Therefore, usable data are obtained from 729 hospitals representing 11.6% of all the U.S. hospitals at this stage.

5.2.3 Three Databases to Calculate Physician Cost Component

We used three databases to facilitate calculation of the market value of physician cost. These databases are the CPT4-ICD9CM crosswalk file, HCFA physician fees and the State Occupation and Employment and Wage Estimates 1999 from the Bureau of Labor Statistics.[76] The NIS database records the physicians' claimable procedures coded using ICD9CM. We obtained the price for each procedure from the HCFA Physician Fees database but HCFA does not use the ICD9CM to code procedure in this price list. The code that HCFA uses is the healthcare common procedure coding system (i.e. HCPCS), which is a modified form of the common procedure terminology version 4 (i.e. CPT-4)[77] code. We mapped the quantity (coded in ICD9CM) to the price (coded in CPT-4) to calculate total cost. We achieved this mapping with the ICD9CM-CPT4 crosswalk file.[78]

5.2.4 HCFA DRG weight File 1997

The final database is the DRG weight file which HCFA uses to determine the payment to hospitals. The file was available at http://www.hcfa.gov/stats/pufiles.htm (The HCFA website migrated when the organization changes its name to CMS).

---

[76] The database is available from URL: http://www.bls.gov/blswage and is reproduced in Appendix C. The data were collected from the National Compensation Survey, Occupational Employment Statistics Survey and Current Population Survey.
[77] Recall that the CPT-4 code is administered and maintained by the American Medical Association
[78] This file is obtained from the commercial coding software called CodeBreaker©. Info-X Inc., the compnay that produce this product, generously provided this database for my research.

5.2.5 Software Selection

Microsoft Assess 2000 is used to manipulate and prepare the databases for analysis. The choice of statistical software is limited to what the university has: namely SPSS, SAS, Minitab, EView, Gauss and TSP. Frontier, a specialized freeware for stochastic frontier analysis written by Professor Tim Coelli, can be also be obtained.

There is an inherent tradeoff between convenience and flexibility in the design of econometric software. Menu driven software such as SPSS, EView, Minitab and Frontier is easy to learn and convenient to use for common statistical analysis. However, they are difficult (if not impossible) for the user to customize.[79] Mason (1992) argues that the command line interface is better than menu interface for econometric work to enable program modification. Gauss, SAS and TSP are all competent software which can serve this purpose. The TSP user manual has an example of procedure for estimating a stochastic production frontier using maximum likelihood method. Modifying this procedure is easier than to program SAS or Gauss from scratch. Furthermore, the author for TSP provides technical support for user. TSP is used to estimate the cost frontier to benefit from its power and flexibility. Minitab is used to test the hypotheses to benefit from the convenience for testing standard procedures.

5.3 Variable Specification

The data described in section 5.2 is used to calculate the variables described in this section. These variables are then fed to TSP version 4.3A to calculate hospital efficiency scores based on the method described in section 5.1.

[79] EView is menu driven but allow some programming. Users do not have access to the source code for Frontier to program it.

5.3.1 Measuring Variable Cost

The conventional definition for variable cost is total cost minus fixed cost, where fixed cost is commonly measured by using capital cost. The disadvantage of this approach is that the measure of capital cost is usually the sum of depreciation and interest expenses. The arbitrariness in depreciation policy introduces some degree of measurement error.[80] The alternative approach is to use labor cost to proxy for variable cost. However, the tradeoff for this approach is the introduction of weak endogeneity because hospital labor wage is also a regressor. The proxy *Variable Cost = Total Cost – Capital Cost* is used because the endogeneity problem is more serious than the measurement error.

We subtracted the physician labor cost from the total cost in AHA survey to correct the problem that the total cost in integrated hospitals contains physician labor while the others do not. From this common baseline, we added back physician input expenditure to obtain the variable cost. We shall discuss the reversal of physician labor cost (in integrated hospitals) from the total cost in this section, and postpone our discussion of physician input expenditure to section 5.3.3. The AHA survey provides information on the number of physician each hospital hires[81] (in full-time-equivalent, i.e. FTE[82]). This number is zero in the segregate hospitals. We calculated the average physician wage by state using the data from the Bureau of Labor Statistics.[83] For most states, the salary per physician FTE in that state is the weighted average of the mean salary of internists, pediatricians, surgeons, obstetricians and gynecologists. The weight is the fraction of the specialists who responded to the survey. The estimated physician cost is the product of the number of physician in the hospital and weighted salary of physician in the state, *i.e. Physician Cost = Physician FTE in the hospital*Mean Salary in the State.* Therefore, *Variable Cost = Total Cost - Depreciation & Interest Expenses - Physician Wage + Physician Inputs.*

---

[80] Recall that we choose the stochastic frontier over DEA because we expect measurement errors.
[81] In the foundation model, the physicians are not directly hired by the hospital (they are hired by the foundation). Therefore, the FTE of physician was for the salary and equity models
[82] AHA assumed that each part time physician means half unit. Despite possible error of this approximation, it is still better than using headcount.
[83] The specific source is the State Occupation and Employment and Wage Estimates 1999.

5.3.2 Measuring Hospital Labor Wage Rate

We first calculated the hospital labor expenditure by subtracting the estimated physician cost (obtained in 5.3.1) from the hospital labor cost. Next, we calculated the number of full time equivalent (FTE) of non physician staff by subtracting the physician FTE from the hospital total FTE. We divided the hospital labor expenditure by the FTE of non-physician staff to obtain the average hospital labor wage rate.

5.3.3 Measuring Physician Input Price

The physician input price (w_M) for each hospital is $w_M = \frac{\text{market value of physician services}}{\text{number of FTE of nurses}}$. The NIS database records every procedure which physician bills the patient (see section 5.2.3 for details). This information gives us the number of each procedure done in the hospitals. The pricelist which HCFA uses to pay for the Medicare-Medicaid programs gives us the market price of each procedure. We calculated the market value of physician services from the sum of all physician revenue streams (i.e. price multiple by quantity) in that hospital. The coding systems for price and quantity data are different: specifically, quantity uses ICD9CM and price uses a modified CPT4 code known as HCPCS.[84] We mapped the two databases to a common coding system using a crosswalk file. Each CPT4 code corresponds to several ICD9CM codes, and each ICD9CM code corresponds to several CPT4 codes (i.e. the relation is many-to-many).

We use the example of cisternal puncture to illustrate the mapping procedure. Cisternal puncture means insertion of a needle into the space between the spines (for diagnostic or treatment purposes). This procedure can be used to collect fluid between the spines (call cerebral spinal fluid) for analysis, or inject a contrast media for radiography for diagnosis. Similarly, we can inject medicine for treatment (such as brain cancer). We started by examining the procedures in our NIS database and count the number of cisternal puncture

[84] ICD9CM is the international classification of disease version 9 (clinical modification), CPT4 is the current procedure terminology version 4, and HCPCS is the healthcare common procedure coding system.

(ICD 101) in each hospital. This ICD code corresponds to CPT 61050 (cisternal or cervical puncture without injection of substance) and CPT 61055 (cisternal or cervical puncture with injection of substance). Together with various local and national modifiers, there are 92 prices where the mean for CPT 61050 is 101.12 and the mean for CPT 61055 is 146.21. The price for ICD 101 is therefore 123.67, the average of these two prices. Note that prices for these CPT codes appear again in ICD codes. CPT 61050 corresponds to ICD 101 and ICD 392 (spinal canal injection); and CPT 61055 corresponds to ICD 101, 118 (other brain treatment procedure) and 392 (Rheumatic chorea). After obtaining the count of procedures and their (weighted) prices, we calculated the total market value of physician insurance billing in each hospital. If the market for physician labor is competitive[85], all the hospitals will get zero economic profit in hiring physician, i.e. total physician revenue stream equals to physician wage expenditure.

The estimated (average) physician wage rate $= \frac{\text{physician wage expenditure}}{\text{number of physician (in full time equivalent terms)}}$. However, we did not have the number of physicians who produced the revenue stream in each hospital. The number of physician under payroll is zero for segregate hospitals. The number of physicians as medical staff (i.e. physicians with admitting privileges) is not a useful proxy because some physicians had admitting privileges to several hospitals. Therefore, there is no consistent way to allocate the time these physicians spend in each hospital. We used the nursing full time equivalent in each hospital to proxy for the physician full time equivalent. This substitution requires the bold assumption that the ratio of the number of nurses to physicians is stable across hospitals. Variations in the specialty mix across hospitals may reduce the reliability of this proxy. For example, there are few nurses for each physician in psychiatric practice, but many nurses to each surgeon in cardiac surgery. Eliminating government and teaching hospitals from our sample reduce this effect to some extent.

5.3.4 Measuring Capital Quantity

[85] The competition for a physician to be the hospital's medical staff by employment or by admission privileges is strong. This is different from the monopolistic competition among physicians for patients.

Capital input is constant in the short run and is a form of hospital capacity constraint. Most researchers studying hospital cost use the bed size to proxy for capital quantity (some examples are Wagstaff, 1989, Bays, 1980; Granneman, Brown and Pauly, 1986). Zuckerman, Hadley and Iezzoni (1994) use depreciation and interest expense per bed to calculate capital price, hence indirectly using bed size to proxy capital stock. Carey (1997) uses total fixed asset minus depreciation to proxy capital stock. Bilodeau, Cremieux and Ouellette (2000) use building area and an index of furniture and equipment to proxy for capital stock. The more precise measure has three disadvantages. First, the additional inputs increase the number of estimable parameters in the translog form and increase the burden of multicollinearity. Second, the data are difficult to obtain. Third, including depreciation rate into capital stock introduces an inconsistent measure. This is because a wide range of depreciation policy (e.g. recognition of what constitute a fixed asset and the method of depreciation) is acceptable with accounting policies. Furthermore, there is a systematic variation in the depreciation policy in each group of hospital because of the local tax law, ownership type and firm growth. For-profit hospitals have the incentive to use aggressive depreciation to obtain depreciation tax shield. While it is true that an asset can only be 100% depreciated, an early tax shield is the same as having an interest free loan from the government.

There are also some disadvantages in using bed size as the proxy for quantity of capital. Hospitals with the same bed size can differ in the amount of capital equipment. Eliminating teaching hospitals from the sample reduces this effect. Deeble (1983) argues that bed size measures the capacity to accommodate and not the capacity to treat patient. It is the latter that is more important in a hospital. Berki (1972) also points out that some types of hospital beds are not substitutable (e.g. we cannot replace intensive care bed with surgical bed). We evaluated the tradeoffs and used bed size to proxy for capital stock. While the variable is measured with error, we did not expect systematic variation after eliminating teaching hospitals.

5.3.5 Measuring Risk Adjusted Mortality Index

Desharnais *et al* (1988) use a mixture of logistic regressions and contingency tables to calculate RAMI. However, we used only the contingency table method for simplicity. This is a three-step procedure. First, we calculated the actual and expected mortality rates for each DRG in each hospital. The actual mortality rate is the ratio of the number of deaths to the number of cases for a given DRG in a hospital. The expected mortality rate is the same ratio (i.e. death/case) for a given DRG in all hospitals in our sample. Second, we calculated a quality index by dividing actual over expected mortality for each DRG in each hospital. This index exceeds one if the hospital has higher mortality than expected for a DRG. Finally, we summed all the quality indices by hospital to obtain the hospital's RAMI. This simplified index is adjusted for casemix differences among hospitals, but does not consider the patient characteristics such as age and gender. High RAMI indicates low care quality. Therefore, our first quality index is the inverse of RAMI, i.e. $\frac{1}{RAMI}$.

5.3.6 Measuring Surgical Complication Rate

We used the surgical complication lists from Desharnais *et al* (1988) to identify cases of surgical complication in each hospital (We reproduce this list in Appendix D). Next, we counted the surgical cases in each hospital. The surgical complication rate is the number of complication divided by the number or surgery. Technically, a patient can have more than one complication for each surgery. Furthermore, long staying patients can have more than one surgery. A high complication rate indicates low care quality. Therefore, our second quality index is the inverse of complication rate.

5.3.7 Measuring Aggregate Hospital Output

Hospitals deliver "quantitative" patient care through inpatient and outpatient services. We aggregated all inpatient services using DRG weighted discharge. First, we used Assess 2000 to match the DRG of all discharges in National Inpatient Sample database to HCFA DRG weight file. Next, we calculated the scalar DRG weighted discharge Y where Y= Σ(count of discharge * DRG weights) for each hospital.

Since we knew inpatient, outpatient and total hospital revenue, we used the revenue as weights to include outpatient services, i.e. $\text{output} = \frac{\text{(total hospital revenue)*DRG weighted discharges}}{\text{inpatient revenue}}$. This way of aggregating outpatient services (instead of treating outpatient as another output) avoids the problem of zero output in translog functional form.

5.4 Regression Quality Control

We faced substantial difficulties in deriving reliable measures for most variables in the cost frontier. Measuring constructs is always a problem in econometrics and becomes an extreme challenge when analyzing hospital cost. Measurement errors compound the problem of random shocks to produce outliers. Removing these outliers could improve the reliability of the regression result.

Belsey, Kuh and Welsch (1980) suggest identifying outliers using the hat matrix. The diagonal of the hat matrix indicates the amount of influence an observation has on the regression. Observations with large hat diagonals have great influence on the regression. For a regression with p variables and n sample, the authors suggest that observations with a hat diagonal greater than $2p/n$ are outliers. Removing outliers reduces the sample size. Data quality is improved if outliers are due to coding errors. However, the risk of bias sample increases if the removed observations are systematically distributed. For example, if only inefficient hospitals provide wrongly coded data, removing these observations would improve data quality but produce bias sample.

5.5 Testing the Hypotheses

Testing the hypotheses is straightforward after we have classified the hospitals and obtained the cost efficiency scores. The statistical principle for evaluating the hypotheses involves testing the difference between two means for independent groups, assuming unequal sample size and unknown true standard deviations. The null hypothesis is: “the means are the same". The alternate hypotheses depend on the question, so we shall leave the discussion to Chapter 6. Minitab version 13 is used to test the hypotheses.

6. Result and Discussion

This chapter consists of the analysis of our cost frontier estimation, the result of the hypotheses testing, and examination of the threats to the result's validity and reliability. Following customary convention, we shall report the test significance based on its *p-value* (the probability of rejecting a true statement, i.e. Type 1 error) as:

Not significant: $p > 0.1$
Significant: $0.1 \geq p > 0.01$
Very significant: $p \leq 0.01$

Our final sample comprises 313 private hospitals representing 5% of the 6,284 U.S. hospitals who have responded to the American Hospital Association 1997 Survey. The 1997 National Inpatient Sample has patient level data for 1,012 hospitals who have participated in the HCUP project. Only 17 out of the 22 participating States provide the link to the AHA identifier. We are left with 729 available hospitals with complete data.

Some hospitals from the 729 available hospitals are excluded to increase internal validity. These include 137 government hospitals and 168 private teaching hospitals. From the remaining 424 hospitals, 28 hospitals are removed because of missing data elements[86], and three hospitals because of zero mortality or surgical complication rates. As discussed earlier, we need to run the cost frontier using ordinary least square to provide the starting value for the maximum likelihood procedure. This initial step is also used to identify observations with hat matrix diagonal exceeding $2p/n$ where p is the number of parameter and n is the number of observations. Thirty two outliers are identified and removed before proceeding to the second stage of the estimation. Finally, 361 valid hospitals are used to construct the efficiency model.

[86] The missing elements are bed size, hospital identifier and total cost.

We need to classify the hospitals to proceed with hypotheses testing: among the 361 valid hospitals, 56 hospitals are hybrid, i.e. these hospitals report both integrated model and network arrangements (e.g. salary model and integrated physician association). Eleven of these network arrangements involve only primary care physicians and not hospital specialists. As primary care physicians have limited rights for patient admission, we treat these eight hospitals as integrated. We eliminate the remaining 48 hospitals from hypotheses testing. Hence, we have 313 observations. The sample hospitals have the following characteristics:

Sample	For-Profit	Nonprofit	Total	Sample %	For-Profit	Nonprofit	Total
FIO	2	29	31	FIO	0.6	9.3	9.9
Network	23	83	106	Network	7.3	26.5	33.8
Segregate	47	129	176	Segregate	15.1	41.2	56.3
Total	72	241	313	Total	23.0	77.0	100.0

Table 1: Sample Hospital Characteristics

6.1 Parameter Estimates and Cost Efficiency

We used the translog cost function in this study. We tested and rejected the functional form for constant return to scale in quantity output (y_Y) using an F-test (Test Statistics = 5.776, upper tail area = .00016), and similarly the Cobb Douglas cost function (Test Statistics = 5.432, upper tail area = .00029). Our frontier model has high *R-square* (0.8764) and *adjusted R-square* (0.8711), but many parameter estimates are not statistically significant. The result is symptomatic of the multicollinearity problem as we expected. In fact, the conditional number is 5375. Belsley, Kuh, and Welsch (1980) suggest that condition numbers larger than 30 indicate serious multicollinearity, while numbers larger than 100 imply severe multicollinearity. [The famous Longley dataset (Longley, 1967) has condition number 43,275]. Multicollinearity is a problem when we are interested in parameters estimates, but is of lesser concern when we wish to obtain cost efficiency.

Using Sheppard's lemma, cost share $S_i = \frac{\partial \ln C}{\partial \ln w_i} = \frac{w_i}{C}\frac{\partial C}{\partial w_i} = \frac{w_i x_i}{C} = a_i + \sum a_{ij} \ln w_j + \sum g_{mn} \ln y_n$ where $i=L,M$ for hospital labor (L) and physician labor (M). For the hospital labor cost share, the regression for the stochastic frontier provides the parameter estimates for a_{ij} and g_{mn}. Actual data for lnw_i and lny_j are than substituted into the cost share equation to derive the estimated cost share. The procedure for physician labor cost share is similar except that the regression for the stochastic frontier does not provide the parameter estimates directly. These parameter estimates are derived from the imposed restrictions of symmetry and homogeneity in input prices. The estimated cost shares for hospital labor and physician labor are positive for all the hospitals. Therefore, the cost function is monotonic in inputs.

The estimated output parameters for DRG weighted discharge ($\hat{b}_Y$ =0.15) and inverse of surgical complication ($\hat{b}_A$= 0.45) are positive but not statistically significant. However, the estimated parameter for inverse of RAMI is negative ($\hat{b}_D$= -1.21). To evaluate the monotonicity in output, the cost function is differentiated with respect to the outputs to obtain $\frac{\partial \ln C}{\partial \ln y_i} = b_i + b_{ii} \ln y_i + \sum b_{ij} \ln y_j + \sum g_{ij} \ln w_j = MC_i * \frac{y_i}{C}$. Using the estimated parameters (b_i, b_{ii}, b_{ij}, and g_{ij}) and actual data (lny_i and lnw_j), the (estimated) marginal cost for each output in each hospital is obtained. It turns out that the marginal costs for DRG-weighted discharges and inverse of surgical complication are positive in all the hospitals. The marginal cost for the inverse of RAMI is negative for all the hospitals (see Appendix H). The DRG-weighted discharge is expected to have positive marginal cost, but there is no *a priori* information about the cost elasticity of care quality. The result for the inverse of RAMI is consistent with the findings of Neuhauser (1971), Shortell, Becker and Neuhauser (1976), and Longest (1978). We can speculate about the plausible reason that the care quality indicators have opposite marginal costs: Physicians facing product uncertainty and seriously ill patients prefer aggressive treatment than doing nothing. Some treatments are beneficial, some are useless, and some are even harmful. For the physician, saving life is the highest priority.[87] However, heroic attempts often only delay

[87] A common guide for emergency room practice is to "save life, save (bodily) function and then save limb" in that order of priority.

death and increase cost. On the other hand, the inverse of surgical complication rate is associated with low cost. This result is consistent with Scott, Forrest and Brown (1976) implying it is costly to deliver quality.

Bilodeau, Cremieux and Ouellette (2000) report checking the curvature of their cost function by calculating the eigenvalues of the Hessian matrix with respect to inputs and outputs. The following matrices are evaluated by hand if they are negative semidefinite:

$$\begin{pmatrix} aLL & aML \\ aLM & aMM \end{pmatrix} = \begin{pmatrix} -0.05596 & 0.05596 \\ 0.05596 & -0.05596 \end{pmatrix} \qquad \begin{pmatrix} bDD & bAD & bYD \\ bDA & bAA & bYA \\ bDY & bAY & bYY \end{pmatrix} = \begin{pmatrix} -0.2949 & 0.1639 & 0.0923 \\ 0.1639 & 0.01506 & -0.1043 \\ 0.0923 & -0.1043 & 0.2415 \end{pmatrix}$$

Input Hessian Matrix Output Hessian Matrix

The input Hessian is negative definite and only depends on the value of a_{LL}. This matrix is actually singular (i.e. determinant =0) because we have imposed homogeneity in input prices. Both matrices are symmetrical because we have imposed the symmetry conditions. The first, second and third leading principle minors of the output matrix are -0.295, -0.031, -0.008 respectively. Hence, the output Hessian is negative definite (and hence negative semi-definite). The concavity conditions are satisfied in the cost function.

We used the log likelihood ratio test to check on the joint restrictions of parameters symmetry and homogenous in input prices. We estimated the unrestricted cost function (i.e. Equation 1) by maximum likelihood and recovered the log likelihood from TSP as L_0. We then repeated the procedure with the restricted cost function (i.e. Equation 2) and recovered L_1. The test statistics is $TS= -2(lnL_0 - lnL_1) \sim \chi^2$ with 13 degree of freedom. Our test statistic is 39.83 > 22.36 (critical value at 5%). Hence, the restrictions are reasonable[88].

[88] Note that we are using a single equation and not an equation system of cost and cost shares. This choice (to avoid the Greene's problem) precludes the possibility of estimating the unrestricted cost function without imposing the symmetry condition. For clarity, estimating $y=f(.)+a_{ij}x_ix_j+a_{ji}x_jx_i$ will produce $y=f(.)+b_{ij}x_ix_j$. This is because $x_ix_j= x_jx_i$ mathematically, and regressing both variables will produce the perfect multicollinearity problem. The homogeneity in input prices is a joint restriction, and it is not possible to implement restriction on each input price individually (see Berndt (1991) p. 469 for detail), i.e. the joint restrictions imposed are: The sum of the input parameters equals one, and parameters for the cross products of inputs-outputs are zero.

Our most important result from the cost frontier is the cost efficiency in each hospital. The mean cost efficiency is 87.4% (median = 87.8%), the standard deviation is 3.6%, and the range is 75.8% to 95.4%. The efficiency scores are comparable to most recent studies using stochastic cost frontier such as Patel, Needleman and Zechauser (1994), Sloan, Taylor, Picone and Chou (1998), Zuckerman, Hadley and Iezzoni (1994). A histogram of the efficiency score distribution is shown as follows:

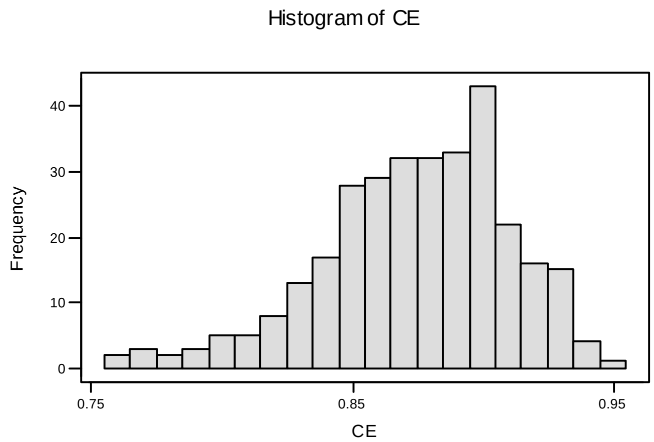

Figure 1: Histogram of Cost Efficiency in Sample Hospitals

6.2 Results for Hypotheses Testing

Our results show that the 29 nonprofit FIOs have higher mean cost efficiency (0.876 versus 0.857) than the 83 nonprofit network hospital (*p-value=0.028).* The *p-value* denotes the chance of Type 1 error for rejecting the hypothesis that the mean efficiency score is the same. This is a one tail test because the alternate hypothesis is: the first mean is higher than the second. The nonprofit FIO also have higher efficiency (0.876 versus 0.853) than the 129 nonprofit segregate hospitals (*p-value=0.040).* Holmstrom (1982) argues that a budget breaking incentive scheme can overcome problem of shirking when individual marginal product is not measurable. This hypothesis (H1) is empirically supported.

Our results show that physician hospital integration does not increase the cost efficiency in for-profit sector. Specifically, the mean efficiency of two for-profit FIOs in our sample is 0.879, the mean efficiency of 23 for-profit network hospitals is 0.891 (*p-value=0.729* for comparing for-profit FIOs and network hospitals*)*, and the mean efficiency of 47 for-profit segregate hospitals is 0.891 (*p-value=0.733* for comparing for-profit FIOs and segregate hospitals*).* These are two-tail tests because the (common) alternate hypothesis is: the first mean is the same as the second mean. Eswaran and Kotwal (1984) argue that the principal in the for-profit firm faces moral hazard problem herself. She can gain higher profit by offering higher payoff to one of the agents in the team so that the team will miss the team bonus. We find empirical support for this hypothesis (H2). We view this hypothesis jointly with the first one as a control. Integration improves cost efficiency in the absence of principal moral hazard, but the effect disappears when the principal's profit incentive comes into the picture.

The next question is: what will happen if the principal cannot implement a budget breaking scheme? This scenario happens in network and segregate hospital because physicians are not hospital employees. Arrow (1963) argues that the physician-patient relation is society's answer to the inherent uncertainty in medical care. Therefore, physicians and hospitals are separate legal entities (and financially independent) to ensure physicians act in the patients' best interests. The physician-patient relation prevents (segregate) hospitals from providing direct financial incentive (let alone a budget breaking one) to physicians, driving the Medical Arms Race among hospitals in the 1970s. Managed care organization in network hospital may provide financial incentive to physician via selective contracting. Physicians remain as independent professionals and some are medical staff for several hospitals. Managed care organizations therefore administer incentive (if any) to the individual hospital or physician, but not to both as a group. Although incentive payments are made, receiving the bonus does not depend on the team's performance (it depends on the individual's performance). Hence, the mechanism that Holmstrom describes does not occur in this case. Alchian and Demsetz (1972) argue that monitoring is a plausible mechanism for firms to be efficient. Frech (1976) argues that for-profit firms are more efficient because there is a clear residual claimant. Our analysis differs from most analyses in the current literature by excluding FIOs where budget breaking mechanism can influence efficiency.

From the results, the 23 for-profit network hospitals, with mean efficiency 0.891, are more cost efficient than the 83 nonprofit network hospitals in our sample, with mean efficiency 0.856, (*p-value=0.000).* In addition, the 47 for-profit segregate hospitals, with mean efficiency 0.891, are also more efficient than the 129 nonprofit segregate hospitals with mean efficiency 0.853, *(p-value=0.000).* Therefore, we find evidence to support the property rights theory (H3).

An interesting question remains unanswered: the team agency theory predicts that nonprofit FIOs are more efficient than for-profit ones, but the property rights theory predicts just the opposite. When two forces act in opposite directions, the comparative static will depend on the relative strength of the forces. Our result shows that the mean efficiency in nonprofit FIO (0.876) is statistically similar to for-profit FIO (0.879). The hypothesis of equal efficiency has a *p-value=0.916.* We argue that budget breaking and monitoring mechanisms are not mutually exclusive in this case.[89] After all, firms that administer bonus scheme often monitor their employees.

6.3 Review of Research Validity

According to Trochim (2000), we can review the validity of research from four angles: internal validity (research design), external validity (inference), constructs validity (measurement) and conclusion validity (assumption).

6.3.1 Internal Validity

Internal validity relates to the ability of the research to infer causal relationship. Cook and Campbell (1979) state that we can infer causal relation by satisfying three conditions:

[89] The groups have equal incentive for monitoring in hypotheses 1 and 2; the groups in hypothesis 3 cannot administer a budget breaking incentive; the groups in the current situation can do both.

temporal precedence, correlation, and the absence of alternate explanation. Controlled experiment has the strongest internal validity. Our research question does not allow us to administer treatment effect (such as ownership or organizational form) to the hospital and hence precludes true experiment. The second best alternative is the natural experiment for hospital conversion and compare *ex ante* with *ex post* efficiency. There are two possible variations here. The hospital acts as its own control for other factors which may influence cost efficiency in the first variation. However, a hospital changing organizational structure can take a long time to stabilize. In the interim, environmental shocks can influence the hospital's cost efficiency. There are also fewer samples in this case than observing a cross section of hospitals. Small sample size, especially in for-profit FIOs, produces the risk of obtaining an empty group. Measuring cost efficiency is also less reliable with stochastic frontier (and DEA) using small sample size. In the second variation, a matched pair is used as control to avoid the problem of environmental shocks. However, there is no sensible way in choosing matched pairs. Halving the sample size in the matched pair method increases the chance of missing groups.

Our most feasible choice is to use a correlation design (which is also the most common in empirical analysis of hospital cost efficiency). The tradeoff in this decision is the inability to establish the temporal precedence of effects because independent variables cannot be manipulated. This disadvantage increases the risk of alternative explanation in causing the observed effect. Eliminating teaching and government hospitals from our sample reduces this risk. Although correlation alone cannot prove a theory, the lack of the expected correlation can refute it. We have falsifiable hypotheses to test the application of team agency theory to the research question. There is a need to replicate this research with other research methodologies to prove the robustness of the application.

6.3.2 External Validity

How valid is our conclusion outside our sample? We started with a census of over 6,000 U.S. hospitals, obtained 729 usable ones derived from stratified random sampling, and finally used 313 hospitals to test our hypotheses. The external validity of our result depends on whether the 313 hospitals are representative U.S. hospitals. Due to the data limitation for casemix, we compared our sample with the 729 usable data on dimensions which may affect cost or cost efficiency: bed size, average length of stay, occupancy, geography, integration, and ownership and casemix.[90] Since the 729 usable hospitals are derived from rigorous stratified sampling to represent 20% of U.S. community hospitals, establishing comparable characteristics to this (reference) group ensure certain level of external validity. We tabulate the results as follows:

Characteristics	Reference (N=729)	Sample (N=313)	*P-values*
Mean Bed Size	196	154	*0.000****
Mean Length of Stay (Days)	7.80	6.98	*0.205*
Mean Occupancy (%)	56.7	54.5	*0.060**
Nonprofit hospital %	80.8%	77.2%	*0.216*
Integrated hospital %	14%	10.8%	*0.177*
Region (1-9)*	(See Text)	(See Text)	*0.998*
Casemix (MDC Level)*	(See Text)	(See Text)	*1.000*

Table 2: Comparison of Sample and Reference Population

Our sample hospitals have significantly less beds than the reference population. This is due to the elimination of teaching hospitals which tend to have more beds. There is no significant difference in the length of stay, although the occupancy rate in the sample is lower. There is no statistically significant difference in the composition of nonprofit or integrated hospitals between the sample and reference populations. These *p-values* are calculated from *t*-statistics. We tested the sample-reference hospital differences for geographical regions and casemix at two levels. For geographical region, we first calculated the *t*-statistics for the difference in proportion between the sample and reference population in each of the nine geographic regions. The proportions of sample and reference hospitals in each region are not significantly different (*p-values* ranges from 0.339 to 0.927). We then calculated the χ^2 using the sample hospital as observed data and reference hospital as expected result for all nine regions. The *p-value = 0.998*. These results show no statistical difference in geography between the sample and

[90] Bed size may influence the economy of scale; treatment aggressiveness may affect the length of stay; occupancy rate affects the average fixed cost; geography may affect wage rates, and HCFA payment allows for geographic adjustment; we argue that integration and ownership influence cost efficiency in this dissertation; casemix affect cost. We leave out teaching status.

reference hospitals for each region and as a whole. We used the same procedure to analyze casemix at the MDC (major disease classification for the DRG system) level. The proportion of digestive disorder cases (MDC 6) in the sample hospitals is statically smaller than the reference hospitals (*p-value=0.042).* The proportion of alcohol/drug induced mental disorder (MDC 20) cases is also lower in the sample hospitals than reference hospitals (*p-value=0.022*). The low proportion of the MDC 20 cases arose from eliminating public hospitals that treat most of the mental disorder cases. There was no obvious reason for the digestive disorder cases (MDC 6). From the result of χ^2 test, casemix difference between sample and reference population at the joint level is not statistically different. Overall, our sample is comparable to the reference, and hence the population of private U.S. community hospital.

6.3.3 Construct Validity

Classifying hospitals into for-profit or nonprofit is unambiguous. When a hospital reports both an integrated model and a network model (e.g. a salary model and also an independent physician association, IPA), establishing if the hospital has an employment relationship with the specialist physicians is less obvious. There are two possibilities. First, the hospital may hire specialists and form an IPA with primary care physicians who refer patients to the specialists. In this case, the hospital remains as a FIO because the primary care physicians do not admit patients to the hospital. Second, the hospital may hire some specialists and form an IPA with more specialists and primary care physicians. These hybrid hospitals were removed from our sample unless we established the first scenario was true. We did not remove all the hybrid hospitals to conserve the number of observations. We also removed all the hybrid hospitals and re-tested the hypotheses. The results remain unchanged.

The main threat to construct validity is the measurement of cost efficiency. The calculation of the dependent variable *(VC)* requires three approximations. First, we subtracted physician salary from the total cost by a quantity *Q= FTE*average wage. FTE* is the number of physician full-time-equivalent in the salary and equity models. *Average* *wage* is the specialty weighted average physician salary in each state, or the national weighted mean for some states. This measure is unbiased because of the *Law of Large Number*, but is statistically noisy. Second, the market value of physician inputs is added back. The mapping of one code system to another (i.e. cross-walking) which involved many-to-many relations is required in this step. The reliability of the measure is reduced by the complicated mapping process. Third, the capital cost is added back as the sum of interest and depreciation expenses. This measure of fixed cost can vary systematically with the ownership type and location of the hospitals. These three sources of errors are expected to be uncorrelated. The other acute measurement problem is the calculation of the physician input price (more accurately the shadow price). The noise in the calculation of the market values of physician services is carried over to the physician input price. The second source of noise in the physician input price is the use of nursing full time equivalent to proxy for physician. The physician input prices are expected to be noisy because physician/nurse ratio can vary by clinical specialty. Nevertheless, using the nursing full-time-equivalent is more reliable than the reported number of physicians. This is because the reported number of physicians means very different things: the number of specialist physicians who can admit patient in integrated models; the number of primary care and specialist physicians (many with admitting privileges in several hospitals) in network hospitals; zero in the segregate hospitals. Our regression parameter for physician wage is not significant.

Strong multicollinearity is another measurement issue. The parameter estimates remain unbiased but inferences using the parameter are imprecise. Fortunately, this problem does not affect the cost efficiency measure if the cost frontier is specified correctly. Our counter-measures to mitigate the effects of potential measurement errors are to use the stochastic frontier (instead of data envelopment analysis) and remove outliers using the hat matrix. Our counter-measure for possible specification error in the cost kernel is to use a flexible functional form, which also contributes to multicollinearity.

6.3.4 Conclusion Validity

The conclusion validity rest on the assumptions we made in the empirical strategy. We assumed a stable technical relationship exists in hospital production which allows the specification of a cost function. The results show that our cost frontier is monotonic in inputs and (most) outputs, concave, and conform to restrictions in parameter symmetry and homogeneity in input prices. However, the mortality measure (RAMI) is not monotonic in output because of the special characteristics of the hospital industry (physician should commit infinite resource to reduce mortality). Otherwise, the cost function is reasonably well behaved.

We made a strong assumption that the composed error was normal-half normal. This is a tradeoff between the risk of specification error and econometric tractability. Using panel data techniques circumscribe the need for specifying a distribution structure for the efficiency term, but introduce new problems: the assumption of time-invariant firm effects in fixed effect model, and the problem of error-parameter independence in the random effect model. In the context of the hospital market, there is no clear preference about assumptions for error specification, time invariance and efficiency-parameter independence. We acknowledge the limitation of our specifications and shall leave the panel data technique for further research.

7. Conclusion

This final chapter concludes the research question in the context of existing research. We shall explain how the result can contribute to theory, policy and managerial practices. Finally, we shall identify areas which require further investigations.

7.1 Conclusions about Research Problem

Our result shows the empirical evidence in the nonprofit hospital to support Holmstrom (1982) hypothesis: Administering a budget breaking incentive scheme to the production team can overcome moral hazard (i.e. shirking) when individual effort cannot be metered. This effect disappears in the for-profit hospital where the principal’s moral hazard problem arises (maximizing the residual is not socially optimal). The result extends the research if for-profit hospitals are indeed more efficient than nonprofit ones by adding an integration dimension. However, we cannot administer the budget breaking incentive in network and segregate hospital. In the absence of this effect, we find that for-profit hospitals are more efficient than nonprofit ones. An indeterminate situation arises when comparing FIOs where the theory of property rights and team agency predict opposite effects. We find no difference in cost efficiency in this case.

7.2 Implications for Theory and Research

The research on the relative cost efficiency of for-profit and nonprofit U.S hospitals in the 1980s has reached consensus by the late 1990s. Heightened competition in the hospital market would eliminate any difference which might exist. At the same time, the increased competition in the market produces new organizational alliances that attracted the attention of researchers. Are these new alliances more cost efficient? Why? Cuellar and Gertler (2006) examine whether these alliances exercise market power to raise price without improving care quality. They find that "integration" in the network setting is about exercising market power and not about the gain from transaction economies. They also find that FIO is different. We believe their FIO sample comprises mostly nonprofit hospitals, and their result is related to our first hypothesis.[91] Our rationale for the cost efficiency gain arise from reducing agency cost (a form of transaction cost), the pro-competitive mechanism highlighted in Varney (1995) healthcare antitrust speech. We have introduced team agency to enrich the analysis of the relative cost efficiency in the for-profit and nonprofit hospitals. There is a clear trend for physician hospital integration in the 1990s. However, the empirical research on the relative efficiency of these integrated organizations (e.g. Wang, Wan, Clement and Begun, 2000) requires a theoretical framework for support.

We have specified the hospital cost function differently from most analyses. First, we examine the theoretical literature on hospitals (Newhouse, 1970; Pauly and Redisch, 1973; Harris, 1977) and conclude that we need to include the physician components in cost and input. Second, we rely on Arrow (1963) to argue that the hospital's output is really patient care, which needs quantity and quality as joint proxies.[92] The structural cost function is appealing if we can accommodate variables that affect cost (This is the motivation for the paper by Granneman *et al*). For the restricted group of hospitals excluding teaching and government hospitals, we have bridged part of the gap between theoretical framework and empirical specification of the hospital cost function.

7.3 Implications for Policy

In this dissertation, we apply agency theory to predict that nonprofit hospitals are more cost efficient if they integrate, but the effect is absent in for-profit hospitals. Our results support these predictions. Increasing cost efficiency is welfare enhancing, and we identify the source of this efficiency in agency cost. Federal Trade Commissioner Varney

[91] However, our measure of efficiency was different from Cuellar and Gertler (2006).
[92] Newhouse (1980) also sees physician as jointly maximizing quantity and quality in hospital production.

says in 1995: "… the greatest pro-competitive, efficiency-enhancing benefits are those that selectively contract so that they can control costs by ensuring that providers have the strongest incentive not to over treat or over utilize". Therefore, physician hospital integration in the nonprofit sector is more likely to generate efficiency enhancing effects than integration in the for-profit sector.

7.4 Areas for Further Research

The areas for further research arise from the limitations of this study. First, the small number of integrated for-profit hospitals weakens our result. Using an expanded data set can overcome this problem. The national inpatient sample expanded from 22 states in 1997, to 33 states in 2001. Second, cross section data, fixed effect model and random effect model have their own tradeoffs. There is a need to apply different techniques to confirm the result is robust. Third, the correlation design has weaker internal validity than natural experiment. There is yet any reported research which use natural experiments of organizational conversion to verify our hypothesis. Duke University (2002) reported to have used cross section comparison and natural experiment to examine the effects of ownership conversion. Examining our proposition at the organizational level, especially for hybrid form, by using the case research method can provide insight for firm level "micro-factor" which can influence organizational cost efficiency. Case research method complements industry study: the former is weaker in external validity because the firm level driver may not generalize to other firm; the latter is weaker in internal validity because many firm level factors are not captured.

7.5 Concluding Remarks

This research extends the property rights view that for-profit hospitals are more cost efficient than nonprofit ones. This is the first research we know that uses the team agency theory to explain cost efficiency differences among hospitals. As more and more U.S.

hospitals start to “integrate” physicians as alliances or employees, there is a need to bring theoretical insight in analyzing cost efficiency of these new organizational forms. This research is a first step in this direction and replication with other methods is needed to confirm the robustness of the result. Section 6.3 explains the threats to the research validities and our counter-measures. Despite the difficult measurement issues, we have obtained a reasonably “true and fair” view of the research question.

Appendix A: Summary of Notations and Symbols.

The notations and symbols used in this dissertation are similar to most econometric textbooks, especially Berndt (1991). In general, lower case denotes function or variable and upper case specify input/output. For example, *f(.)* denotes production function; w denotes input price and K denote capital stock. The subscript enables the variable to become specific. For example, w_L denotes wage or input price of labor. The derivate is denoted with superscript and subscript when we have simple terms. For example, f_L' means first derivative of the production function with respect to labor input, i.e. the marginal product of labor. There is also a convention in the coefficients of the functional forms. The intercept is always a_0. In the translog form, the coefficient for input is a_i ($i \neq 0$), output coefficient is b_i, and the cross products coefficients are a_{ij}, b_{ij} and g_{ij} respectively for cross products of inputs, outputs and mixture of both. In regression, the subscripts make the coefficients intuitive, especially for cross products. The following is a listing of the notations and symbols used in this thesis:

f(.)	Production Function
t(.)	Transformation Function, the multi-product production function
c(.)	Cost Function, minimize cost subject to *f(.)*
h(.)	Input Aggregator Function
g(.)	Output Aggregation Function
y	Output Quantity
p	Output Price
q	Quantity
x	Input Quantity
w	Input Price
θ	Constant
K	Capital
L	Labor
M	Medical Labor (i.e. physician)
D	Death Rate (i.e. risk adjusted mortality index)
A	Adverse Reaction (i.e. complication rate)
Y	Aggregate Hospital Output
S	Share Equation
s	Elasticity of Scale
σ	Allen's Elasticity of Substitution or Standard Deviation
Δ	Discrete difference of two quantities
f_L'	First derivative of f with respect to L i.e. $\partial f/\partial L$
f_L''	Second derivative of f with respect to L i.e. $\partial^2 f/\partial^2 L$
e	Composed Error Term (also known as composite error)
v	Random Error Term (i.e. white noise)
u	One-sided Error Term (i.e. related to efficiency measure)
η	Error Term for Cost share/Input demand in equation system
λ	Ratio of σ_u^2/σ_v^2
E[.]	Expected Value (i.e. mean)

var(.) or σ^2	Variance
$\sum$	Multivariate covariance matrix
z(.)	Probability density function
$\phi(.)$	Standard Normal Function
$\Phi(.)$	Standard Cumulative Normal

The following are abbreviations used in this dissertation in alphabetical order

AHA	American Hospital Association
AHRQ	Agency for Healthcare Research and Quality
AMA	American Medical Association
AUTOGRP	A bio-statistical package/grouper
BLS	Bureau of Labor Statistics, a government agency
COLS	Corrected Ordinary Least Square
CPT, CPT4	Common Procedural Terminology, CPTE denotes 4th version
DEA	Data Envelopment Analysis
DRG	Diagnosis Related Group
FTE	Full Time Equivalent
GLS	Generalized Least Square
GPWW	Group Practice without Wall
HCFA	Healthcare Financing Administration, renamed as CMS
HCPCS	HCFA Common Procedure Coding System
HMO	Health Maintenance Organization
ICD9CM	International Classification of Disease, version 9, Clinical Modification
ISM	Integrated Salary Model
JCAH	Joint Commission on the Accreditation of Hospitals
LSDV	Least Square Dummy Variable
MAR	Medical Arms Race
MLE	Maximum Likelihood Estimator
MOLS	Modified Ordinary Least Square
MRI	Magnetic Resonance Imaging
MSO	Management Service Organization
NIS	National Inpatient Sample
OLS	Ordinary Least Square
PHO	Physician Hospital Organization
PPS	Prospective Payment System
RAMI	Risk Adjusted Mortality Index
RB-RVS	Resource Based Relative Value System
SID	States Inpatient Database
SUR	Seemingly Unrelated Regression
TSP	Time Series Processor, an econometric program
VA	Veteran Affairs

Appendix B: Hospitals in the United States

	1980	1990	2000
Industry Capacity	6,965 (1,364,516)	6,649 (1,213,327)	5,810 (983,628)
Hospitals (Beds)			
Ownership: Hospital (Beds)			
Federal	359 (117,328)	337 (98,255)	245 (53,067)
Nonprofit	3,322 (692,459)	3,191 (656, 755)	3,003 (582,988)
For-profit	730 (87, 033)	749 (101,377)	749 (109,883)
State Government	1,778 (208, 895)	1,444 (169,228)	1,163 (130.689)
Community Hospital*			
Bed Size Distribution			
6-199 Beds	4,120 (70.6%)	3,770 (70.0%)	3,489 (71.1%)
200-499 Beds	1,393 (23.9%)	1,369 (25.4%)	1,179 (24.0%)
Above 500 Beds	317 (5.5%)	285 (4.6%)	247 (4.9%)
Total	5,830 (100%)	5,384 (100%)	4,908 (100%)
Occupancy			
Federal	80.1	72.9	68.2
Nonprofit	78.2	69.8	65.5
For-profit	65.2	52.8	55.9
State Government	71.1	65.3	63.2
Admission (‘000s)			
Federal	2,044	1,759	1,034
Nonprofit	25,566	22,878	24,453
For-profit	3,165	3,066	4,141
State Government	7,413	5,236	4,496
Total	38,892	33,774	34,891
Average length of stay			
Federal	16.8	14.9	12.8
Nonprofit	7.7	7.3	5.7
For-profit	6.5	6.4	5.4
State Government	7.3	7.7	6.7
All Hospital mean	9.9	9.1	6.8
Outpatient Visits (‘000s)			
Federal	50,566	58,527	63,402
Nonprofit	142,156	221,073	393,168
For-profit	9,696	20,110	43,378
State Government	50,459	60,146	84,858
% Outpatient Surgery	16.3	50.5	62.7

* Community hospital includes all but Federal hospitals

Source: US DHSS, Health United States 2003 (Table 95, 106)

Appendix C: Mean Physician Wage from Bureau of Labor Statistics [58]

	Number of Physician					Mean Physician Salary					
States	Internist	OBGYN	Pediatrician	Psychiatrist	Surgeon	Internist	OBGYN	Pediatrician	Psychiatrist	Surgeon	Mean
Arizona	770	0	110	90	210	114,170	140,870	121,720	90,700	127,510	115,458
California	2,870	1,110	1,600	1,530	2,470	131,390	137,170	84,550	110,160	121,170	118,211
Colorado*	48,740	18,780	18,940	17,870	48,450	123,280	135,430	112,760	103,660	135,660	125,100
Connecticut	500	490	330	NA	1,560	124,520	134,280	96,100	NA	133,420	127,745
Florida	2,560	2,090	1,490	450	2,060	132,710	145,020	122,640	119,100	143,310	135,766
Georgia	NA	NA	120	140	NA	127,650	134,880	104,980	114,980	NA	110,365
Illinois	960	NA	580	690	3,510	96,740	NA	87,040	94,650	106,890	101,715
Iowa*	48,740	18,780	18,940	17,870	48,450	123,280	135,430	112,760	103,660	135,660	125,100
Kansas	NA	190	NA	80	NA	110,440	138,680	NA	122,990	145,520	134,031
Maryland	NA	560	NA	430	NA	115,750	104,300	93,130	110,610	144,640	107,041
Massachusetts	NA	380	610	330	NA	126,810	139,480	101,310	102,630	140,710	112,628
Missouri	1,010	200	340	530	540	116,870	111,150	119,660	114,360	143,700	121,818
New Jersey	NA	NA	NA	280	5,560	NA	131,730	NA	112,210	137,320	136,116
New York	3,570	1,200	1,570	2,360	3,320	100,400	141,100	113,680	111,910	130,650	116,813
Oregon	48,740	18,780	18,940	17,870	48,450	123,280	135,430	112,760	103,660	135,660	125,100
Pennsylvania*	3,950	530	880	1,310	2,270	128,800	138,300	117,520	110,020	139,140	128,126
South Carolina	390	250	120	220	1,020	136,660	123,900	119,330	95,330	138,830	130,586
Tennessee	1,460	580	570	NA	1,780	129,940	143,390	104,780	NA	143,640	134,005
Utah	260	150	270	NA	230	124,730	130,830	107,400	NA	124,090	120,432
Washington	2,750	NA	NA	710	1,100	135,850	140,720	117,370	86,450	138,420	128,778
Wisconsin	1,070	70	250	NA	270	140,260	128,170	119,390	NA	140,250	136,605

[58] We denote some data as NA (Not Available) because either they are not available to BLS, or the standard errors are too large. If these occur, we do include the missing category to compute the mean (last column). In some states, the * denote the use of US average because the data are not available at the state level.

Appendix D: Complication List in Desharnais *et al* (1988)

SN	ICD9CM	Description
1	2513	Post-surgical hypo-insulinemia
2	3490	Lumbar puncture reaction
3	3491	Central nervous system complication from surgically implanted device
4	47830	Vocal cord paralysis unspecified
5	47831	Vocal cord paralysis unilateral partial
6	47832	Vocal cord paralysis unilateral total
7	47833	Vocal cord paralysis unilateral partial
8	47834	Vocal cord paralysis bilateral total
9	5070	Food / Vomit pneumonitis
10	5304	Perforation of esophagus
11	65930	Septicemia in labor
12	66800	Pulmonary complication in delivery unspecified
13	66802	Pulmonary complication - delivered with postpartum complications
14	66804	Pulmonary complication postpartum
15	66880	Anesthetic complication delivery *episode of care unspecified*
16	66881	Anesthetic complication delivery [not] mentioned ante partum condition
17	66882	Anesthetic complication delivered with postpartum complication *mentioned*
18	66883	Anesthetic complication ante partum
19	66884	Anesthetic complication ante partum
20	66890	Unspecified anesthetic complication delivery episode of care unspecified
21	66891	Unspecified anesthetic complication delivery [not] mentioned ante partum condition
22	66892	*Unspecified* anesthetic delivered with postpartum complication *mentioned*
23	66893	*Unspecified* anesthetic complication ante partum
24	66894	*Unspecified* anesthetic complication postpartum
25	66912	Obstetric shock - delivered with postpartum complications *mentioned*
26	66914	Obstetric shock - with postpartum complications
27	66930	Acute renal failure with delivery *episode of care* unspecified
28	66932	Acute renal failure delivered with postpartum complications
29	66934	Acute renal failure postpartum
30	67002	Major puerperal infection - delivered with postpartum complication
31	67004	Major puerperal infection with postpartum complication
32	67300	Obstetrical air embolism *episode of care* unspecified
33	67302	Obstetrical air embolism delivered with postpartum complication
34	67304	Obstetrical air embolism postpartum complication
35	67400	Cerebrovascular disorders in the puerperium *episode of care* unspecified
36	67402	Cerebrovascular disorders delivered with postpartum complication
37	67404	Cerebrovascular disorders postpartum complications
38	67410	Disruption of cesarean wound unspecified
39	67412	Disruption of cesarean wound delivered with postpartum complications
40	67420	Disruption of perineal wound unspecified
41	67422	Disruption of perineal wound delivered with postpartum complications
42	67424	Disruption of perineal wound delivered with postpartum
43	67512	Breast abscess delivered with postpartum complications
44	9954	Shock due to anesthesia

SN	ICD9CM	Description
45	9970	Surgical complication - central nervous system
46	9971	Surgical complication - heart
47	9972	Surgical complication - Peripheral vascular system
48	9973	Surgical complication - respiratory system
49	9974	Surgical complication - gastrointestinal tract
50	9975	Surgical complication - urinary tract
51	99762	Infection amputation stump
52	9979	Surgical complication - other specified body systems
53	9980	Postoperative shock
54	9981	Hemorrhage complicating a procedure
55	9982	Accidental laceration during a procedure
56	9983	Postoperative wound disruption
57	9984	Foreign body left during a procedure
58	9985	Postoperative infection
59	9986	Persistent postoperative fistula
60	9987	Postoperative foreign substance reaction
61	9988	Other specified surgical complication
62	9989	Other surgical complication unspecified
63	9991	Air embolism following medical care
64	9992	Other vascular complication following medical care
65	9993	Other infection as a complication of medical care
66	9994	Anaphylactic shock - serum
67	9995	Other serum reaction
68	9996	ABO incompatibility reaction
69	9997	Rh. incompatibility reaction
70	9998	Other transfusion reaction

* Careful readers may notice some differences between Desharnais *et al* (1988) Appendix and the above. There is a correction of a typographic error in SN37, the ICD9CM code is 67404 not 67402, i.e. the ICD9CM code in SN 36 and 37 are different. Some words are inserted in italics after examining the ICD9CM to ensure that the meanings are clear. Specifically, for four-digit codes 668.0 (pulmonary complications), 668.8 (Other complications of anesthesia or other sedation in labor and delivery) and 668.9 (Unspecified complication of anesthesia and other sedation), the fifth-digit sub-classification is (for use with categories 660-669) denotes the current episode of care: 0 (unspecified as to episode of care or not applicable), 1 (delivered, with or without mention of ante partum condition), 2 (delivered, with mention of postpartum complication), 3 (ante partum condition or complication) and 4 postpartum condition or complication).

Appendix E: Proofs relating to μ_* and σ_*^2

Let us define four identities: $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \sigma_u/\sigma_v$, $\mu_* = e\sigma_u^2/\sigma^2$, $\sigma_*^2 = \sigma_u^2\sigma_v^2/\sigma^2$

We wish to re-write μ_* and σ_*^2 in terms of e, λ and σ, so that we can use the TSP result to calculate firm efficiency. We have:

$\sigma^2 = \sigma_u^2 + \sigma_v^2 \Rightarrow \sigma_v^2 = \sigma^2 - \sigma_u^2$……..(1)

$\lambda = \sigma_u/\sigma_v \Rightarrow \sigma_v = \sigma_u/\lambda$ ....(2)

Substitute (2) in (1) obtains $\sigma^2 = \sigma_u^2 + (\sigma_u^2/\lambda^2) = \sigma_u^2(1+1/\lambda^2)$

$\Rightarrow \sigma_u^2/\sigma^2 = \lambda^2/(1+\lambda^2)$ ....(3)

Therefore, $\mu_{i*} = e\lambda^2/(1+\lambda^2)$

Next, $\sigma_*^2 = \sigma_u^2\sigma_v^2/\sigma^2$

Substitute (1) in above obtains $\sigma_*^2 = \sigma_u^2(\sigma^2 - \sigma_u^2)/\sigma^2$

Substitute (3) in above obtains $\sigma_*^2 = \lambda^2/(1+\lambda^2) * (\sigma^2 - \lambda^2\sigma^2/(1+\lambda^2))$

$= \lambda^2/(1+\lambda^2) * ((1+\lambda^2-\lambda^2)/(1+\lambda^2))\sigma^2 = \lambda^2\sigma^2/(1+\lambda^2)^2$

Therefore, $\mu_{i*}^2 = e\lambda^2/(1+\lambda^2)$ and $\sigma_*^2 = \lambda^2\sigma^2/(1+\lambda^2)^2$

Q.E.D.

Appendix F: TSP Program Code for Stochastic Frontier

```
? This is batch file to run TSP. Any line starting with (?) is a remark
line and will not be run.
? Any line that ends (;) is recognized as (FORTRAN) command and will be
executed.
?
OPTIONS CRT, MEMORY =8;
READ (FILE='INITIAL.WKS') AHAID,HOSPID,VC,wL,wM,yK,yD,yA,yY;
?
? After reading the file, generate the variable.
?
? GENR Restricted Cost Frontier Variables,
? NOTE: VCM and L differ from the Unrestricted Set
?
GENR VCM=LOG(VC/wM);     ? Note This
GENR L=LOG(wL/wM);       ? Note This
GENR K=LOG(yK);
GENR D=LOG(yD);
GENR A=LOG(yA);
GENR Y=LOG(yY);
GENR LL=0.5*L*L;         ? Note the factor of 0.5
GENR KD=K*D;
GENR KA=K*A;
GENR KY=K*Y;
GENR DA=D*A;
GENR DY=D*Y;
GENR AY=A*Y;
GENR KK=0.5*K*K;         ? Note the factor of 0.5
GENR DD=0.5*D*D;         ? Note the factor of 0.5
GENR AA=0.5*A*A;         ? Note the factor of 0.5
GENR YY=0.5*Y*Y;         ? Note the factor of 0.5
GENR LK=L*K;
GENR LD=L*D;
GENR LA=L*A;
GENR LY=L*Y;
GENR VCMY=VCM-Y ;
?
? Do functional form testing
?
TITLE 'TEST CONSTANT RETURN TO SCALE IN YY' ;
OLSQ(SILENT)VCM
C,L,K,D,A,Y,LL,KD,KA,KY,DA,DY,AY,KK,DD,AA,YY,LK,LD,LA,LY,MK,MD,MA,MY;
SET SSRU=@SSR; SET DFU=@NOB-@NCID;
OLSQ(SILENT) VCMY C,L,K,D,A,LL,KD,KA,DA,KK,DD,AA;
SET SSRR=@SSR; :THERE ARE 13 RESTRICTIONS
SET FSTAT=((SSRR-SSRU)/13)/(SSRU/DFU);
CDF (F,DF1=4,DF2=DFU)FSTAT;
?
TITLE 'TEST COBB DOUGLAS FUNCTION' ;
OLSQ(SILENT)VCM
C,L,K,D,A,Y,LL,KD,KA,KY,DA,DY,AY,KK,DD,AA,YY,LK,LD,LA,LY,MK,MD,MA,MY;
SET SSRU=@SSR; SET DFU=@NOB-@NCID;
OLSQ(SILENT) VCMY C,L,K,D,A,KD,KA,DA,KK,DD,AA;
SET SSRR=@SSR; :THERE ARE 14 RESTRICTIONS
SET FSTAT=((SSRR-SSRU)/14)/(SSRU/DFU);
CDF (F,DF1=4,DF2=DFU)FSTAT;
?
? OLS Restricted Cost Function, Identify Outlier, Conditional Number
?
Title 'OLS Regression for Restricted Cost Function to get Outliers';
OLSQ (ROBUST,HI,SILENT)
```

```
VCM C,L,K,D,A,Y,LL,KD,KA,KY,DA,DY,AY,KK,DD,AA,YY,LK,LD,LA,LY;
SELECT @HI>2*@NCOEFF/@NOB;
GENER DINO=@HI/(@COEFF/@NOB);
TITLE 'WRITE OUTLIER WORKSHEET';
WRITE (FILE=USUAL.WKS) AHAID,@RES,DINO;
?
?
? Check Multicollinearity of OLSQ (Conditional Index)
?
?
CCN;
PROC CCN;
MAT XPX = @VCOV";              ? @VCOV = @S2*(X'X)",XPX = X'X
MAT D = (SQRT(DIAG(XPX)))";
MAT EVAL = EIGVAL(D'XPX*D);
MAT CONDNUM = SQRT(MAX(EVAL)/MIN(EVAL));
PRINT CONDNUM;
ENDPROC;
END;
?
? Eliminate Outlier to get new data. Do stochastic frontier
? See TSP user manual 9.6.4
?
?
GENR VCM=LOG(VC/wM);     ? Note This
GENR L=LOG(wL/wM);       ? Note This
GENR K=LOG(yK);
GENR D=LOG(yD);
GENR A=LOG(yA);
GENR Y=LOG(yY);
GENR LL=0.5*L*L;         ? Note the factor of 0.5
GENR KD=K*D;
GENR KA=K*A;
GENR KY=K*Y;
GENR DA=D*A;
GENR DY=D*Y;
GENR AY=A*Y;
GENR KK=0.5*K*K;         ? Note the factor of 0.5
GENR DD=0.5*D*D;         ? Note the factor of 0.5
GENR AA=0.5*A*A;         ? Note the factor of 0.5
GENR YY=0.5*Y*Y;         ? Note the factor of 0.5
GENR LK=L*K;
GENR LD=L*D;
GENR LA=L*A;
GENR LY=L*Y;
?
FRML    RESID    E=VCM-A0-AL*L-BK*K-BD*D-BA*A-BY*Y-ALL*LL-BKD*KD-BKA*KA-
BKY*KY-BDA*DA-BDY*DY-BAY*AY-BKK*KK-BDD*DD-BAA*AA-BYY*YY-GLK*LK-GLD*LD-
GLA*LA-GLY*LY;
PARAM
A0,AL,BK,BD,BA,BY,ALL,BKD,BKA,BKY,BDA,BDY,BAY,BKK,BDD,BAA,BYY,GLK,GLD,GL
A,GLY;
FRML FRONTP LOGL=LOG(2)+LOG(SIGI)+LNORM(E*SIGI)+LCNORM(E*LAMDA*SIGI);
PARAM LAMDA,SIGI;
EQSUB FRONTP RESID;
REGOPT (PVPRINT STAR) T;
OLS VCM C,L,K,D,A,Y,LL,KD,KA,KY,DA,DY,AY,KK,DD,AA,YY,LK,LD,LA,LY;
UNMAKE @COEFF
A0,AL,BK,BD,BA,BY,ALL,BKD,BKA,BKY,BDA,BDY,BAY,BKK,BDD,BAA,BYY,GLK,GLD,GL
A,GLY;
SET SIGI=1/@S; LAMDA=1;
ML(HITER=N)FRONTP;      ?HITER=N FOR FAST CONVERGENCE
?
? Use Battese and Coelli (1988) Estimator
```

```
?
GENR RESID @RES; ?Get @RES from ML, else will use OLSQ value
SET @S=1/SIGI;
GENR MUS=@RES*LAMDA**2/(1+LAMDA**2);      ? SEE Appendix E
SET SIGS=LAMDA/(SIGI*(1+LAMDA**2));       ? SEE Appendix E
GENR U=LCNORM((MUS/SIGS)-SIGS)+(-MUS+0.5*SIGS**2)-LCNORM(MUS/SIGS);
GNER CE=EXP(U);
WRITE (FILE=FINALCE.WKS) HOSPID, AHAID, CE;
?
?
TITLE 'LOG LIKELIHOOD TEST OF RESTRICTION';
?
? Run the restricted cost function first because already there.
? This is basically repeating the SFE without the OLS start value
?
FRML  RESID2  E2=VCM-A2-AL*L-BK*K-BD*D-BA*A-BY*Y-ALL*L*L-BKD*K*D-BKA*K*A-
BKY*K*Y-BDA*D*A-BDY*D*Y-BAY*A*Y-BKK*K*K-*BDD*D*D-BAA*A*A-BYY*Y*Y-
GLK*L*K-GLD*L*D-GLA*L*A-GLY*L*Y;
PARAM
A2,AL,BK,BD,BA,BY,ALL,BKD,BKA,BKY,BDA,BDY,BAY,BKK,BDD,BAA,BYY,GLK,GLD,GL
A,GLY;
FRML FRONTP2
LOGL=LOG(2)+LOG(SIGI2)+LNORM(E2*SIGI2)+LCNORM(E2*LAMDA2*SIGI2);
PARAM LAMDA2,SIGI2;
EQSUB FRONTP2 RESID2;
SET SIGI2=1/@S; LAMDA2=1;
ML(HITER=N)FRONTP2;      ?HITER=N FOR FAST CONVERGENCE
SET L0=@LOGL; PRINT L0;
?
? Need to GENR the variables again because some items are different
? with the unrestricted form.
? Symmetry will be imposed because A*B=B*A mathematically
?
GENR VC=LOG(VC);
GENR L=LOG(wL);
GENR M=LOG(wM);
GENR K=LOG(yK);
GENR D=LOG(yD);
GENR A=LOG(yA);
GENR Y=LOG(yY);
?
FRML RESID3 E3=VC-A3-AL*L-AM*M-BK*K-BD*D-BA*A-BY*Y-
0.5*(ALL*LL+AMM*MM+2*AML*ML)-BKD*KD-BKA*KA-BKY*KY-BDA*DA-BDY*DY-BAY*AY-
0.5*(BKK*KK+*BDD*DD+BAA*AA+BYY*YY)-GLK*LK-GLD*LD-GLA*LA-GLY*LY-GMK*MK-
GMD*MD-GMA*MA-GMY*MY;
PARAM
A3,AL,AM,BK,BD,BA,GLK,GLD,GLA,GLY,BY,ALL,AMM,AML,BKD,BKA,BKY,BDA,BDY,BAY
,BKK,BDD,BAA,BYY,GLK,GLD,GLA,GLY,GMK,GMD,GMA,GMY;
FRML FRONTP3
LOGL=LOG(2)+LOG(SIGI)+LNORM(E3*SIGI)+LCNORM(E3*LAMDA3*SIGI);
PARAM LAMDA3,SIGI3;
EQSUB FRONTP3 RESID3;
SET SIGI3=1/@S; LAMDA3=1;
ML(HITER=N)FRONTP3;      ?HITER=N FOR FAST CONVERGENCE
SET L1=@LOGL; PRINT L1;
?
?
?
```

Appendix G: Stochastic Frontier Parameters

DESCRIPTION	Coefficient	First Stage	p-value	Second Stage	p-value
Constant	A0	0.0925	0.966	0.0343	0.994
Labor	AL	0.9613	0.000***	0.9642	0.000***
Capital	BK	1.2558	0.056*	1.1675	0.225
1/RAMI	BD	-1.3502	0.115	-1.2049	0.502
1/Complication Rate	BA	0.4591	0.175	0.4539	0.312
DRG weighted Discharge	BY	0.1797	0.708	0.1501	0.934
Labor cross product	ALL	-0.0612	0.350	-0.0560	0.376
Capital*1/RAMI	BKD	0.1943	0.179	0.1664	0.489
Capital*1/Complication Rate	BKA	-0.0226	0.691	-0.0195	0.757
Capital*DRG weighted Discharge	BKY	-0.3656	0.005***	-0.3439	0.088*
1/RAMI*1/Complication Rate	BDA	0.1598	0.046**	0.1639	0.139
1/RAMI*DRG weighted Discharge	BDY	0.1180	0.368	0.0923	0.808
1/Complication Rate*DRG weighted Discharge	BAY	-0.1008	0.176	-0.1043	0.297
Capital cross product	BKK	0.3810	0.014***	0.3766	0.031**
1/RAMI cross product	BDD	-0.3540	0.126	-0.2949	0.493
1/Complication rate cross product	BAA	0.0122	0.816	0.0151	0.779
DRG weighted discharge cross product	BYY	0.2382	0.000***	0.2415	0.555
Labor*Capital	GLK	0.0454	0.456	0.0305	0.657
Labor*1/RAMI	GLD	0.0311	0.543	-0.0005	0.734
Labor*1/Complication Rate	GLA	-0.0155	0.932	-0.0191	0.845
Labor*DRG weighted discharge	GLY	-0.0435	0.548	0.0004	0.458
Error Standard Deviation = SQRT($\sigma_u^2+\sigma_v^2$)	SIGI	-	-	14.4884	0.000***
Ratio of sigma(u)/sigma(v)	LAMDA	-	-	3.4023	0.001***

Appendix H: Marginal Costs

The following table contains the calculated marginal costs for the 313 hospitals in our sample. We report the marginal cost for quantity output as MCY and the cost elasticity of 1/RAMI as ED, and cost elasticity of 1/(surgical complication) as EA. We can easily convert elasticity to marginal cost by: marginal cost = elasticity*VC/quality. Since we need to know only the sign of the marginal cost, checking the sign of the elasticity is sufficient. As it is hard to interpret a unit of the quality measures, the marginal costs are hard to interpret. For the quantity output (Y), we expect most marginal costs to exceed one for several reasons: (1) the cost captures physician component while the quantity does not (2) the cost captures outpatient component while the quantity does not.

SN	yD	yA	VC	ED	EA	MCY
1	0.5703	0.0800	158,474,514	0.93	-1.66995	3.63
2	1.4538	0.0420	90,195,664	1.05	-1.53005	3.49
3	1.3990	0.0437	37,453,736	0.45	-1.43228	3.38
4	0.4056	0.0200	19,065,661	0.56	-1.48024	5.80
5	1.7439	0.0392	76,459,886	0.83	-1.62499	4.80
6	1.2372	0.0679	80,176,093	1.00	-1.79026	2.89
7	1.4336	0.0282	124,678,871	0.66	-1.55209	4.55
8	1.4737	0.0444	84,696,074	0.11	-1.26057	3.13
9	0.4367	0.0456	24,006,694	1.07	-1.72497	7.72
10	0.7708	0.0210	42,198,176	0.86	-1.59723	4.50
11	1.3926	0.0718	23,737,445	0.47	-1.48977	4.43
12	1.2596	0.0755	267,873,021	0.64	-1.47913	5.67
13	1.5199	0.0318	115,526,140	0.46	-1.39606	5.52
14	3.7483	0.0370	76,659,456	0.89	-1.57149	4.31
15	1.6755	0.0769	76,451,108	0.88	-1.55342	4.89
16	1.0457	0.0214	47,847,324	1.02	-1.65005	5.16
17	1.3331	0.0782	75,860,409	1.02	-1.67732	3.01
18	0.7292	0.0489	29,335,800	0.94	-1.61046	5.48
19	0.5372	0.0384	17,382,482	0.68	-1.52939	5.42
20	5.0237	0.0485	213,641,602	0.78	-1.65424	3.87
21	0.9543	0.0231	32,597,682	1.03	-1.67636	10.20
22	0.6977	0.0375	57,173,251	1.06	-1.68339	5.77
23	1.6666	0.0516	74,182,636	0.99	-1.7044	3.03
24	2.2712	0.0354	41,023,209	0.54	-1.44788	5.28
25	1.3081	0.0508	102,934,316	0.90	-1.70441	2.39
26	2.0683	0.0184	111,255,361	0.31	-1.4704	7.85
27	0.7762	0.0448	43,167,761	0.81	-1.6136	3.60
28	0.4070	0.0304	15,247,902	0.84	-1.69709	5.72
29	1.4807	0.0272	353,744,328	0.60	-1.5387	7.26
30	1.0569	0.0252	70,435,506	0.51	-1.43004	4.08
31	0.9520	0.0720	33,827,314	0.99	-1.56872	4.48
32	0.8500	0.0659	24,902,874	0.82	-1.63428	5.60
33	1.9624	0.0570	186,412,627	0.74	-1.57509	2.48
34	0.8723	0.0444	81,851,904	0.73	-1.47486	4.22
35	1.1943	0.0454	83,639,661	0.75	-1.53984	5.02

SN	yD	yA	VC	ED	EA	MCY
36	2.4952	0.0565	68,467,422	0.75	-1.52832	1.63
37	0.5740	0.0745	77,246,501	1.22	-1.77695	3.10
38	0.8397	0.0173	245,008,765	0.19	-1.32252	2.97
39	0.8006	0.0337	50,445,586	1.04	-1.70375	3.23
40	1.1337	0.0296	62,211,827	0.82	-1.58291	2.55
41	1.2793	0.0386	212,800,460	1.13	-1.80204	2.97
42	2.2822	0.0216	137,326,473	0.92	-1.75317	3.20
43	1.6922	0.0965	16,233,001	0.72	-1.5738	2.45
44	0.5473	0.0296	157,603,999	0.45	-1.36819	1.96
45	0.6842	0.0213	28,040,675	0.58	-1.46868	4.88
46	0.3629	0.0220	27,408,125	1.04	-1.6767	5.22
47	1.4827	0.0205	40,418,546	0.43	-1.44401	4.42
48	0.5802	0.0281	29,544,415	1.11	-1.69644	5.26
49	1.8227	0.0574	152,811,651	0.53	-1.4942	3.57
50	1.0928	0.0325	51,027,938	1.07	-1.71233	3.50
51	0.5050	0.0084	14,840,393	0.68	-1.68712	6.77
52	2.3964	0.0309	44,935,751	0.52	-1.46003	3.53
53	1.1225	0.0568	188,635,764	0.63	-1.60697	1.93
54	1.2287	0.0304	77,979,737	1.28	-1.77864	3.74
55	0.3643	0.0321	29,086,849	0.95	-1.72879	2.25
56	1.1204	0.0468	55,432,034	0.88	-1.71757	2.96
57	0.6552	0.0408	79,762,616	0.63	-1.50289	2.27
58	1.3101	0.0348	51,048,289	0.74	-1.46828	2.98
59	1.0944	0.0517	41,436,619	0.83	-1.63124	4.79
60	0.5917	0.0336	62,934,807	0.79	-1.6148	2.17
61	0.4596	0.0273	26,660,429	1.69	-2.11186	2.31
62	0.5846	0.1514	253,132,544	0.91	-1.81188	3.91
63	0.8018	0.0137	27,015,749	0.90	-1.72561	4.66
64	0.4799	0.1755	22,205,365	1.21	-1.79344	6.47
65	1.1642	0.1077	44,066,176	1.07	-1.79731	3.82
66	0.7949	0.0356	120,680,537	0.58	-1.43871	4.26
67	1.3361	0.0324	217,362,484	0.24	-1.35093	4.85
68	0.9845	0.0307	129,905,383	0.46	-1.39606	4.94
69	0.4002	0.0549	29,589,753	0.97	-1.62779	5.51
70	1.3430	0.0497	114,100,361	1.11	-1.6886	4.40
71	1.1402	0.3760	20,689,820	0.92	-1.71346	6.08
72	1.5786	0.0339	76,737,103	1.05	-1.7164	5.62
73	0.6741	0.0277	78,916,419	0.95	-1.79711	5.62
74	0.8635	0.0304	45,898,426	1.13	-1.77818	3.88
75	1.0148	0.0367	65,530,908	1.10	-1.72662	4.28
76	0.7634	0.0768	39,591,528	0.76	-1.55638	4.69
77	0.9461	0.0403	24,888,803	0.84	-1.52829	5.68
78	0.4715	0.0517	113,833,694	0.79	-1.53958	5.58
79	0.7841	0.0303	155,026,761	0.93	-1.66222	4.21
80	1.3203	0.0276	165,233,702	0.75	-1.5973	5.02
81	0.8753	0.1824	166,264,057	1.05	-1.73901	4.36
82	0.2003	0.0340	19,817,935	1.07	-1.69522	3.74
83	1.5702	0.0204	18,535,898	0.82	-1.6129	3.19
84	1.0232	0.0951	80,679,647	0.79	-1.65547	1.91
85	0.7922	0.0484	39,167,663	0.61	-1.56275	2.32
86	0.2092	0.0695	26,975,447	0.64	-1.4973	1.70

SN	yD	yA	VC	ED	EA	MCY
87	1.2881	0.0520	167,717,800	0.85	-1.7007	1.40
88	1.0728	0.1029	41,949,518	0.82	-1.58636	1.15
89	0.5603	0.0847	34,548,177	0.94	-1.58986	1.72
90	1.9128	0.0469	331,136,650	0.48	-1.426	1.92
91	1.5732	0.0782	35,206,573	0.55	-1.4348	2.10
92	0.5726	0.0169	61,600,188	0.64	-1.44061	3.98
93	0.3188	0.0320	29,374,571	0.46	-1.42602	1.96
94	0.7001	0.0770	44,860,524	1.11	-1.77843	2.20
95	7.2852	0.0349	31,529,587	0.75	-1.48582	1.77
96	0.5311	0.0346	77,411,489	0.77	-1.54195	2.53
97	0.5012	0.0371	70,224,305	1.21	-1.79929	1.23
98	1.9936	0.0723	88,129,940	1.09	-1.78426	3.50
99	0.8232	0.0421	61,782,849	0.87	-1.67004	2.14
100	2.3839	0.0191	116,658,163	0.69	-1.56172	1.89
101	0.2137	0.0194	20,461,760	0.71	-1.46792	2.51
102	1.6449	0.0422	98,008,977	0.94	-1.70321	2.38
103	0.5699	0.0414	94,378,095	1.20	-1.7526	1.26
104	1.3406	0.0523	15,695,878	0.59	-1.54846	1.82
105	0.6944	0.0253	63,139,100	0.85	-1.59501	2.01
106	0.6386	0.0525	82,218,072	0.51	-1.56607	3.73
107	0.6116	0.0271	241,419,174	0.44	-1.43474	3.01
108	0.7764	0.0219	218,167,583	1.17	-1.81916	2.10
109	2.1817	0.0451	151,082,478	0.10	-1.21298	1.79
110	0.5757	0.0704	34,132,135	0.81	-1.70552	2.96
111	0.3655	0.0445	61,279,231	0.95	-1.64426	1.59
112	0.5947	0.0340	20,911,024	0.66	-1.53128	2.12
113	0.5256	0.0553	25,032,570	0.64	-1.44528	2.26
114	0.6819	0.0260	267,299,534	0.78	-1.30322	2.24
115	1.1528	0.0227	76,444,000	0.76	-1.58053	1.27
116	0.4577	0.0417	53,404,969	1.04	-1.67259	1.44
117	0.9843	0.0396	220,582,187	0.82	-1.57849	2.63
118	2.2904	0.0186	43,684,851	1.21	-1.62794	1.39
119	2.6759	0.0491	161,766,725	0.70	-1.50416	2.25
120	1.4471	0.0243	281,712,372	0.80	-1.52783	2.25
121	4.1289	0.0249	62,301,764	0.75	-1.5254	1.94
122	0.5835	0.0117	24,008,823	0.82	-1.4603	2.39
123	0.1649	0.0171	17,953,668	0.89	-1.61879	1.49
124	0.3373	0.0542	36,887,530	0.80	-1.50083	2.37
125	0.7974	0.0313	693,911,465	0.87	-1.53622	2.05
126	1.2160	0.0296	88,025,077	0.79	-1.60218	2.18
127	0.6458	0.0400	45,098,960	0.57	-1.51934	1.70
128	2.7378	0.0434	46,087,446	0.55	-1.446	1.48
129	0.5104	0.0654	73,112,108	0.84	-1.57732	1.60
130	0.6657	0.0231	56,470,327	1.03	-1.60894	1.36
131	0.6469	0.0604	66,916,064	0.65	-1.51495	2.22
132	1.8647	0.0354	37,566,652	0.42	-1.32096	2.01
133	1.3013	0.0459	82,785,244	0.95	-1.6069	2.10
134	1.2875	0.0371	99,700,533	0.76	-1.5614	1.48
135	0.8958	0.0292	56,124,354	0.85	-1.59332	1.82
136	0.9217	0.0435	37,228,499	0.74	-1.5199	1.53
137	0.6643	0.0780	79,620,986	0.76	-1.52918	1.71

SN	yD	yA	VC	ED	EA	MCY
138	1.1392	0.0748	226,490,791	0.75	-1.5462	2.30
139	1.3469	0.0175	58,634,790	0.74	-1.50738	1.80
140	0.7805	0.0634	67,581,508	0.72	-1.49689	2.11
141	2.3424	0.0252	81,230,454	0.32	-1.36639	3.57
142	1.4298	0.0459	107,308,900	0.18	-1.33704	2.33
143	2.1220	0.0580	189,846,388	0.56	-1.4684	2.05
144	0.3997	0.0432	37,053,093	0.63	-1.41332	2.04
145	1.1056	0.1037	30,620,085	0.66	-1.50057	1.35
146	1.0863	0.0745	27,075,799	0.51	-1.33074	2.12
147	0.9534	0.0200	127,024,799	0.38	-1.32548	2.65
148	7.4657	0.0320	117,863,967	0.90	-1.5517	2.12
149	1.5467	0.0163	65,402,762	0.73	-1.51892	3.03
150	1.1455	0.0384	66,887,840	0.65	-1.44461	3.47
151	0.5573	0.0253	72,206,977	0.46	-1.42997	4.92
152	1.0432	0.0260	31,317,533	0.54	-1.46439	1.82
153	0.7512	0.0269	69,055,394	0.80	-1.44032	1.38
154	1.4851	0.0383	90,405,237	0.36	-1.36443	1.36
155	1.3057	0.0488	50,684,744	0.75	-1.5215	2.11
156	0.9421	0.0445	190,107,291	0.87	-1.5921	2.52
157	0.5701	0.0474	42,436,156	0.84	-1.52663	5.13
158	1.2217	0.0285	126,507,069	0.65	-1.44711	2.50
159	1.3732	0.0741	170,347,319	0.60	-1.41981	3.10
160	0.7683	0.0382	48,221,697	0.75	-1.46213	4.19
161	1.6854	0.0723	28,926,665	0.64	-1.40506	3.59
162	1.2304	0.0302	34,896,170	0.59	-1.47281	4.29
163	0.6988	0.0277	96,116,792	0.97	-1.66383	2.22
164	0.7036	0.0641	86,140,297	0.77	-1.56491	3.41
165	1.6604	0.0270	40,923,251	0.96	-1.68498	2.93
166	0.4822	0.0442	83,703,927	1.15	-1.70604	2.37
167	1.2281	0.0366	123,748,092	0.96	-1.67339	2.62
168	0.2998	0.0503	23,988,587	1.04	-1.65354	5.01
169	1.0194	0.0225	66,822,601	0.75	-1.63809	3.64
170	0.5453	0.0460	25,096,299	0.86	-1.54797	6.19
171	1.2083	0.0228	191,690,947	0.86	-1.57547	2.68
172	1.7115	0.0439	110,526,448	0.95	-1.62722	1.97
173	0.7596	0.0285	27,580,050	0.92	-1.66734	3.02
174	3.5733	0.2848	47,378,928	0.78	-1.58756	2.62
175	0.5818	0.0227	80,394,470	0.91	-1.59052	4.20
176	0.6077	0.0282	22,488,374	1.20	-1.82404	4.55
177	0.7232	0.0484	52,549,558	0.50	-1.48332	4.04
178	0.9907	0.0219	7,754,245	0.29	-1.37786	4.62
179	0.7816	0.0077	35,830,023	1.07	-1.70077	3.72
180	1.8724	0.0272	87,123,173	0.95	-1.57789	3.86
181	0.9449	0.0227	70,534,145	1.15	-1.91464	3.89
182	0.2227	0.0549	6,018,196	0.84	-1.62157	7.17
183	0.8760	0.0288	21,706,672	0.89	-1.64285	3.86
184	0.1178	0.0346	10,192,226	1.40	-1.93317	5.69
185	0.3061	0.0173	6,999,150	0.81	-1.59164	6.72
186	1.2491	0.0522	181,733,527	0.91	-1.60541	4.42
187	0.8880	0.0104	164,018,396	1.06	-1.62977	3.28
188	1.4787	0.0403	200,322,696	0.90	-1.5908	3.47

SN	yD	yA	VC	ED	EA	MCY
189	1.1946	0.0290	66,621,253	1.17	-1.75428	3.90
190	0.7318	0.0278	10,451,959	1.10	-1.72549	5.63
191	0.7189	0.0306	64,109,132	1.09	-1.70336	3.60
192	1.2410	0.0958	215,543,232	0.98	-1.63653	3.00
193	2.4321	0.0728	137,203,542	0.78	-1.52712	3.63
194	1.1004	0.0323	99,259,020	0.33	-1.34136	3.20
195	0.8794	0.0192	54,431,865	0.47	-1.49864	6.70
196	0.2525	0.0400	10,904,187	0.80	-1.53883	5.68
197	0.4218	0.0119	36,702,589	0.71	-1.52558	4.42
198	0.1984	0.0317	16,357,226	0.48	-1.42585	4.88
199	0.6369	0.0156	41,498,582	0.51	-1.35735	3.93
200	0.4277	0.0330	6,476,922	0.83	-1.57216	5.21
201	0.5378	0.0410	43,176,233	0.86	-1.57229	4.99
202	0.5407	0.0316	16,376,788	0.84	-1.58554	5.68
203	1.0304	0.0507	159,318,683	1.07	-1.68722	3.97
204	0.3034	0.0186	17,183,295	1.08	-1.66599	6.16
205	0.3126	0.0209	12,387,460	0.75	-1.61094	7.24
206	2.2147	0.0329	104,446,076	0.62	-1.45551	3.65
207	0.4354	0.0184	28,191,395	0.47	-1.3657	4.99
208	1.0510	0.0740	216,171,927	0.52	-1.49822	2.98
209	0.8110	0.0389	56,561,347	0.87	-1.6809	3.73
210	0.4727	0.0467	19,693,316	1.14	-1.73626	5.13
211	1.0820	0.0179	102,622,359	0.77	-1.53099	4.85
212	0.6166	0.0586	151,491,100	0.83	-1.62937	3.85
213	0.8979	0.0526	41,420,842	0.77	-1.53199	3.34
214	0.8522	0.0258	54,563,179	0.70	-1.5503	3.48
215	0.1483	0.0571	3,383,650	0.78	-1.59098	12.75
216	0.8355	0.0227	18,493,690	0.82	-1.52707	4.82
217	0.9451	0.0421	22,909,902	1.05	-1.72371	5.44
218	1.2280	0.0330	23,326,095	0.51	-1.42362	3.90
219	0.3875	0.0382	36,486,995	0.77	-1.53814	5.48
220	0.1406	0.1724	5,484,292	0.84	-1.69687	10.42
221	0.1723	0.0075	4,373,525	0.96	-1.63291	6.88
222	0.8002	0.0122	96,709,415	0.97	-1.61246	3.92
223	1.5255	0.0188	74,177,413	0.91	-1.69403	4.08
224	0.8965	0.0229	157,242,732	0.57	-1.29466	3.94
225	0.4687	0.0200	37,410,890	0.64	-1.55371	3.68
226	0.5287	0.0493	14,579,969	0.86	-1.62718	4.15
227	0.7483	0.0516	69,811,312	0.48	-1.39848	4.42
228	0.8486	0.0221	107,983,594	0.82	-1.62966	3.49
229	0.4536	0.0189	68,707,924	0.88	-1.6803	2.91
230	0.7935	0.0524	75,355,149	0.67	-1.4765	2.83
231	1.5422	0.0124	161,484,018	0.48	-1.43387	2.86
232	0.6676	0.0322	14,956,215	0.84	-1.56929	2.63
233	0.4584	0.0400	21,907,218	0.47	-1.38403	5.20
234	1.7991	0.0363	52,819,195	0.81	-1.52075	3.64
235	2.6098	0.0523	17,958,460	0.83	-1.58976	3.98
236	0.9147	0.0406	92,541,430	0.62	-1.55358	2.81
237	2.2999	0.1429	19,942,252	1.17	-1.71742	4.07
238	1.7417	0.0473	135,851,585	0.71	-1.57251	2.72
239	1.7787	0.0298	171,402,325	0.51	-1.49829	2.51

SN	yD	yA	VC	ED	EA	MCY
240	1.1641	0.0695	302,642,769	1.00	-1.46233	3.30
241	0.9006	0.0264	53,475,436	0.70	-1.47625	3.47
242	0.7053	0.0148	165,112,636	0.68	-1.62302	3.90
243	0.5111	0.0435	166,879,762	0.43	-1.48231	2.23
244	1.2518	0.0355	41,845,149	0.38	-1.39035	2.28
245	0.7349	0.0799	22,073,721	0.71	-1.38619	2.24
246	0.9299	0.0243	92,561,832	0.44	-1.38155	3.02
247	0.0135	0.0330	45,328,434	1.12	-1.68151	8.15
248	1.9988	0.0199	39,458,480	0.93	-1.66541	2.40
249	0.5986	0.0466	85,198,148	0.52	-1.39046	2.78
250	1.2443	0.0284	89,074,660	0.82	-1.44918	1.84
251	0.7977	0.0378	28,454,716	0.27	-1.37816	3.41
252	0.1602	0.2388	6,931,570	0.60	-1.44436	4.47
253	0.0279	0.0833	3,026,676	0.67	-1.51753	14.35
254	0.1940	0.0123	5,443,665	0.57	-1.4424	9.42
255	0.5789	0.0364	156,026,780	0.35	-1.48895	4.80
256	0.6783	0.0545	10,671,580	1.25	-1.74024	5.17
257	0.6884	0.0300	74,787,502	0.71	-1.51382	4.16
258	0.4033	0.0286	14,821,839	0.66	-1.46556	6.24
259	0.2323	0.0028	13,658,695	0.32	-1.35706	8.60
260	0.8092	0.0229	41,443,309	0.80	-1.61813	3.90
261	0.8863	0.0196	120,139,256	0.82	-1.57528	4.79
262	1.5597	0.0256	381,259,813	0.69	-1.53177	6.19
263	0.7448	0.0272	41,598,523	0.21	-1.21718	4.66
264	0.7262	0.0273	41,691,896	0.82	-1.64794	4.39
265	1.2468	0.0406	230,104,318	0.77	-1.54165	3.45
266	0.7820	0.0350	65,902,381	0.68	-1.62936	4.19
267	1.0125	0.0661	26,195,545	0.50	-1.20651	6.00
268	0.8406	0.0321	171,437,070	0.49	-1.40263	4.23
269	0.2894	0.0330	52,989,986	1.16	-1.6707	3.04
270	1.1328	0.0355	33,138,751	0.95	-1.71761	5.06
271	0.6442	0.0325	18,280,957	0.84	-1.6283	4.14
272	0.3694	0.0371	75,755,173	0.81	-1.57371	3.21
273	0.5395	0.0235	23,542,411	0.69	-1.61059	3.12
274	1.3212	0.0713	239,277,096	0.57	-1.46622	2.36
275	1.3781	0.0102	23,382,127	0.63	-1.50013	1.82
276	0.4477	0.0277	45,619,636	0.59	-1.45042	1.64
277	1.3063	0.0319	45,717,438	0.59	-1.4747	2.82
278	1.2432	0.0403	63,798,724	0.64	-1.404	1.70
279	0.5781	0.0183	27,544,846	0.95	-1.61441	1.82
280	1.0754	0.0338	185,322,235	0.86	-1.67583	2.03
281	0.3365	0.0261	7,585,813	0.50	-1.41445	3.69
282	0.6181	0.0623	28,340,658	0.76	-1.56595	2.46
283	1.2600	0.0629	21,135,624	0.72	-1.56831	3.04
284	1.0093	0.0396	172,319,990	0.76	-1.5716	2.19
285	1.2138	0.0813	173,951,536	0.83	-1.7428	1.95
286	0.9219	0.0480	50,410,898	1.00	-1.67876	1.95
287	0.2372	0.0106	12,362,746	0.97	-1.76899	1.83
288	2.2761	0.0340	19,443,030	0.66	-1.48585	3.81
289	0.7090	0.0619	102,951,105	1.06	-1.64435	2.39
290	0.6517	0.0310	300,967,962	1.00	-1.73789	3.06

SN	yD	yA	VC	ED	EA	MCY
291	1.0360	0.0439	219,456,957	0.61	-1.46181	2.00
292	1.4129	0.0442	19,924,958	0.38	-1.45614	1.51
293	0.4808	0.0605	88,516,442	0.41	-1.38123	2.73
294	0.7549	0.0180	167,198,552	0.50	-1.34215	2.21
295	0.7682	0.0225	57,316,227	0.98	-1.69656	1.61
296	0.6696	0.0598	94,214,388	0.80	-1.59865	2.00
297	0.3529	0.0591	25,332,373	0.54	-1.51974	3.43
298	2.5281	0.0368	125,762,607	0.30	-1.42406	1.75
299	0.7640	0.0540	243,722,350	0.35	-1.41885	2.31
300	0.2761	0.0195	19,501,526	0.70	-1.52464	4.39
301	0.6687	0.0881	37,844,128	0.90	-1.67422	4.98
302	2.7812	0.0402	158,365,952	0.65	-1.43891	2.05
303	1.6049	0.0828	78,479,873	0.11	-1.29295	1.69
304	0.6488	0.0141	9,314,905	0.49	-1.32932	5.30
305	2.6352	0.0383	73,299,456	0.68	-1.51337	3.20
306	0.8570	0.0197	120,290,185	1.11	-1.82844	2.36
307	2.1509	0.0459	109,546,022	0.68	-1.50598	2.53
308	0.4228	0.0664	24,806,730	0.68	-1.58068	1.29
309	0.8009	0.0521	40,525,088	0.72	-1.53393	3.30
310	0.5294	0.0547	57,221,012	0.52	-1.48396	2.40
311	0.9507	0.0831	119,424,276	1.02	-1.69139	1.85
312	0.2202	0.0402	19,748,018	0.90	-1.56367	4.97
313	0.7536	0.0464	72,008,595	0.79	-1.63593	3.89

Bibliography

Advisory Board (1993). The Grand Alliance. Washington, DC: Advisory Board.

Afriat S. (1972). Efficiency Estimation of Production Functions. International Economic Review.

Aigner D.J., Lovell C.A.K. and Schmidt P. (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. Journal of Econometrics. Vol. 6: 21-37.

Alchian A.A. (1961). Some Economics of Property Rights. Santa Monica, California: RAND Corp.

Alchian A.A. and Demsetz H. (1972). Production, Information Costs and Economic Organization. American Economic Review. Vol. 62(5): 777-95.

Allen R.G.D. (1938). Mathematical Analysis for Economics. MacMillan, London.

Ament R.P. (1976a). Resource Need Index and Average Charge per Resource Need Unit: Distribution in SPC Hospitals. PAS Reporter. Vol.14(8): 1-7.

Ament R.P. (1976b). The Use of Casemix Figures in Analyzing Average Charges for Inpatients. PAS Reporter. Vol. 14(3): 1-5.

Arrow K.J. (1963). Uncertainty and the Welfare Economics of Medical Care. American Economic Review. Vol. 53(5):941-73.

Arrow K.J., Chenery H.B., Minhas B.S. and Solow R.M. (1961). Capital Labor Substitution and Economic Efficiency. Review of Economics and Statistics. Vol. 43(3) August: 225-50.

Barer M.L. (1982). Casemix Adjustment in Hospitals Cost Analysis. Journal of Health Economics. Vol. 1: 53-80.

Barten A.P. (1969). Maximum Likelihood Estimation of a Complete System of Demand Equations. European Economic Review, Vol. 1: 7-73.

Battese G.E. and Coelli T.J. (1988). Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data. Journal of Econometrics, Vol. 38:387-99.

Battese G.E. and Coelli T. (1992). Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Framers in India. Journal of Productivity Analysis. Vol. 3(1/2): 153-69.

Bauer P.W. (1985). An Analysis of Multi-product Technology and Efficiency using Joint Cost Function and Panel Data: An Application to the U.S. Airline Industry. Unpublished Ph.D. Dissertation, University of North Carolina, Chapel Hill, NC.

Bauer P.W. (1990). Recent Development in the Econometric Estimation of Frontiers. Journal of Econometrics. Vol. 46: 39-56).

Baumol W.J., Panzar J.C. and Willig R.D. (1982). Contestable Market and the Theory of Industry Structure. Harcourt Brace Jovanovich, New York.

Bays C.W. (1980). Specification Error in the Estimation of Hospitals Cost Functions. Review of Economics and Statistics. Vol. 62(2):302-5.

Becker E.R. and Sloan A.S. (1985). Hospitals Ownership and Performance. Economic Inquiry. Vol. 23: 21-36.

Becker E.R. and Steinwald B. (1981). Determinants of Hospitals Casemix Complexity. Health Service Research. Vol. 16(4): 439-58.

Becker G.S. (1975). Human Capital: A Theoretical and Empirical Analysis with special reference to Training (2nd ed). Columbia University Press, New York.

Belsley D.A. Kuh E. and Welsch R.E. (1980). Regression Diagnostics: Identifying Influential Observations and Sources of Collinearity. John Wiley & Sons, New York.

Berger A.N. 91993). Distribution Free Estimates of Efficiency in the U.S. Banking Industry and Tests of the Standard Distribution Assumptions. Journal of Productivity Analysis, Vol. 4(3) September: 261-92.

Berger A.N. and Humphrey D.B. (1991). The Dominance of Inefficiencies over Scale and Product Mix Economies in Banking. Journal of Monetary Economics, Vol. 28: 117-48.

Berger A. N and Humphrey D.B. (1992). Measurement and Efficiency Issues in Commercial Banking. In Griliches Z. (ed.) Output Measurement in the Service Sectors. National Bureau of Economics Research Studies in Income and Wealth, Chicago: University of Chicago Press.

Berki S.E. (1972). Hospital Economics. Lexington Books, Massachusetts.

Berndt E.R. (1976). Reconciling Alternative Estimates of the Elasticity of Substitution. Review of Economics and Statistics. Vol. 58(1): 59-68.

Berndt E.R. (1991). The Practice of Econometrics: Classic and Contemporary. Addison-Wesley.

Berndt, E.R. and Hesse, D. (1986). Measuring and Assessing Capacity Utilization in the Manufacturing Sectors of Nine OECD Countries. European Economic Review, Vol. 30: 961-89.

Bilodeau D., Cremieux P.Y. and Ouellette P. (2000). Hospital Cost Function in a Non-Market Health Care System. Review of Economics and Statistics. Vol. 82(3) August: 489-98.

Blackorby C. and Diewert W. E. (1979). Expenditure Functions, Local Duality mad Second Order Approximations. Econometrica. Vol. 47: 579-603.

Blackorby C., Primont D. and Russell R. R. (178). Duality, Separability and Functional Structure: Theory and Economic Applications. North Holland, Amsterdam.

Breyer F. (1987). The specification of a Hospitals Cost Function. Journal of Health Economics. Vol. 6: 147-57.

Brown R.S., Caves D.W. and Christensen L.R. (1979). Modeling the Structure and Cost for Multi-product Firms. Southern Economics Journal. Vol. 46(1): 256-73.

Bureau of Labor Statistics (2001). Wage by Area and Occupation. URL: <http://www.bls.gov.oes/1999/oessrcst.htm> (version current as of May 2001)

Burns L.R., Gimm G. and Nicholson S.(2005). The financial performance of integrated health organizations. Journal of Healthcare Management. Vol. 50(3):191-211

Burns L.R and Thorpe D.P. (1993). Trends and Models in Physician Hospital Organization. Healthcare Management Review. Vol. 18(4): 7-20.

Butler R.R.G. (1995). Hospitals Cost Analysis: development in Health Economics and Public Policy Vol. 3. Kluwer Academic Publishers, Dordrecht, Netherlands.

Cameron J.M. (1985). The Indirect cost of Graduate Medical Education. New England Journal of Medicine. Vol. 312(19): 1233-8.

Carey, K. (1997). A Panel Data Design for Estimation of Hospital Cost Functions. Review of Economics and Statistics. Vol. 79(3): 443-53.

Carlson S. (1939). A Study in the Pure Theory of Production. P.S. King & Son: London.

Cassels J.M. (1936). On the Law of Variable Proportions, Exploration in Economics. McGraw Hill, New York: 223-36.

Caves D.W., Christensen L.R. and Tretheway M.W. (1980). Flexible Cost Functions for Multi-product Firms. Review of Economics and Statistics. Vol. 62(3): 477-81.

Chamberlain E. (1933). The Theory of Monopolistic Competition. Harvard University Press: Cambridge MA.

Chambers R.G. (1988). Applied Production Analysis: A Dual Approach. Cambridge University Press, Cambridge.

Charnes A., Cooper W.W. and Rhodes E. (1978). Measuring the Efficiency of Decision Making Units. European Journal of Operational Research. Vol. 2(6): 429-444.

Christensen L.R. and Greene W. (1976). Economies of Scale in US Electric Power Generation. Journal of Political Economy, Vol. 84(4) August: 655-76.

Christensen L.R., Jorgenson D.W. and Lau L.J. (1973). Transcendental Logarithmic Production Function Frontiers. Review of Economics and Statistics. Vol. 55 February: 29-45.

Cobb C.W. and Douglas P.H. (1928). A Theory of Production. American Economic Review. Vol. 18: 139-65.

Conrad R.F. and Strauss R.P. (1983). A Multiple-output Multiple-input Model of the Hospitals Industry in North Carolina. Applied Economics. Vol. 15(3): 341-52.

Cook, T.D. and Campbell, D.T. (1979). Quasi-experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally.

Cornwell C.P., Schmidt P. and Sickles R. (1990). Production Frontiers with Cross sectional and Time Series Variation in Efficiency Levels. Journal of Econometrics. Vol. 46: 185-200.

Cowing T.G. and Holtmann A.G. (1983). Multi-product Short-Run Hospitals Cost Functions: Empirical Evidence and Policy Implications from Cross-Section Data. Southern Economic Journal. Vol. 49(3): 637-53.

Cowing T.G., Holtmann A.G. and Powers S. (1983). Hospitals Cost Analysis: A Survey and Evaluation of Recent Studies. In Scheffler R.M. and Rossiter L.F. (eds.). Advances in Health Economics and Health Services Research. Vol. 4, JAI Press, Connecticut: 257-303.

Cuellar A.E. and Gertler P.J. (2006). Strategic Integration of Hospitals and Physicians. Journal of Health Economics. Vol. 25(1):1-28.

Culyer A.J., Wiseman J., Drummond M.F. and West P.A. (1978). What Accounts for the Higher Cost of Teaching Hospitals? Social and Economic Administration. Vol. 12(1): 20-30.

Cummins J. D. and Zi H. (1998). Measuring Cost Efficiency in the U.S. Life Insurance Industry: Econometric and Mathematical Programming Approaches. Journal of Productivity Analysis 10: 131-152.

De Alessi L. (1983). Property Rights, Transaction Costs and X-Efficiency: An Essay in Economic Theory. American Economic Review. Vol. 73(1): 64-81.

Debreu G. (1951). The Coefficient of Resource Utilization. Econometrica, Vol. 19(3) July: 273-292.

Deeble J.S. (1983). Hospitals Cost and Reimbursement Formula. In Tatchell P.M. (ed.), Economics and Health: 1982 Proceedings of the Forth Australian Conference of Health Economists, Technical Paper No. 3. Health Research Project, Australian National University, Canberra: 320-2.

Desharnais S., Chesney J.D., Wroblewski R.T., Fleming S.T. and McMahon L.F. (1988). The Risk Adjusted Mortality Index: A New Measure of Hospitals Performance. Medical Care. Vol. 26(12): 1129-1148.

Diewert W.E. (1971). An Application of the Shepard Duality Theorem: A Generalized Leontief Production Function. Journal of Political Economy. Vol 79(3), May/June: 481-507.

Diewert W.E. (1973). Functional forms for Profit and Transformation Functions. Journal of Economic Theory. Vol. 6(3) June: 284-316.

Diewert W.E. (1974). Applications of Duality Theory. In Intrilligator M.d. and Kendrick D.A. (eds.), Frontiers of Quantitative Economics, Volume II. North Holland, Amsterdam: 106-71.

Diewert W.E. (1982). Duality Approaches to Microeconomic Theory. In Arrow K.J. and Intrilligator M.D. (eds.), Handbook of Mathematical Economics, Volume II. North Holland, Amsterdam: 535-99.

Donabedian A. (1980). The Definition of Quality and Approaches to its Assessment. Ann Arbor: Health Administration Press.

Donabedian A. (1988). Quality and Cost: Choices and Responsibilities. Inquiry. Vol. 31: 90-9.

Donabedian A., Wheeler and Wyszewianski (1982). Quality, Cost and Health: An Integrative Model. Medical Care. Vol. 20: 975-92.

Dor A. (1994). Non minimum cost functions and Stochastic Frontier: On Application to Health Care Provider. Journal of Health Economics. Vol. 13: 329-34.

Dowling W.L (1995). Strategic Alliances as a Structure for Integrated Delivery Systems. In Kaluzny A., Zuckerman H.S. and Ricketts III T.C. (eds.), Partners for the Dance: Forming Strategic Alliance in Health Care. Chicago: Health Administration Press.

Dranove D. (1987). Rate Setting by Diagnostic Related Groups and Hospital Specialization. RAND Journal of Economics. Vol. 18(3): 417-27.

Dranove D., Shanley M. and White W. (1993). Price and Concentration in Hospital Market: the Switch from Patient Driven to payer Driven Competition. Journal of Law and Economics. 34: 179-204.

Duke University (2002) Health policy Cyberexchange: Hospital Conversion. Durham, NC. URL: <http//www.hpolicy.duke.edu/cyberexchange/conversion/ruecon.html> (current at May 2002)

Dynan L., Bazzoli G.J. and Burns L. (1998). Assessing the extent of integration achieved through physician-hospital arrangements. Journal of healthcare Management. May/Jun 1998; 43(3): 242-62.

Eaton c. and Lipsey R. (1989). Product Differentiation. In handbook of Industrial Organization Vol. 1. North Holland, Amsterdam: 723-70.

Enthovan A. (1978). Consumer Choice Health Plan. New England Journal of Medicine. Vol. 298(12): 650-8.

Eswaran M. and Kotwal A. (1984) On the Moral Hazard of Budget-Breaking. Rand Journal of Economics, Winter, 1984: 578-581.

Evans R.G. (1971) behavioral cost function for Hospitals. Canadian Journal of Economics. Vol. 4(3), May: 198-215.

Evans, R.G. and Walker H.D. (1972). Information theory and the analysis of hospital cost structure. Canadian Journal of Economics. Vol. 5: 398-418.

Fama E.F. (1980) Agency Problem and the Theory of the Firm. Journal of Political Economy. Vol. 88(2): 288-307.

Fama E.F. and Jensen M. (1983). Agency Problems and Residue Claims. Journal of Law and Economics, Vol. 26: 327-49.

Farrar D.E. and Glauber R.R. (1967). Multicollinearity in Regression Analysis: The Problem Revisited. Review of Economics and Statistics. Vol. 49: 92-107.

Farrell M.J. (1957). The Measurement of Productive Efficiency. Journal of the Royal Statistical Society, Series A. Vol. 120:253-81.

Feldstein M.S. (1967). Economic Analysis for Health Service Efficiency. North Holland, Amsterdam.

Ferguson C.E (1969). The Neoclassical Theory of Production and Distribution. Cambridge University Press: Cambridge.

Fetter R.B., Shin Y., Freeman J.L., Averill R.F. and Thompson J.D. (1980). Casemix Definition by Diagnostic Related Groups. Medical Care. Vol. 18(2), February Supplement: 1-15.

Fleming S.T., McMahon L.F. Jr., Desharnais S.I., Chesney J.D., Wroblewski R.T. (1991). The Measurement of Mortality: A Risk-adjusted Variable Time Window Approach. Medical Care. Vol. 29(9) September: 815-28.

Frech H.E. III (1976). The Property Rights Theory of the Firm: Empirical Evidence from Natural Experiment. Journal of Political Economy. Vol. 84: 143-52.

Frick A.P., Martin S.G. and Swartz M. 9185). Casemix and cost Differences between Teaching and Non-teaching Hospitals. Medical Care. Vol. 23(4), April: 283-95.

Frisch R.A (1965). Theory of Production. Dordrecht: D. Reidel

Fuchs V.R. (1974). Who Shall Live? Basic Books, New York.

Fuss M. and McFadden D. (1978). Production Economics: A Dual Approach to Theory and Applications, Volume 1 and 2. North Holland, Amsterdam: 219-68.

Fuss M. and McFadden D. and Mundlak Y. (1978). A survey of Functional Forms in the Economic Analysis of Production. In Fuss M. and McFadden D. (eds.), Production Economics: A Dual Approach to Theory and Applications, Volume 1. North Holland, Amsterdam: 219-68.

Garber A.M. Fuchs V.R. and Silverman J.F. (1984). Casemix, Costs, and Outcomes: Differences between Faculty and Community Services in a University Hospital. New England Journal of Medicine. Vol. 310: 1231-7.

Garbrielsen A. (1975). On estimating Efficient Production Functions. Working Paper A-85 Chr. Michelsen Institute. Department of Humanities and social Sciences, Bergen: Norway.

Gaumer G. (1986). Medicare Patient Outcome and Hospital Organizational Mission. In Gray B.H. (ed), For Profit Enterprise in Healthcare. Ebook at URL: <http://www.nap.edu/books/0309036437/html>, National Academy Press: Washington DC.

Gaynor M. and Anderson G.F. (1995). Uncertain Demand, the Structure of Hospitals Costs, and the Cost of Empty Beds. Journal of Health Economics. Vol. 14: 291-317.

Glied S. (2000). Managed Care. In culyer A.J. and Newhouse J.P. (eds.), Handbook of Health Economics Vol. 1A. Elsevier Science B.V., Amsterdam, Netherlands: 707-53.

Granneman T.W., Brown R.S. and Pauly M.V. (1986). Estimating Hospitals Costs: A Multi-output Analysis. Journal of Health Economics. Vol. 5: 107-27.

Greene W. (1980). On estimation of a Flexible Frontier Production Model. Journal of Econometrics. Vol. 13: 27-56.

Greene W. (1997). Frontier Production Function. In Persaran M.H. and Schmidt P. (eds.), Handbook of Applied Econometrics Volume 2: Microeconomics. Blackwell Publishers: UK.

Grimaldi P.L. and Micheletti J.A. (1982). Homogeneity Revisited: the new DRGs. Journal of the American Medical Record Association. Vol. 53(4), April: 56-70.

Grossman M. (1972). On the concept of Health Capital and the Demand for Health. Journal of Political Economy. Vol. 80(2): 223-55.

Grossman S. and Hart O. (1986). The cost and Benefits of Ownership: A Theory of Vertical Integration. Journal of Political Economy, Vol. 94:691-719.

Hadley J. (1983). Teaching and Hospitals Costs. Journal of Health Economics. Vol. 2(1), March: 75-9.

Hall R.E. (1973). The Specification of Technology with Several Kinds of Output. Journal of Political Economy. Vol. 81(4) July/August: 878-92.

Hanoch G. (1970). Homotheticity in Joint Production. Journal of Economic Theory. Vol. 2(4) December: 423-6.

Hardwick J. (1986). Hospitals Casemix Standardization: A Comparison of Resource Need Index and Information Theory Measures. In Butler J.R.G. and Doessel D.P. (eds.), Economic and Health 1985: Proceedings of the Seventh Australian Conference of Health Economists. Australian Studies in Health Service Administration No. 56, School of Health Administration, University of New South Wales, Sydney: 36-63.

Harris J.E. (1977). The Internal Organization of Hospitals: Some Economic Implications. Bell Journal of Economics. Vol. 8: 467-82.

Harvey A.C. (1977). Some Comments on Multicollinearity in Regression. Applied Statistics. Vol. 26(2):188-191

Hausman J.A. (1978). Specification Tests in Econometrics. Econometrica Vol. 46: 1251-71.

Hausman J.A. and Taylor W. (1981) Panel Data and Unobservable Individual Effects. Econometrica, Vol. 49: 1377-1398.

Herzlinger R.E. and Krasker W. (1987). Who Profits from Non-profit? Harvard Business Review: 93-106.

Hicks J.R. (1932). The Theory of Wages, 1963 edition. London: Macmillan.

Hicks J.R. (1939). Value and Capital: an inquiry into some Fundamental Principles of Economic Theory, 1946 edition. Oxford: Clarendon Press.

Holmstrom B. (1982). Moral Hazard in Teams. Bell Journal of Economics, Vol. 13: 324-40.

Holmstrom B. and Milgrom P. (1991). Multitask principal Agent Analysis: Incentive Contracts, Asset Ownership and Job Design. Journal of Law, Economics and Organization. Vol. 7 (Special Issue); 24-52.

Horn S.D. and Schumacher D.N. (1979). An Analysis of Casemix Complexity using Information Theory and Diagnostic Related Grouping. Medical Care. Vol. 17(4), April: 382-9.

Horn S.D. and Sharkey P.D. (1983). Measuring Severity of Illness to Predict Patient Resource Use with DRGs. Inquiry. Vol. 26(4), Winter: 314-21.

Hornbrook M.C. (1982a). Hospitals Casemix: Its Definition, Measurement and Use: Part I. The Conceptual Framework. Medical Care Review. Vol. 39(1), Spring: 1-43.

Hornbrook M.C. (1982b). Hospitals Casemix: Its Definition, Measurement and Use: Part II. Review of Alternative Measures. Medical Care Review. Vol. 39(2) Summer: 73-123.

Hosek J.R. and Palmer A.R. (1983). Teaching and Hospitals Costs. Journal of Health Economics. Vol. 2(1) March: 29-46.

Hotelling H. (1932). Edgeworth Taxation Paradox and the Nature of Demand and Supply Functions. Journal of Political Economy. Vol. 40: 577-616.

Jenks S.F., Williams D.K. and Kay T.L. (1988). Assessing Hospitals Associated Death From Discharge Data: The Role of Length of Stay and Co-morbidities. Journal of the American Medical Association. Vol. 260: 2240-6.

Jensen M.C. and Meckling W.H. (1976). The Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. Journal of Financial Economics. Vol. 3: 305-360.

Jondrow J.K., Lovell I., Materov and Schmidt P. (1982). On the Estimation of Technical Efficiency in the Stochastic Frontier Production Function Model. Journal of Econometrics. Vol. 19: 233-8.

Jones K.R. (1985). Predicting Hospitals Charge and Stay Variation. Medical Care. Vol. 23(3), March: 220-35.

Klastorin T.D. and Watss C.A (1980). On the Measurement of Hospitals Casemix. Medical Care. Vol. 18(6), June: 675-85.

Kmenta J. and Gilbert R.F. (1968). Small Sample Properties of Alternative Estimators of SURs. Journal of American Statistical Association, vol. 63: 1180-200.

Kolsen, H.M. (1968). The Economics and Control of Road-Rail Competition. Sydney University Press, Sydney.

Koopmans T.C. (1951). An Analysis of Production as an Efficient Combination of Activities. In Koopmans T.C (ed.), Activity Analysis Production and Allocation: Cowles Commission for Research in Economics, Monograph No. 13, New York, John Wiley, 1951.

Kooreman P. (1994) Nursing home care in The Netherlands: a nonparametric efficiency analysis. Journal of Health Economics. Vol. 13 (3): 301-16.

Kumbhakar S. (1990). Production Frontier and Panel Data, and Time Varying Technical Inefficiency. Journal of Econometrics. Vol. 46: 201-11.

Kumbhakar S.C. (1991). The Measurement and Decomposition of Cost Efficiency: the Translog Cost System. Oxford Economics Papers, Vol. 43: 667-83.

Kumbhakar S.C. and Lovell C.A.K. (2000). Stochastic Frontier Analysis. Cambridge University Press.

Laitinen K. (1980). A Theory of Multi-product Firm. North Holland, Amsterdam.

Lee Y.H. and Schmidt P. (1993). A Production Frontier Model with Flexible Temporal Variation in Technical Efficiency. In fried H.K., Lovell C.A.K. and Schmidt P. (eds.), The Measurement of Productive Efficiency. Oxford University Press: New York.

Leontief W.W. (1936). Composite Commodities and the Problem of Index Numbers. Econometrica. Vol. 4: 39-59.

Lindsay C.M. (1976). A Theory of Government Enterprise. Journal of Political Economy. Vol. 84(5) October: 1061-78.

Linna M., Häkkinen U. and Linnako E. (1998). An Econometric Study of Costs of Teaching and Research in Finnish Hospitals. Health Economics. Vol. 7: 291-305.

Longest B.B. (1978). An Empirical Analysis of the Cost/Quality Relationship. Health and Health Service Administration. Vol. 23: 20-35.

Longley J.W. (1967) An Appraisal of Least Squares Programs for the Electronic Computer from the point of view of the User. Journal of the American Statistical Association, Vol. 62 (319): 819-41

Ma A. (1994). Healthcare Payment Systems: cost and Quality Incentives. Journal of Economics and Management Strategy. Vol. 3(1): 93-112.

Mason J.M. (1992). Econometric Software: A User's View. Journal of Economic Perspectives. Fall 1992: 165-187.

Meeusen W.J. and van den Broeck (1977). Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error. International Economic Review, Vol. 2: 435-444.

Melfi C.A. (1984). Estimation Decomposition of Productive Efficiency in a Panel Data Model: An Application to Electric Utilities. Unpublished Ph.D. Dissertation, University of North Carolina, Chapel Hill, NC.

Mundlak Y. (1963). Specification and Estimation of Multi-product Production Functions. Journal of Farm Economics. Vol. 45(2), May: 433-43.

Nadiri M.I. (1982). Producers Theory. In K.J. Arrow and M.D. Intrilligator (eds.), Handbook of Mathematical Economics, Vol. II. Amsterdam: North-Holland.

National Center fro Health Statistics (2003). Health, United States, 2003. Hyattsville, Maryland: Public Health Service. 2003.

Nerlove M. (1967). Recent Empirical Studies of the CES and related Production Function. In Brown M. (ed.), The Theory and Empirical Analysis of Production. Columbia University Press, New York.

Neuhauser D. (1971). The Relationship between Administrative Activities and Hospitals Performance, Research Series No. 28. Chicago: Center for Health Administration Studies, University of Chicago.

Newhouse J.P. (1970). Towards a Theory of Nonprofit Institutions: An Economic Model of a Hospital. American Economic Review. Vol. 60: 64-74.

Newhouse. J.P. (1994). Frontier Estimation: How Useful a Tool for Health Economics? Journal of Health Economics. Vol. 13(3): 317-22.

Patel J., Needleman J. and Zechauser R. (1994). Changing Fortunes, Hospital Behaviors, and Ownership Forms. Faculty Research Working Paper No. R94-17. John F Kennedy School of Government, Harvard University.

Pauly M.V. (1978). Medical Staff Characteristics and Hospitals Costs. Journal of Human Resources. Vol. 13, Supplement: 77-111.

Pauly M.V. and Redisch M. (1973). The Nonprofit Hospitals as a Physician's Cooperative. American Economic Review. Vol. 63: 87-100.

Pitt M. and Lee L.F. (1981). The Measurement and Sources of Technical Inefficiency in Indonesia Weaving Industry. Journal of Development Economics, Vol. 9:43-64.

Pollack M.M., Ruttimann U.E. and Geston P.R. (1987). Accurate Prediction of the Outcome of Pediatric Intensive Care: A New Quantitative Method. New England Journal of Medicine. Vol. 316: 134-9.

Register C.A. and Burning E.R. (1987). Profit Incentive and Technical Efficiency in the Production of Hospitals Care. Southern Economic Journal. Vol. 53: 899-914.

Richmond J. (1974). Estimating the Efficiency of Production. International Economic Review. Vol. 15: 515-21.

Robinson J.C., Garnick J. and McPhee S. (1987). Market and Regulatory Influences on the Availability of Coronary angioplasty and Bypass Surgery in the US Hospitals. New England Journal of Medicine. 317: 85-90.

Robinson J.C. and Luft H.S. (1985). The Impact of Hospitals Market Structure on Patient Volume, Average Length of stay, and the Cost of Care. Journal of Health Economics. Vol. 4: 333-56

Ruud P. (2000). An Introduction to Classical Econometric Theory. Oxford University Press: New York.

Samuelson P.A. (1947). Foundations of Economic Analysis, Cambridge MA: Harvard University Press.

Samuelson P.A (1966). The Fundamental Singularity Theorem for Non-joint Production. International Economic Review. Vol. 7(1), January: 34-41.

Sato K. (1967). A Two-Level Constant Elasticity of Substitution Production Function. Review of Economic Studies. Vol. 34: 201-18.

Schlesinger M., Marmor T.R. And Sithey R. (1987). Non-profit and For-profit Medical Care: Shifting Roles and Implications for Health Policy. Journal of Health Politics, Policy and Law. Vol. 12(3) Fall: 427-57.

Schmidt P. (1984). An Error Structure for Systems of Translog Cost and Share Equations, Econometrics Workshop Paper 8309, Department of Economics, Michigan State University, East Lansing, MI.

Schmidt P. (1986). Frontier Production Function. Econometric Review. Vol. 4(2): 289-328.

Schmidt P. and Lovell C.A.K. (1979). Estimating Technical and Allocative Inefficiency Relative to Stochastic Production and Cost Frontiers. Journal of Econometrics Vol. 9: 343-66.

Schmidt P. and Sickles R.C. (1984). Production Frontiers and Panel Data. Journal of Business and Economic Statistics. Vol. 2: 367-74.

Scott W.R., Forrest W.H. and Brown B.W. (1976). Hospitals Structure and Postoperative Mortality and Morbidity. In Shortell S.M. and Brown M. (eds.), Organizational Research in Hospitals. Chicago: Inquiry Books, Blue Cross Association.

Shepard R.W. (1953). Theory of Cost and Production Function. Princeton NJ; Princeton University Press.

Shleifer A. (1985). A Theory of Yardstick Competition. RAND Journal of Economics. 16: 319-27.

Shortell S.M., Becker S.W. and Neuhauser D. (1976). The Effects of Management Practice on Hospitals Efficiency and Quality of Care. In Shortell S.M. and Brown M. (eds.), Organizational Research in Hospitals. Chicago: Inquiry Books, Blue Cross Association.

Shortell S.M., Gillies R. and Andersen D. (1996). The New American Healthcare: Creating Organized Delivery Systems. San Francisco: Jossey-Bass Publishers.

Shortell S.M. and Hughes E.F.X. (1988). The Effects of Regulation, Competition and Ownership on mortality Rates among Hospital Inpatients. New England Journal of Medicine. Vol. 318: 1100-7.

Shortell, S.M. and Hull, K.E. (1996). The New Organization of the Health Care Delivery System: In S.H. Altman and U.E. Reinhardt (eds.) Strategic Choices for a Changing Health Care System. Chicago: Health Administration Press.

Sloan F.A. (2000). Not-for-profit Ownership and Hospital Behavior. In Culyer A.J. and Newhouse J.P. (eds.), Handbook of Health Economics Vol. 1B. Elsevier Science B.V., Amsterdam, Netherlands: 1141-74.

Sloan F.A., Feldman R.D. and Steinwald A.B. (1983). Effects of Teaching on Hospitals Costs. Journal of Health Economics. Vol. 2(1), March: 1-28.

Sloan F.A. and Steinwald A.B. (1980). Insurance Regulation and Hospitals Costs. D.C. Health and Co., Massachusetts.

Sloan F.A., Taylor D.H., Picone G. and Chou S.Y. (1998). Hospital Ownership, Cost and Quality of Care: Is there a Dime's Worth of Difference? NBER Working Paper #W6706. National Bureau of Economic Research.

Sloan F.A. and Vraiciu R.A. (1983). Investor Owned and Not-for-profit Hospitals: Addressing Some issues. Health Affairs. Vol. 2: 25-37.

Solberg E.J. (1982). Intermediate Microeconomics. Business Publication Inc., Texas.

Stevenson R. (1980). Likelihood Functions for Generalized Stochastic Frontier Estimation. Journal of Econometrics. Vol. 13: 58-66.

Theil H. (1967). Economics and Information Theory. North Holland: Amsterdam.

Thomson L. (1975). The Lives of a Cell. New York: Bantam Books.

Trochim, W. (2000). The Research Methods Knowledge Base, 2nd edition. Atomic Dog Publishing, Cincinnati, OH.

Uzawa H. (1962). Production Function with Constant Elasticities of Substitution. Review of Economic Studies. Vol. 29(4): 291-9.

Valdmanis V.G. (1990). Ownership and Technical Efficiency of Hospitals. Medical Care. Vol. 28: 552-61.

Varney, C.A. (1995). Efficiency Justifications in Hospital Mergers and Vertical Integration Concerns. Prepared Remarks of Commissioner Christine A. Varney (Federal Trade Commission), at Healthcare Antitrust Forum. Holiday Inn, Chicago, Illinois (2 May 1995).

Wagstaff A. (1989). Estimating Efficiency in Hospitals Sector: A Comparison of Three Statistical Cost Frontier Methods. Applied Economics. Vol. 21: 659-72.

Wang B.B., Wan T.T., Clement J. and Begun J. (2001). Managed Care, Vertical Integration Strategies and Hospital Performance. Health Care Management Science. Vol. 4(3) September: 181-91.

Watts C.A. and Klastorin T.D. (1980). The Impact of Casemix on Hospitals cost: A Comparative Analysis. Inquiry. Vol. 23(4), Winter 357-67.

Weinstein M. (1964). The Sum of Values from a Normal and A truncated Normal Distribution. Technometrics. Vol. 6: 104-5, 469-70.

Weisbrod B.A. (1991). The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care and Cost Containment. Journal of Economic Literature, 29: 523-552.

Wilson G.W. and Jadlow M.J. (1982). Competition, Profit Incentives, and Technical Efficiency in the Provision of Nuclear Medicine Services. Bell Journal of Economics. Vol. 13: 472-82.

Wilson, J. Q. (1989). Bureaucracy. New York: Basic Books.

Winstein C. (1957). Discussion of Mr. Farrell's Paper. Journal of Royal Statistical Society Series A, General. Vol. 120: 282-4.

Zuckerman S., Hadley J. and Iezzoni L. (1994). Measuring Hospitals Efficiency with Frontier Cost Functions. Journal of Health Economics. Vol. 13: 255-80.