2009

HOW EFFICIENT ARE MILITARY HOSPITALS? A COMPARISON OF TECHNICAL EFFICIENCY USING STOCHASTIC FRONTIER ANALYSIS

Linda Gail Kimsey

University of Kentucky, lgkimsey@gmail.com

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Recommended Citation


https://uknowledge.uky.edu/gradschool_diss/743

This Dissertation is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Doctoral Dissertations by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsvaky.edu.
ABSTRACT OF DISSERTATION

Linda Gail Kimsey

The Graduate School

University of Kentucky

2009
HOW EFFICIENT ARE MILITARY HOSPITALS? A COMPARISON OF TECHNICAL EFFICIENCY USING STOCHASTIC FRONTIER ANALYSIS

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Martin School of Public Policy and Administration at the University of Kentucky

By
Linda Gail Kimsey

Lexington, Kentucky

Co-Directors: Dr. J. S. Butler, Professor of Public Policy and Administration
and Dr. E.F. Toma, Professor of Public Policy and Administration

Lexington, Kentucky

2009

Copyright © Linda Gail Kimsey 2009
ABSTRACT OF DISSERTATION

HOW EFFICIENT ARE MILITARY HOSPITALS? A COMPARISON OF TECHNICAL EFFICIENCY USING STOCHASTIC FRONTIER ANALYSIS

Attainment of greater efficiency in hospital operations has become a goal highly sought after as a result of several factors including skyrocketing costs. The possibility that the different incentives associated with ownership type might affect efficiency has been covered thoroughly in the literature. There are numerous studies comparing for-profit to not-for-profit hospitals or public to private hospitals. Analysis of federal ownership, however, has been less studied. In particular, comparisons involving military hospitals are non-existent, attributed to data availability and an assumption that military hospitals are too different from civilian facilities.

This dissertation employs a cross-sectional Stochastic Frontier Analysis (“SFA”) of 2006 data to compare the technical efficiency of military, for-profit, not-for-profit, and other government hospitals, controlling for differences in patients, scope of work, physician-hospital working arrangements, and other structural characteristics. Four model specifications are examined, varying the method of accounting for heterogeneity of case mix. One of the specifications uses a distance function technique to allow for specific inclusion of multiple outputs, namely inpatient and outpatient workload. Results obtained using SFA are validated using Data Envelopment Analysis (“DEA”) and compared with results produced through simple ratio analysis.

Estimates of overall technical efficiency ranged from 76% to 80%. The analysis found no significant correlation between ownership category and technical efficiency. Factors found to be significantly correlated with greater technical efficiency include younger average patient age, more female patients, percentage of surgical inpatient work, percentage of circulatory system-based work, accreditation, and having all credentialed
physicians (i.e. no physician employees). Pooled-vs.-partitioned analysis showed that military hospitals are indeed different, but not enough to render comparisons meaningless. Data Envelopment Analysis produced comparable individual hospital efficiency scores (correlations of approximately 0.6 between like specifications using SFA and DEA) and comparable average efficiency (~87%). Ratio analysis results were sensitive to the specific ratio analyzed.

This dissertation adds to the body of literature on the relationship between ownership and hospital technical efficiency. It is the first comparison of military and civilian hospital technical efficiency.

KEYWORDS: Hospital Efficiency; Stochastic Frontier Analysis; Technical Efficiency; Military Hospitals; Ownership

Linda Gail Kimsey
Student Signature

May 12, 2009
Date
HOW EFFICIENT ARE MILITARY HOSPITALS? A COMPARISON OF TECHNICAL EFFICIENCY USING STOCHASTIC FRONTIER ANALYSIS

By

Linda Gail Kimsey

Dr. J. S. Butler, Ph.D.
Co-Director of Dissertation and Director of Graduate Studies

Dr. E. F. Toma, Ph.D.
Co-Director of Dissertation

May 12, 2009
RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor’s degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but with quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

Name

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

Date
DISSERTATION

Linda Gail Kimsey

The Graduate School

University of Kentucky

2009
HOW EFFICIENT ARE MILITARY HOSPITALS? A COMPARISON OF TECHNICAL EFFICIENCY USING STOCHASTIC FRONTIER ANALYSIS

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Martin School of Public Policy and Administration at the University of Kentucky

By

Linda Gail Kimsey

Lexington, Kentucky

Co-Directors: Dr. J. S. Butler, Professor of Public Policy and Administration and Dr. E.F. Toma, Professor of Public Policy and Administration

Lexington, Kentucky

2009

Copyright © Linda Gail Kimsey 2009
To my mother.
ACKNOWLEDGMENTS

I gratefully acknowledge the U.S. Navy for affording me the opportunity to earn my Ph.D. I also gratefully acknowledge my dissertation co-chairs, Dr. J.S. Butler and Dr. Eugenia Toma for their support in finishing my degree within strict time limits.

Additionally, Dr. Butler provided expert guidance on the estimation methods as well as ceaseless enthusiasm for my chosen topic, while Dr. Toma provided expert assistance on the literature foundations for my work. Next, I thank the other members of my dissertation committee, Dr. Jeff Talbert, Dr. Dwight Denison, and Dr. Aaron Yelowitz for their insights and recommendations. DeAnn Farr willingly donated her time to pull the military data I needed to perform my analysis. Finally, my sister Laura and two friends, Sandi and Yelena, provided both emotional support and proofreading assistance that were invaluable to me in completing this degree.
TABLE OF CONTENTS

ACKNOWLEDGMENTS...........................................................................................................iii
LIST OF TABLES...................................................................................................................viii
LIST OF FIGURES....................................................................................................................ix
CHAPTER 1 - INTRODUCTION..............................................................................................1
  Background..........................................................................................................................1
  Focus.................................................................................................................................4
  Goals...............................................................................................................................5
CHAPTER 2 - LITERATURE REVIEW....................................................................................7
  Ownership..........................................................................................................................7
    Property Rights...............................................................................................................8
    Monitoring......................................................................................................................9
    The Role of Not-for-Profits..........................................................................................9
    Government Ownership...............................................................................................11
  Factors Attenuating the Effects of Ownership.................................................................12
    Competition..................................................................................................................12
    Size...............................................................................................................................15
  Other Hospital-Specific Factors Influencing Efficiency .....................................................16
    Quality..........................................................................................................................16
    Physician Characteristics.............................................................................................17
    Patient Characteristics.................................................................................................19
  Progression of Hospital Ownership/Efficiency Studies......................................................20
    Pre-Frontier Analysis.....................................................................................................20
    Pre-Frontier Methods.....................................................................................................20
    Pre-Frontier Studies.......................................................................................................21
  Frontier-based research....................................................................................................23
    Frontier Fundamentals.................................................................................................23
    Frontier Methods of Measuring Efficiency.....................................................................27
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Envelopment Analysis</td>
<td>125</td>
</tr>
<tr>
<td>Variables</td>
<td>126</td>
</tr>
<tr>
<td>Methods</td>
<td>126</td>
</tr>
<tr>
<td>Results</td>
<td>127</td>
</tr>
<tr>
<td>Correlation with Traditional Ratio Measures</td>
<td>128</td>
</tr>
<tr>
<td>Summary</td>
<td>133</td>
</tr>
<tr>
<td>CHAPTER 10 - SUMMARY AND CONCLUSIONS</td>
<td>135</td>
</tr>
<tr>
<td>Summary</td>
<td>135</td>
</tr>
<tr>
<td>Limitations</td>
<td>137</td>
</tr>
<tr>
<td>Recommendations for Future Research</td>
<td>140</td>
</tr>
<tr>
<td>Conclusions</td>
<td>142</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>143</td>
</tr>
<tr>
<td>VITA</td>
<td>153</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 6.1: Final Data Set Variables ................................................................. 78
Table 6.2: General U.S. Short-Term Medical/Surgical Hospital Raw Data ............. 80
Table 6.3: Final Sample Descriptive Statistics .................................................. 81
Table 6.4: Sample Mean Descriptive Statistics By Ownership ................................ 84
Table 6.5: Top Ten Military Inpatient DRGs ..................................................... 88
Table 6.6: Top Ten Civilian DRGs ................................................................. 88
Table 6.7: Complexity of Work ..................................................................... 89
Table 6.8: Inpatient Work Profile .................................................................. 90
Table 7.2: Hypothesis Testing Summary ............................................................ 101
Table 8.1: Production Function ..................................................................... 106
Table 8.2: Inefficiency Results ...................................................................... 109
Table 8.3: Tests of Significance for Differences By Ownership Category .............. 110
Table 8.4: Average Technical Inefficiency by Ownership Type ............................. 115
Table 8.5: Average Technical Inefficiency by Ownership Type and Size ............... 116
Table 8.6: Average Technical Inefficiency by Military Service ............................ 117
Table 8.7: Average Technical Inefficiency by Ownership and Median Case Mix Index.. 118
Table 8.8: Average Technical Inefficiency by Volume of MDC 14 Workload .......... 118
Table 8.9: Partitioned Results ....................................................................... 120
Table 9.1: Ranking Correlation Matrix of SFA and DEA Specifications .................. 128
Table 9.2: Ownership Coefficients Based on Regression Using Ratios .................... 131
Table 9.3: Significance Tests for Differences By Ownership Category – Ratio Analysis 132
Table 9.4: Correlation of SFA, DEA, and Ratio Rankings – CMI and Raw .............. 133
LIST OF FIGURES

Figure 2.1: Production Possibilities Frontier .............................................................. 24
Figure 2.2: Graphical Depiction of Technical and Allocative Efficiencies .................. 26
Figure 4.1: Conceptual Model of Hospital Technical Efficiency ................................. 50
Figure 7.1: Error Term Distributions ........................................................................ 99
Figure 8.10: Scatter Plot of Residuals .................................................................. 122
CHAPTER 1 - INTRODUCTION

Background
In 2006, healthcare expenditures in the U.S. comprised 16% of GDP, reaching $2.1 trillion (Centers for Medicare and Medicaid Services, 2008). Even more alarmingly, one economist has projected this share of GDP to increase to 29% by 2040 (Fogel, 2008). The causes cited for this growth are multifaceted. An aging population is believed to consume greater amounts of healthcare. Governmental regulations aimed at ensuring minimum levels of quality and access raise costs. Diagnostic and therapeutic technological advances increase both cost and quantity demanded (Feldstein, 2007). Furthermore, if physicians perform these diagnostic and therapeutic measures not for the patient’s health, but to safeguard against malpractice, they engage in defensive medicine, which can also drive growth in healthcare expenditures (U.S. Congress, Office of Technology Assessment, 1994). Health insurance encourages moral hazard, increasing demand for health services because they are valued at less than their actual cost by those insured (Feldstein, 2007). Additionally, interaction of these phenomena can result in even further expenditure increases. Technological improvements increase both costs and demand for insurance. Simultaneously, expanded insurance encourages research and development of more advanced (and costly) technology (Weisbrod, 1991), and these processes feed each other in an upward spiral of costs. While causes for these cost increases have been identified, solutions have not. Furthermore, if (as some have asserted) much of the U.S. healthcare system now operates in the “flat of the curve”, where marginal increases in medical expenditures produce increasingly smaller marginal improvements (or even decrements) in overall health (Feldstein, 2007), the effectiveness of these expected increases in expenditures is uncertain.
Given the growing share of GDP that healthcare occupies, it seems natural for efficiency (commonly referred to as “doing more with less”) to be an important pursuit for healthcare organizations. Quantifying and improving efficiency has become and accepted strategy for controlling an organization’s overall expenditures. What exactly is efficiency, and why is it important? Broadly speaking, efficiency can be defined by an index of ratios of output quantities to input quantities. The firm that produces more output with the same quantity of inputs (i.e. is closer to possible production levels) is more efficient. These same concepts apply to costs as well as quantities. The firm that produces more output at a lower cost (i.e. chooses the less costly allocation of inputs) is more efficient (Salerno, 2003). Efficiency of organizations is always vital from an economic perspective because of the basic principle of scarcity of resources. Given the environment of increasing costs just discussed, this holds true especially in the healthcare sector.

Expenditures on hospital care are the largest category of healthcare costs, accounting for 31% of total national healthcare expenditures (Centers for Medicare and Medicaid Services, 2008). Since hospitals represent the largest segment of healthcare costs and they provide the highest cost healthcare services, they provide a potentially fruitful venue for investigating both the quantity and possible causes of inefficiency.

It may be that certain specific organizational characteristics correlate with higher levels of efficiency and arguably, one of the most fundamental organizational characteristics is ownership type. Is it possible that systematic structural differences between for-profit, not-for-profit, and government hospitals create differences in levels of efficiency achieved? If so, this knowledge would be useful in implementing policies that encourage success of the most efficient ownership type.
Military hospitals are an infrequently studied ownership type – not state, county, or local-level government – but federal government ownership. While the majority of federal spending on health is for Medicare and Medicaid insurance programs, $130 billion was spent by the federal government in 2005 for other healthcare (Feldstein, 2007), including direct health care provision for veterans, military members and their families, Indians on reservations, Alaskan natives, and prisoners. This care is provided within the Department of Defense (DOD), Veteran Affairs (VA), Public Health Services (PHS), and the Department of Justice (DOJ) (Harrison, Coppola, & Wakefield, 2004). In 2007, the combined VA and DOD health program budgets were an estimated $70 billion. Additionally, the outlay for Indian Health Services’ budget was $3.26 billion resulting in a total federal budget of approximately $73.3 billion in 2007, exclusive of Bureau of Justice healthcare and other miscellaneous categories not specifically identified (Office of Management and Budget, 2007). If ownership affects efficiency, what effect does federal control have? Could federally-controlled military hospitals be inherently more efficient (or inefficient) than hospitals controlled by other types of ownership, ceteris paribus? This dissertation, using Stochastic Frontier Analysis (“SFA”) as the primary method of estimation to examine the most fundamental type of efficiency – technical efficiency, will investigate this question.

Studies of hospital efficiency abound, and many of these have considered – to some degree – the effects of ownership on efficiency. Military hospitals have yet to significantly inform this discussion, perhaps because they are often viewed as unusual and thus less relevant for comparisons. Yet military hospitals care for a highly-valued segment of citizens – those who volunteer to defend our country – and all citizens fund this care with taxes. If military hospitals provide equivalent care in a more efficient manner, perhaps other healthcare providers can learn from military practices and procedures identified as efficiency-promoting. Alternatively, if military hospitals are not capable of providing high quality health care in an efficient manner, perhaps policymakers should consider pursuing other alternatives for providing healthcare to
military beneficiaries. Additionally, military hospitals – like Veterans Health Administration (“VA”) hospitals – provide an opportunity to study a functioning system of socialized medicine in the U.S.

Focus

Technological (or productive) efficiency is the most fundamental efficiency measure: it refers strictly to the maximum output possible given a set of inputs or to minimum inputs possible given a set output level. It does not imply cost minimization or benefit maximization. There are other ways to measure efficiency. Allocative efficiency is a broader concept that factors in the cost of resources. It refers to “the extent to which inefficiency occurs because an institution is using the ‘wrong’ combination of inputs given what they cost to purchase”. Cost efficiency jointly considers technical and allocative efficiency, and is calculated by multiplying the two together (Salerno, 2003).

These concepts, pertinent to iso-cost and iso-quant curves, are applicable to revenues and outputs and their respective iso-revenue and production possibilities curves as well (Coelli, Prasada Rao, O'Donnell, & Battese, 2005). Scale efficiency is a component of technical efficiency. Constant returns to scale – the long-run outcome among competitive firms – signify perfect scale efficiency. If a firm is operating at either increasing or decreasing returns to scale, it is not scale efficient (Salerno, 2003).

Some argue that technical efficiency is the most appropriate basis for comparison when public organizations are considered. Public organizations have goals other than efficiency, namely, equity, financial balance, and macroeconomics. Equity refers to effects on income redistribution; financial balance refers to deficits incurred and the reasons for them; macroeconomics refers to unemployment rates, trade deficits, etc. Which of these goals is most important depends on the type of organization. However, “productive efficiency is the only objective whose attainment does not impede the realization of other goals. Producing too little or employing too many factors as
compared to what is technically feasible cannot be justified in the name of any other objective” (Pestieau & Tulkens, 1990). Focusing on technical efficiency eliminates concerns over pricing and costing differences, and non-financial performance data is likely to be more reliable (Pestieau & Tulkens, 1990). It also allows for the possibility that public facilities might have different constraints that place allocative efficiency at a different point along the frontier of production possibilities. A technical efficiency focus is also appropriate because achieving it is thought to produce the highest gains. This is consistent with Leibenstein’s concept of X-efficiency, the idea that greater gains in efficiency were achievable from reaching the production frontier than from movement on the frontier (i.e. improvements in allocative efficiency) (Leibenstein, 1966). This dissertation will focus on comparing technical (productive) efficiency across four ownership types – for-profit, not-for-profit, state/local, and military (federal) ownership.

**Goals**

The intended results of this dissertation are twofold:

1) It should inform military leaders (medical and non-medical) on the efficiency of military hospitals in comparison to civilian facilities – not-for-profit, for-profit, and other governmental hospitals. Both significant and non-significant findings will be of interest. If military hospitals are found to be at least as efficient as civilian facilities, military leadership should consider focusing their efficiency-seeking efforts elsewhere, such as strategies to control increases in benefits provided or pure cost-reduction initiatives. At a minimum, the results could be used to counter the bad press” that the Military Health System sometimes receives. If they are found to be less efficient than civilian facilities, perhaps policymakers should pursue other alternatives for healthcare provision for military beneficiaries that utilize the more efficient civilian sector.

2) It should add to the current body of literature on the effects of ownership – including that of the federal government – on hospital productive efficiency. By either capitalizing on their “unusual” characteristics (rather than shunning them) as
explanatory variables or controlling for them where necessary, the inclusion of military hospitals in this analysis allows for exploration of the factors thought to influence efficiency from a new direction.

Chapter 2 is a comprehensive literature review that addresses ownership and other factors that may affect efficiency and their treatment in the literature, studies of ownership and efficiency in healthcare prior to development of frontier techniques, the introduction of frontier techniques to healthcare, and studies of ownership and efficiency using these frontier techniques. Chapter 3 describes the Military Health System, and discusses studies of efficiency involving military facilities. This chapter concludes with the research questions that this dissertation intends to answer. Chapter 4 develops a conceptual model of hospital healthcare production and efficiency, and Chapter 5 extends this conceptual model to derive an empirical model. Chapter 6 describes the data and manipulation of it to create the variables that comprise the actual data set. This chapter also presents descriptive statistics. Chapter 7 discusses methodology. Chapter 8 presents the results, and Chapter 9 validates them. Finally, Chapter 10 summarizes the findings of this dissertation, discusses limitations, and provides suggestions for further research.
CHAPTER 2 - LITERATURE REVIEW

The focus of this dissertation – comparing technical efficiency of military hospitals to that of other hospital types – requires the following:

- an understanding of the theories that have been developed as to why one form of ownership might produce higher (or lower) levels of efficiency than others,
- knowledge of what previous studies have found regarding ownership and hospital efficiency, and
- an understanding of the distinctive nature of the Military Health System.

This chapter first reviews the existing literature on theories of ownership, both in general and in the specific context of hospital-provided healthcare. The second portion of this chapter reviews other characteristics that have been theorized to affect hospital efficiency. The final portion discusses the progression of hospital efficiency studies, focusing on those that analyzed the effects of ownership, and the methods these studies used. Chapter 3 will discuss the unusual characteristics of the Military Health System.

Ownership

Theories about the effects of ownership on organizational operations – including efficiency – are well-established in the literature. Essentially, one can view this literature as a debate on whether structure (including ownership) or environment is more important in determining efficiency. The conclusions of this research have varied.

Tullock discussed the difficulties in identifying the role that ownership (control) might play in shaping organizational performance. Viewed as black boxes within black boxes, the operation of bureaucracies (governmental or not) is difficult to analyze because it requires identifying and measuring the specific constraints actually faced by the
managers of (and within) these nested black boxes (Tullock, 1977). However, two concepts – property rights and monitoring – appear in the literature as agreed-upon explanations for differences in performance (including efficiency) related to ownership.

Property Rights
The idea that ownership might affect efficiency (as well as other performance indicators) is grounded in property rights theory. Property rights create profit incentives (Alchian, 1965), and differences in these property rights produce different incentives. In a for-profit organization, residual profits – the difference between earnings and costs – flow to its shareholders. Because they have rights to the residual profits, owners should have greater incentive to attain higher efficiency in order to reap greater revenues and increase residual profits. Not-for-profit organizations are barred from making such distributions to those with control over it (Hannsmann, 1980). Government organizations also have no specific residual claimants. Thus, property rights are either attenuated or non-existent in both not-for-profits and public organizations. Said another way, for-profit organizations – interested in maximizing shareholder returns – focus on the bottom line, and this makes them more efficient than either not-for-profit or governmental organizations. Furthermore, property rights provide “incentives to invest and innovate” (Schleifer, 1998) that are not present in the public sector. Given these incentives, government ownership is only potentially preferable when:

- The possibility of cost reductions leading to lessened quality exists;
- Innovation is relatively unimportant;
- Weak competition and ineffective consumer choice are present; and
- There are weak reputational mechanisms.

However, Schleifer points out that both the presence of not-for-profits – a private alternative seen by some as developed by market economies to assuage concerns of decreased quality due to the bottom-line focus of for-profit organizations – and the possibility of corruption of government officials further reduce the instances where
government ownership would be preferable (Schleifer, 1998). Not-for-profits dominate the hospital industry, which therefore suggests a reduced preference for government ownership of hospitals. Not-for-profits are discussed further below.

Monitoring
Additionally, principal-agent theory shapes the effects of ownership on efficiency. The market, which monitors agents’ performance in private organizations, is assumed a better means of monitoring agents’ behavior for principals than the political processes at work in government organizations. When neither a profit motive nor an adequate monitoring mechanism exists, substandard performance can manifest itself in several ways. Niskanen theorized that bureaucrats chose to maximize discretionary budgets rather than profits, leading to inefficiency, overproduction, or some combination thereof (Niskanen, 1975). Migue and Belanger hypothesized that bureaucrats might pursue other goals – such as a preference for larger staffs or greater capital – that would reduce efficiency. Williamson and DeAlessi, as well as Parker, focused this differing-goal theory on a preference for personnel (Orzechowski, 1977). Finally, in developing the concept of “X-efficiency”, Leibenstein (1966) addressed the notion that organizations, especially government ones, may operate below the production possibilities frontier due to organizational schemes that might arise in an environment of ineffective monitoring.

The Role of Not-for-Profits
The dominating role of not-for-profits is a defining characteristic of the hospital industry, and (as noted above) this is a condition that theoretically reduces the instances where government ownership might be preferable to private ownership (Schleifer, 1998). There are three dominant theories as to why not-for-profits are prevalent in healthcare (Folland, Goodman, & Stano, 2006). Weisbrod hypothesized that not-for-profits arose to supply unmet healthcare demands due to the public goods nature of
healthcare. Markets characteristically undersupply goods of a public nature. Additionally, if the median voter theory holds true, government also undersupplies healthcare based on the preferences of some voters. Charitable donations to not-for-profit hospitals provide those individuals one way to act on their desire for more healthcare. In doing so, they receive additional personal benefits external to market processes. Hansmann focused on contract failure theory to explain their growth, asserting as Schleifer (1998) did that quality in healthcare is hard for individuals to discern and profit incentives further obfuscate the picture. Thus, not-for-profit hospitals fulfill a quality-monitoring role for the general population. Finally, Bays attributed the rise of not-for-profits to interest group theory. Physicians as an interest group were (and still are, although perhaps to a lesser degree) very powerful due to their homogeneous desires, the concentrated benefits for which they lobbied, and the diffuse costs of those benefits to the public. Physicians found this power gave them greater influence in not-for-profit hospitals (Folland, Goodman, & Stano, 2006).

Not-for-profits and for-profits differ in their institutional constraints in three key ways: 1) Not-for-profits cannot sell stock; they must rely on donative capital. 2) Not-for-profits cannot pay out residual profits as shareholder dividends. 3) Not-for-profit firms cannot be sold for proceeds that would flow to individual owners (Pauly, 1987). Following the property rights theory discussed above, these constraints should lead to less efficient operations on the part of not-for-profits. The Newhouse utility-maximization model of not-for-profit behavior is based on this theory. If not-for-profits are not profit-maximizers, they must seek to maximize non-monetary utility based on the preferences of individual managers (Folland, Goodman, & Stano, 2006). Similar to government organizations, leaders of not-for-profit organizations are thought to have goals other than profit maximization, such as altruism (Duggan, 2000), and since any monitoring mechanisms are not free-market based, pursuit of these other goals may lead to greater inefficiency. Frech (1976) had similar ideas, theorizing that not-for-
profits are more likely to prefer non-pecuniary benefits, and if these benefits are costly, then not-for-profits may be more inefficient than for-profits. Pauly and Redisch offered a different view, maintaining that profit-maximization is still at work in not-for-profit hospitals, but the physicians are the ones acting as profit maximizers. Harris modeled the hospital as a “non-cooperative oligopolistic game” between administrators and physicians (Folland, Goodman, & Stano, 2006). In all of these theories, goals other than maximizing residual profits likely lead to lessened efficiency.

However, not-for-profit organizations span a wide continuum of operations. Using Hansmann’s four-quadrant typology (donative vs. commercial financing, and mutual vs. entrepreneurial control), not-for-profit hospitals today fall into the entrepreneurial commercial category, generally most similar to for-profit organizations (Hannsmann, 1980). Additionally, hospitals of all ownership types operate in an environment characterized by extensive regulation, information asymmetry and non-price competition (Vining & Boardman, 1992). Thus, perhaps not-for-profit and for-profit hospitals do not function differently but rather react similarly to the environment in which they operate. In a summary of research on differences in for-profits and not-for-profits, Sloan found them to be “far more alike than different” in most measures of performance, including efficiency (Sloan, 2000). A fifteen-year longitudinal regression analysis found convergence in overall for-profit and not-for-profit efficiency as measured by expenses per adjusted admission and Full-Time-Equivalents (“FTEs”) per adjusted census from 1980 to 1994 due in large part to industry-wide regulatory changes (Potter, 2001).

Government Ownership
Theories on the incentives created by property rights and the effectiveness of monitoring from either participation in the market or through principal-agent relationships are used to explain why both government and not-for-profit control should
be inferior to for-profit ownership in terms of efficiency. However, theories as to whether not-for-profit ownership is either superior or inferior to government ownership are not discussed as explicitly in the literature, either in general or in a healthcare context. Furthermore, theories on differences based on the level of government (local, state, or federal) are even harder to find. However, it seems logical to assume that not-for-profit hospitals experience more exposure to market forces (via debt markets) than government organizations, and thus should be more efficient. With respect to differences in performance based on differences in government level, theories regarding size (see discussion below) and hierarchical distance from principal to agent (which would complicate monitoring) would seem to be most applicable. If these theories hold true, federal control should lead to lower levels of efficiency.

Factors Attenuating the Effects of Ownership

If they exist, differences in performance across ownership type due to property rights and monitoring mechanisms are innate and not easily changed. Incentives created by ownership could be costly to alter, but specific strategies involving more controllable organizational characteristics may be implemented that attenuate these differences, at least at the margin. Some of these characteristics include competition and size, and for hospitals, quality and physician and patient characteristics. These characteristics are discussed below.

Competition

One of the characteristics most discussed in the literature as potentially influential on organizational efficiency is competition. In an empirical analysis of 670 U.K. firms that found competition led to increased total factor productivity growth, Nickell (1996) noted that competition has been thought to improve efficiency since the days of Adam Smith. Monopolies provide the opportunity for slack on the part of owners, managers, and employees, and slack reduces efficiency and productivity. Competition lessens the
possibility of slack. Nickell noted three general “broad-brush” examples that evidence the effects of competition:

- A comparison of Eastern Europe’s low productivity with Western Europe’s much higher productivity, demonstrating what repression of competition can do,
- International success by Japan in industries with strong domestic competition compared with lackluster Japanese performance in industries with little or no domestic competition; and
- The considerable gains in productivity of U.S. airlines after de-regulation

Nickell also referred to several studies that found a correlation between greater market concentration and technical efficiency (Nickell, 1996).

Government agencies typically face less competition in their daily operations than for-profit organizations (Leibenstein, 1966), since fewer other organizations perform their tasks. Not-for-profits’ access to donative capital also lessens exposure to market competition. Depending on the industry, however, competition can be introduced into the operation of organizations that do not operate for profit. Competition was one of four elements Spann discussed that could improve efficiency of government entities: a governmental unit must produce its goods or services at a price or quality level equal to its private competitor to maintain customers and continue operating (Spann, 1977).

Niskanen also discussed competition. He theorized that insufficient exposure to competition was one of the reasons bureaucrats in government agencies were able to maximize budgets (rather than profits), and that greater competition for goods/services produced (along with increased contracting with the private sector) could reduce inefficiency. He also noted that competition increased exposure to free-market mechanisms, thereby lessening the need for monitoring of agent performance by the responsible level of government (Niskanen, 1975). Leibenstein – the developer of the idea of x-efficiency – also acknowledged the role of inadequate competition in permitting organizational schemes to exist that might lead to greater inefficiency (Leibenstein, 1966).
While some authors have argued that competition is a more important determinant of efficiency than ownership is, two related analyses of top 500 Canadian non-financial companies (Vining & Boardman, 1992) and top 500 non-U.S. industrial companies (Boardman & Vining, 1989) found evidence to the contrary. Superiority aside, it is plausible that competition attenuates to some degree the influence of ownership on hospital performance by forcing government and not-for-profit facilities to perform at the level of for-profits. Even federal hospitals located in large markets may be affected (albeit to a lesser degree) by the presence of other facilities: although non-eligible beneficiaries may not receive care in these federal facilities, those eligible for care in federal hospitals may choose to receive care at other hospitals. Exploring this notion, Burgess & Wilson utilized a county-level Herfindahl-Hirschman Index ("HHI") to model competition as an explanatory variable of inefficiency for all hospitals, including VA facilities (Burgess & Wilson, 1998). While this variable was not significant in their model of technical efficiency, updated definitions of relevant geographical markets for hospital competition exist that may produce different results (Wong, Zhan, & Mutter, 2005) (The Dartmouth Institute for Health Policy and Clinical Practice, 1999).

Increased price-based competition need not always lead to greater efficiency, as Carroll (1990) pointed out in a theoretical exposition on the behavior of federal agencies. Harris (1977) put forth a similar view of the impact of competition on efficiency of hospitals: the internal organization of hospitals and the role of the physician within its walls create non-price competition to attract physicians to the facility. A greater number of hospitals in a market results in increased competition for physicians. This competition reveals itself through increased purchases of equipment and supplies thought to attract providers, and this leads to greater inefficiency. Supporting this theory, Wilson & Jadlow (1982) found a significant direct correlation between
inefficiency and market concentration in their study of hospital nuclear medicine services.

Size
Spann (1977), Niskanen (1975), and Leibenstein (1966) also addressed the effects of organization size on performance. Size can be more significant than ownership when analyzing certain performance indicators. In an early hospital-specific study, Clarkson noted that fewer for-profit hospitals were accredited. However, he demonstrated that adding a variable for size rendered this association non-significant (Clarkson, 1972). Typical indicators of size are assets, sales, and employees (Boardman & Vining, 1989) (Vining & Boardman, 1992). Spann (1977) discussed how differences in the actual size of the political unit and the optimal scale of operations might negatively affect efficiency. More generally, however, larger size has been theorized to correlate with greater inefficiency. Niskanen (1975) theorized that larger principals and/or agents complicated the monitoring process, and this could lead to inefficiency. Hansmann (1980) and Tullock (1977) discussed the possibility that in larger organizations (whether for-profit, not-for-profit, or governmental), ownership and control tend to be further apart, and this separation results in greater inefficiency.

The competing viewpoint on the impact of size on performance (efficiency) is that economies of scale, if present, produce a correlation between larger size and less inefficiency. With respect to hospital-specific research, Vitaliano and Toren (1996) and Bruning and Register (1989) are examples of studies supportive of economies of scale. Bruning and Register’s 1989 cross-sectional study of 1,254 U.S. hospitals found a significant correlation between more beds and greater efficiency. Furthermore, they found the effects of size to be similar on for-profits and not-for-profits (Bruning & Register, 1989). Two studies focusing on VA hospitals provided empirical support for the opposing view (i.e. Niskanen’s theory on monitoring complications associated with
size), finding larger hospitals to be less efficient (Hao & Pegels, 1994) (Sexton & et.al., 1989).

Other Hospital-Specific Factors Influencing Efficiency

Quality
In most sectors where efficiency studies have been popular, such as banking and utilities, quality is less relevant. However, some recognition of quality in healthcare efficiency studies seems especially important in order to address overall effectiveness in the provision of potentially life-or-death hospital care. Additionally, the theory that not-for-profits fulfill a role of ensuring the provision of high-quality healthcare (Schleifer, 1998) (Hannsmann, 1980) indicates quality should be included in hospital efficiency studies. In the extreme, if all hospital dispositions were due to death certainly, quality would be judged as low, but efficiency could be deemed high. Yet, determining appropriate measures of quality is challenging. Structural measures, such as teaching status, are the easiest to model, while outcome measures, such as mortality, may be more meaningful since they represent a patient’s “bottom line” (Romano & Mutter, 2004). Process measures such as whether aspirin was given to myocardial infarction patients upon arrival may also provide meaningful information on quality, yet the connection to outcomes may not be direct. Past research has modeled quality to varying degrees and has produced inconsistent results regarding its relationship with efficiency. Rosko and Mutter (2008) provided perhaps the best example of explicit modeling of quality in a frontier efficiency study. They included twelve measures of quality as defined by the Agency for Healthcare Research and Quality (“AHRQ”) as explanatory variables of inefficiency. In-hospital mortality due to pneumonia and incidence of infection due to medical care are two examples of the variables examined. While these variables improved overall model fit, they did not have a significant effect on overall cost efficiency (Rosko & Mutter, 2008). Other research also found insignificant quality/efficiency correlations (Zuckerman, Hadley, & Iezzoni, 1994) (Deily,
McKay, & Dorner, 2000). Some research ignored quality, rationalizing that its abstract nature made modeling it too difficult (Burgess & Wilson, 1996) or that it was unlikely to have a significant effect (Rosko, 2001). However, other efficiency research found a significant correlation between increased efficiency and higher quality (Nayar & Ozcan, 2008). Incorporating three process measures of quality as outputs, the authors found that higher efficiency did not have to come at the expense of quality (Nayar & Ozcan, 2008). Finally, some researchers have modeled the quality-efficiency relationship in a completely different way—by using efficiency scores as an independent variable in a regression with a dependent variable representing quality—and have found a direct, significant correlation between greater efficiency and higher quality. In a study using Joint Commission on the Accreditation of Healthcare Organization (“JCAHO”) scores as the dependent variable in the 2-stage analysis, estimated inefficiency scores became the key explanatory variable (Harrison & Coppola, 2007). In a similar approach, McKay and Deily (2008) used estimated efficiency scores as the key explanatory variable in two separate regressions with observed mortality and complications rates as the dependent variables. They found that increased focus on cost-efficiency did not have to result in lessened quality. In healthcare, this approach is intellectually appealing, given the importance of producing high-quality outcomes: efficiency serves merely as means of achieving them.

Physician Characteristics
The characteristics of physicians practicing within a facility are environmental factors unique to healthcare that could influence the facility’s efficiency. Pauly (1980) viewed hospitals as workshops for physicians. Throughout a hospital stay, the physician directs “production” of healthcare (influencing the consumption of hospital labor and supplies) under her command. Harris (1977) noted that the physicians practicing in hospitals are critical members of an administrator-run team, yet the patient-doctor relationship compels doctors to serve a separate managerial role. “The net result is one organization
split into two disjoint pieces, each with its own objectives, managers, pricing strategy and constraints” (Harris, 1977). “There is a special negative externality in an arrangement in which one makes repeated marginal decisions about life and death... Whether or not it is justified, this notion has an important influence on the way the hospital is organized” (Harris, 1977). It has been estimated that physicians are responsible for 80% of hospital resource utilization (Chilingerian, 1995). Clearly, physicians hold a pivotal position in hospital care, and if individual motives and behaviors have any bearing on the production function, physician actions must influence the efficiency of hospitals to some degree.

However, in most civilian hospitals, physicians are not hospital employees: hospitals merely credential the physicians who work within its walls. Thus, the labor of credentialed physician FTEs is not included in the hospitals’ reported labor, yet the workload they manage and the inputs they use are included in the hospitals’ statistics. Reimbursement for a hospital procedure is split between the physician for her services and the hospital for use of the facility (Pauly, 1980). This construct forces researchers to exclude physician labor in most hospital-level efficiency studies. While common, however, this type of hospital-physician relationship is not absolute: physicians are predominantly employees in VA and military facilities, and other hospitals have a mix of physician arrangements.

It seems obvious that efficiency studies should integrate physicians’ motives and behavior into the modeled production function, yet (as just mentioned) data rarely allows this to occur. When explicit modeling of physician characteristics has occurred in efficiency studies, it has typically been at a clinic or patient level. Chilingerian (1995) examined the efficiency of 36 physicians within the same hospital, finding that HMO physicians and specialists were most efficient and that decreasing returns to scale set in as workload increased for physicians who saw higher-severity patients. Gaynor and Pauly (1990) developed a physician behavioral function and integrated it to the
healthcare production function to examine physician group practices. The authors found that while incentives increased output, they did not affect efficiency. They also found that physicians with more experience and those working in a smaller group practice had greater efficiency (Gaynor & Pauly, 1990). Although only limited to a single state, one study was able to utilize individual physician characteristics to assess their contribution to hospital-level efficiency for obstetric procedures. Focusing on Arizona hospital dispositions in 1989 – 1990, Burns, Chilingerian, and Wholey (1994) studied efficiency as defined by how far a patient’s length of stay and charges were below the hospital’s average. They found physician characteristics had a significant impact on hospital efficiency.

Patient Characteristics
The nature of healthcare makes the individual patient’s characteristics potentially important to the process of delivering healthcare, and potentially important determinants of efficiency. Frequently, studies of healthcare efficiency control for patient heterogeneity by adjusting output based on case severity using the average Diagnosis-Related Group (“DRG”) relative weight for Medicare patients, (i.e. the Medicare Case Mix (Rosko & Mutter, 2008)), and less often, all-patient average DRG relative weights. This is perhaps because most studies rely on the American Hospital Annual Survey and publicly available Medicare reports as their primary data sources. Brown (2003) did create an all-patient facility average severity index using DRG data from the Health Care Cost and Utilization Project, Nationwide Inpatient Sample (“NIS”), but made no further explicit patient-level adjustments, even though they were available. Studies that use clinical data sources typically utilize more patient-specific data (Bradford, Kleit, Krousel-Wood, & Re, 2001) (Burns, Chilingerian, & Wholey, 1994), but utilizing this data in a study of many facilities can be computationally demanding. Zuckerman, Hadley, and Iezzoni (1994) determined that the marginal information gained from including patient characteristics as efficiency explanatory variables did not outweigh the cost of obtaining and operationalizing them.
Progression of Hospital Ownership/Efficiency Studies

Pre-Frontier Analysis

A search of efficiency literature revealed that early efficiency research (prior to the mid-eighties) predominantly found for-profit organizations to be most efficient. This work was primarily either observational in nature, comparing certain ratios across ownership types and evaluating significance, or used Ordinary Least Squares Regression ("OLS"). Early studies explored efficiency in a general sense: specific types of efficiency were not examined. A brief discussion of these standard methods and pertinent studies follows.

Pre-Frontier Methods

Simple Ratio Analysis. Efficiency ratios are a staple of traditional financial statement analysis. For example, inventory turnover is a measure of how long merchandise held for resale remains on the premises, and return on assets is a measure of how much income a given level of assets produces. Efficiency ratios are important in non-financial analysis as well. Surgical procedures per provider and occupied bed days are examples in healthcare. Efficiency ratios are informative, but only with respect to the specific aspect measured. They do not allow for consideration of multiple inputs and/or outputs, and they do not accommodate analysis of interactions of these multiple inputs and outputs (Thanassoulis, Boussofiane, & Dyson, 1996). Ratio analysis usually involves comparing results to arbitrary benchmarks such as a percentage cutoff (Rosko & Mutter, 2008). Furthermore, ratio analysis does not focus on the production possibilities frontier.

Ordinary Least Squares Regression. Alternatively, OLS regression allows for consideration of multiple factors. Some research used a given efficiency ratio as the dependent variable, with potential explanatory variables on the right-hand side.
Chirikos and Sear (2000) noted that previous research modeled a specific production function but evaluated it in an OLS framework. Becker and Sloan (1985) examined cost per patient day and cost per admission along with two revenue-to-cost measures. Burns, Chilingerian, and Wholey (1994) compared individual physician mortality rates to hospital average mortality rates. With this type of model, a residual of zero is interpreted as average efficiency: positive residuals represent above average efficiency, and negative residuals represent below average efficiency. Yet this interpretation is questionable since it does not factor in random variation (Hollingsworth & Peacock, 2008). Furthermore, OLS results in information loss due its averaging out effects (Rosko & Mutter, 2008) and usually results in a downward-biased intercept (Coelli, Prasada Rao, O'Donnell, & Battese, 2005). Like ratio analysis, OLS does not focus on the production possibilities frontier. Corrected Ordinary Least Squares Regression (“COLS”) attempts to correct the usual downward bias of OLS by estimating a cost (or production) function and then shifting the OLS regression line to pass through the observation with the lowest residual. However, this technique forces the frontier to be parallel to the OLS regression line, by shifting it to pass through the observation with the smallest residual, and thus produces a deterministic frontier that does not allow for random error (Rosko and Mutter, 2008). Furthermore, by forcing the frontier to be parallel to the OLS regression line, COLS does not estimate the true production possibilities frontier.

Pre-Frontier Studies
Although not healthcare-related, Boardman and Vining (1989) and Vining and Boardman (1992) examined efficiency – specifically sales per employee and sales per asset ratios across ownership types – in top-500 non-U.S. entities using OLS. Both studies found state-owned organizations to be less efficient, however, frontier techniques (discussed next) were available at the time these studies were performed, and this invokes curiosity as to whether their use might have changed the results. Furthermore, in their 1992 work, the authors cited twenty healthcare-related studies finding private
healthcare organizations to be more efficient than public ones, three studies finding no
difference, and only one study finding public healthcare organizations to be more
efficient (Vining & Boardman, 1992). Five of these studies (and one additional article)
are discussed below.

Clarkson (1972) found empirical evidence that non-proprietary hospital managers
behaved in ways that might lead to greater inefficiency. He found they allocated less
time to unpleasant tasks such as personnel management, spent less time working
undesirable second and third shifts, and gave less attention to market information. He
also found nonproprietary hospitals exhibited greater variability in input selection, and
this variability is assumed to directly relate to efficiency (Clarkson, 1972). Lindsay
(1976) found that VA hospitals focused more on easily observed goals under the
purview of Congress (which could actually appear as greater efficiency) and less on
harder-to-observe quality-related characteristics. Herzlinger and Krasker (1987) found
for-profits had lower operating costs and made better use of capital and labor than not-
for-profits, and at the same time (contrary to popular opinion), they did not engage in
“cream-skimming” for better-insured patients and did not deny care to the poor. Frech,
in an analysis of Medicare claims processing organizations, compared “the performance
of firms with different types of property rights in providing a standardized product
under contract to the federal government” (Frech III, 1976). He found for-profits
significantly outperformed non-profits in cost per claim, number of claims processed,
and errors per 1,000 claims, and attributed a portion of this difference to inappropriate
scale of operations.

One of the few articles Vining and Boardman (1992) mentioned as finding no difference
between for-profits and not-for-profits was by Becker and Sloan (1985), who have
researched prolifically the operation of not-for-profit hospitals. They found no
significant differences in efficiency due to differences in ownership based on OLS
regressions of financial performance including cost per patient day, cost per admission, patient revenue to total cost, and total revenue to total cost. Friedman and Shortell (1988) also found insignificant differences in performance between not-for-profit and for-profit hospitals during the transition from cost-based reimbursement to Medicare’s Prospective Payment System (from 1983 to 1985). Additionally, they found that not-for-profits’ slightly lower profitability improved with respect to that of for-profits during the transition due to larger decreases in for-profits’ admission volume.

**Frontier-based research**
The development of two frontier measurement techniques occurred in the late 70s, and application of these techniques to studies of hospitals began in the mid- to late-80s. The first application of Data Envelopment Analysis (“DEA”) to hospitals is generally attributed to Sherman’s (1984) study of seven Massachusetts teaching hospitals, although Wilson and Jadlow employed the method two years earlier in a study of hospital nuclear medicine services (Wilson & Jadlow, 1982). The first application of Stochastic Frontier Analysis (“SFA”) to hospitals occurred in 1989 with a study of 49 Spanish hospitals (Wagstaff, 1989). A very brief introduction of each technique follows. Additionally, SFA will be discussed in more detail in upcoming chapters, and DEA will be discussed in more detail in Chapter 9. Since the introduction of these methods into healthcare efficiency measurement, for-profits no longer win the debate on which type of ownership is most efficient. What follows is a brief explanation of these frontier techniques.

**Frontier Fundamentals**
Frontier analyses involve estimation of the relevant production possibilities frontier. In general, efficiency enumerates the relationship between the inputs (usually land, capital, and labor) and the outputs of the production function, which defines the possible combinations of inputs and the resulting outputs (Hollingsworth & Peacock,
Production possibilities curves (frontiers) are widely studied in economics. All points on a production possibilities frontier curve represent the “maximum output attainable from each input level” (Coelli, Prasada Rao, O'Donnell, & Battese, 2005). In the two-product graph below, points lying beneath the frontier (X, e.g.) can be considered part of the feasible production set: they are achievable, but are inefficient. More of either product could be produced without reducing production of the other so that X could move to the frontier at either point A or C. Points above the frontier (Y, e.g.) are unachievable given current technology.

Figure 2.1: Production Possibilities Frontier

Production functions in healthcare are not straightforward. Individuals demand health – not healthcare – and thus demand for healthcare is derived from the demand for health and well-being. The supply of healthcare is complicated by asymmetry of information, the predominance of non-price competition, and regulations. Using health outcomes as the unit of analysis in an efficiency study presents problems due to lack of a clear link between cause and effect and difficulties in quantification. Because of these issues, intermediate products of healthcare such as inpatient dispositions and outpatient visits are often used as outputs (Hollingsworth & Peacock, 2008).
There are different measures of efficiency. Farrell developed the concept of radial measures of efficiency and divided efficiency into two components: 1) technical and 2) allocative efficiency (Hollingsworth & Peacock, 2008). Technical efficiency refers to maximum output given a set of inputs or to minimum inputs given a set level of output. Because costs are not considered, technical efficiency does not imply cost minimization or benefit maximization. Allocative efficiency is a broader concept, factoring in a cost component. An institution that is allocatively efficient is using the right combination of inputs, given what they cost (Salerno, 2003).

The difference between these two concepts can be seen in the cost frontier diagrams below. The diagrams depict the possible combinations of two inputs (in this case, staff and computers) to produce some level of educational achievement (medical education, e.g.). In the diagram on the left (considering only technical efficiency), points A and J are considered efficient because each lies on the frontier of production possibilities represented by isoquant (B), using a minimal combination of staff and computers (inputs). The diagram on the right considers cost with the inclusion of an isocost line (C). Point A remains technically efficient but it is not allocatively efficient: the line segment between A and A’ represents this allocative inefficiency. On the other hand, point J – located at the point of tangency – is both technically and allocatively efficient.
Figure 2.2: Graphical Depiction of Technical and Allocative Efficiencies

Cost efficiency is the product of technical and allocative efficiency, represented by the ratio of the line segment from the origin to the isocost line to the line segment from the origin to the actual observation. The above concepts are applicable to outputs and revenues with respect to the production possibilities frontier as well (Coelli, Prasada Rao, O’Donnell, & Battese, 2005).

Scale efficiency refers to the overall size of a firm’s operations. Constant returns to scale – the long-run outcome among competitive firms – signifies perfect scale efficiency (Salerno, 2003). If a firm’s scale of operations is too small, it is operating at increasing returns to scale. If a firm’s scale of operations is too large, it is operating at decreasing returns to scale. In both cases, the firm should adjust the size of operations to become perfectly scale efficient (Coelli, Prasada Rao, O’Donnell, & Battese, 2005).

A refinement of the concept of technical efficiency is “x-inefficiency.” Some view X-inefficiency and technical efficiency are interchangeable concepts (Ruggiero, Duncombe,
& Miner, 1995), however Leibenstein (1978)—the developer of x-inefficiency—did not specify this interchangeability. Others note that the two concepts are different, and have attributed the difference in technical and x-inefficiency to be due to their objectives: measurement is the objective of technical efficiency, and identification of the cause is the objective of x-inefficiency (Ruggiero, Duncombe, & Miner, 1995).

Frontier Methods of Measuring Efficiency

Given the flaws of standard methods for measuring efficiency discussed previously, two methods based on the production possibilities frontier have become common for measuring efficiency—Stochastic Frontier Analysis (“SFA”) and Data Envelopment Analysis (“DEA”). Both methods are based on the work of Farrell (1957) on radial measures of efficiency.

“The fundamental assumption is to depart from the assumption of perfect input-output allocation but to allow for inefficient operations. Inefficiency is defined as the distance of a firm from a frontier production function accepted as the benchmark. The basis for this measure is the radial contraction/expansion connecting inefficient observed points with (unobserved) reference points on the production frontier” (Fiorentino, Karmann, & Koetter, 2006).

Both methods attempt to overcome difficulties associated with the use of simple ratio analysis and ordinary least squares regression by focusing on the production possibilities frontier and allowing consideration of multiple variables. These frontier analyses differ from traditional ordinary least squares regression, which estimates average performance rather than possible performance: both methods establish a frontier representing implementation of best practices, based on the observations under study. SFA and DEA each have unique advantages and disadvantages.

Stochastic Frontier Analysis

In 1977, Aigner, Lovell, and Schmidt and Meeusen and van den Broeck proposed the concept of Stochastic Frontier Analysis independently. SFA is based on the idea that a frontier production function represents the maximum output possible, given a set of
inputs. Since the frontier represents an upper bound of production levels, the resulting error due to inefficiency is one-sided – a subtraction from the frontier. SFA is stochastic, meaning that it allows for the possibility of random error. It is also parametric, meaning the researcher must specify a frontier functional form (linear, log-linear, Cobb-Douglas, translog, e.g.) for the model. SFA essentially divides the traditional OLS error term into two pieces – inefficiency and random noise. The inefficiency component \((u_i)\) is assumed strictly positive, and a half-normal distribution is typical, although truncated-normal, exponential, and gamma distributions are also possible. “The \(v_i\)s [stochastic variability of the frontier] are assumed to be independently and identically distributed normal random variables with zero means and variances \(\sigma_v^2\)” (Coelli et al., 2005).

Technical efficiency studies using SFA are generally limited to one output. Focusing on cost as the dependent variable is one way to fix this limitation, and this is the more common approach used in hospital studies, where multiple outputs exist (dispositions, outpatient visits, surgeries, bed-days, e.g.). When studies of efficiency use a cost-focused SFA approach, multiple outputs can be considered: output quantities and input prices become the dependent variables\(^1\). The resulting inefficiency scores cannot be separated into technical or allocative inefficiency, however. In a cost function, the frontier represents a lower bound of cost levels, and the resulting cumulative technical/allocative error due to inefficiency is one-sided – an addition to the cost frontier.

Some researchers have highlighted problems with empirically estimating efficiency using SFA. Skinner (1994) pointed out that deviations from assumptions about the error terms— in particular the assumption of a homoscedastic non-skewed \(v_i\) (the stochastic variability of the frontier) – could bias inefficiency estimates. A \(v_i\) that is in actuality negatively skewed would manifest itself in skewness of the overall error term \((u_i + v_i)\). The required assumption of zero skewness for \(v_i\) would mean that its actual negative

\(^1\) This can also be applied to a revenue frontier.
skew would be attributed to $u_i$, and bias inefficiency estimates downward. Using visual comparisons, he also questioned the ability of maximum likelihood stochastic frontier estimation to detect inefficiency and separate the overall error term into two parts (Skinner, 1994).

To summarize, SFA’s advantages are its consideration of random error in estimating inefficiency and its use of econometric techniques allowing estimation of standard errors. SFA’s disadvantages are that it requires specification of both a functional form for the production function and a distribution of the deterministic inefficiency term, and that it does not easily accommodate multiple outputs.

Data Envelopment Analysis
The development of Data Envelopment Analysis (“DEA”) — a linear programming technique also known as the CCR ratio (after its creators’ names) — is attributed to Charnes, Cooper, and Rhodes (O’Neill, 2008). DEA is a “multi-factor productivity analysis model for measuring the relative efficiencies of a homogenous set of decision-making units (DMUs)” (Srinivas, 2000). For every observation DMU, DEA identifies a comparison “peer” group that produced at least as much output as the reference DMU, but used fewer inputs. Then, using Farrell’s radial efficiency concepts, DEA determines how much that DMU could reduce inputs while maintaining production of the same output quantities. Multiple inputs and outputs are easily accommodated in DEA without having to aggregate into less meaningful indexes (Hollingsworth & Peacock, 2008).

A DEA efficiency score is essentially a ratio of weighed outputs to weighted inputs: a perfectly efficient decision-making unit (“DMU”) would receive a score of one, and would reside on the frontier. DEA is non-parametric, meaning no assumptions as to functional form are required (O’Neill, 2008). Non-parametric methods such as DEA reduce the possibility of specification error that exists in SFA. Its non-parametric nature also means that efficiency is determined solely on the sample observations themselves,
and thus it can be very sensitive to data outliers within the sample. It is also deterministic; meaning no estimation of an error term is involved. The absence of an error component means that the entire distance of the observation to the frontier is attributed to inefficiency: there is no allowance for consideration of “random noise” such as unexpected one-time expenditures out of a hospital’s direct control. DEA directly plots the production frontier from observed inputs and outputs. The frontier is comprised of “linear segments that interpolate between those observations with the highest ratios of output to input. The resulting frontier thus ‘envelops’ all the observations” (Smith & Street, 2005). Since DEA results in this piecewise production frontier of line segments from efficient observation to efficient observation, it is possible for the frontier to contain some pieces that are parallel to the x- or y- axes. The distance from the observed performance to the frontier is technical efficiency: the distance from that derived point on the frontier to the end of the linear segment it lies on would represent slack inefficiency, a concept specific to DEA. However, slack efficiency would disappear if there were infinite observations because the frontier function would become smooth (Coelli, Prasada Rao, O'Donnell, & Battese, 2005).

In general, DEA can only discriminate among inefficient entities – not efficient ones – because efficient entities all receive a score of one (Jacobs, 2001). However, some DEA researchers have developed a “super-efficiency” concept that does distinguish between decision-making units receiving a perfect score. In essence, super-efficiency calculates the reference frontier excluding data for the $i$-th firm. The program is run multiple times (once for each firm), and it becomes possible for the $i$-th firm to be more efficient than the frontier (Coelli, Prasada Rao, O'Donnell, & Battese, 2005). The absence of an error or stochastic disturbance term in DEA means that standard errors (and therefore, confidence intervals) cannot be estimated – a serious econometric problem.

To summarize, DEA’s main advantages are that it accommodates multiple inputs and outputs and that it minimizes the possibility of specification error, requiring neither
specification of a functional form for the production function nor specification of a distribution for the inefficiency term. Its disadvantages are its failure to consider random error, its inability to develop confidence intervals for the inefficiencies it derives, and its sensitivity to outliers.

Both SFA and DEA allow for calculation of inefficiency scores for each observation, creating a temptation to rank observations to be used as the basis for facility-specific funding (or other critical) decisions. However, research has found these point estimates can be sensitive to model specification. This sensitivity means that it may be more appropriate to use SFA and DEA for detecting overall trends (Jacobs, 2001) than for judging performance of a particular observation (hospital). Now that these frontier techniques have been explained, this literature review returns to a discussion of hospital efficiency studies that found different results with respect to for-profit ownership using these techniques.

**Hospital Ownership Frontier Studies**
Hollingsworth (2008) updated both a previous journal article (Hollingsworth, 2003) and a book (Hollingsworth & Peacock, 2008) on the status of efficiency studies in healthcare. By his count, as of 2006 there were 317 published healthcare-related studies of efficiency using frontier techniques. Forty-eight percent of these studies used DEA, with another 19% using DEA scores in a secondary regression (typically to explore possible causes of inefficiency. Only 18% used SFA. Over one-half of these studies were of hospitals. He highlighted four studies involving for-profit, twelve involving not-for-profit, seventeen involving public, and six involving federal facilities, yet other studies have included a categorical ownership variable as a control variable as well. Of the studies that he highlighted with a primary interest in specific ownership type, public hospitals actually produced the highest average efficiency score and for-profits produced the lowest – a result counterintuitive to the theories previously discussed regarding the incentives associated with property rights. Not-for-profits and federal
facilities fell in between. Hollingsworth, however, offered little explanation for these results, speculating the cause was either the nature of healthcare as an unusual economic good or differences in methodology between studies (Hollingsworth B., 2008).

DEA Studies Finding Not-For-Profit and Government Hospitals More Efficient
One of the studies highlighted by Hollingsworth and Peacock (2008) is Valdmanis (1990), which compared public and not-for-profit hospitals. Its exclusion of for-profits is not typical of hospital efficiency studies. This cross-sectional, multiple-input/multiple-output DEA study of 74 Michigan urban hospitals found public efficiency to be 97.8% and not-for-profit efficiency to be 88.1%. Although not specifically identified by Hollingsworth as ownership-focused, another study with similar results is a DEA analysis that found more government hospitals to be fully efficient than either for-profit or nonprofit facilities (Ozcan, Luke, & Haksever, 1992). In this national cross-sectional study of 3,000 urban U.S. hospitals, the authors found 57.1% of government hospitals were fully efficient, while only 43.2% of for-profits and 36.5% of not-for-profits were fully efficient. Additionally, the percentages of highly inefficient (hospitals whose efficiency scores fell in the lowest quartile) were 21% for government hospitals, 22% for not-for-profit hospitals, and 35% for for-profit hospitals. These differences were statistically significant. In addition, the relationship of the performances of government hospitals relative to private ones held through analyses of several control variables – size, competition, system ownership, and region (Ozcan, Luke, & Haksever, 1992). Neither of these studies specifically examined causes of these efficiency differences in a regression framework, however. Both of these studies are representative of how DEA analyses handle variation of sample subjects. Typically, the sample is designed to be as homogeneous as possible with respect to possible confounding characteristics such as size or setting.
SFA Studies Finding Not-For-Profit and Government Hospitals More Efficient

While DEA has been the more popular method for studying hospital efficiency, SFA has been used with increased frequency (Hollingsworth B., 2008). Using SFA, Rosko found for-profit ownership to be positively and significantly correlated with greater inefficiency in two separate studies that examined relationships between certain structural characteristics and efficiency (Rosko, 2001) (Rosko, 2004). Both of these studies applied a cost function to panel data, although the 2004 study focused on teaching hospitals. Rosko (2001) found an overall efficiency of 84.7%, while Rosko (2004) found an overall efficiency of either 86.9% or 88.8%, depending on how the teaching mission was specified. Deily, McKay, and Dorner (2000) also found for-profits were statistically more inefficient than not-for-profits or government hospitals using SFA. While they did not explore causes for the differences, the authors did find that inefficient for-profit hospitals were more likely to exit the market than the other two facility types. Finally, although the primary research interest was the effect of managed care penetration on efficiency, Brown (2003) found for-profit status to be significantly associated with greater inefficiency in four of his five models exploring the effects of variation in patient condition severity. His analysis used an unbalanced panel of 613 hospitals covering five years – from 1992 through 1996, and he utilized a different source, the HCUP NIS, in addition to the usual AHA Survey (Brown, 2003). This was one of the few analyses of hospital technical efficiency using SFA found in the literature.

Studies Finding No Difference In Efficiency By Ownership Type

Burgess & Wilson (1996) (1998) introduced federal ownership as a fourth ownership type by including VA, other government, for-profit, and not-for-profit hospitals in two separate studies using DEA. The first study found statistically significant differences between the four groups. VA hospitals fared well in terms of radial efficiency, but less well in terms of scale and slack efficiencies. However, the authors could not quantify differences in overall technical efficiency and they did not analyze possible causes of the inefficiencies (Burgess & Wilson, 1996). Their 1998 study delved further into possible
causes of inefficiency by using a two-step analysis, which estimates technical efficiency in the first step and regresses potentially related variables against the efficiency scores in the second step. These variables included competition, length of stay, percentage of nurses who were RNs, and administrative costs per bed day. This research did not find significant differences across ownership types after controlling for these other variables (Burgess & Wilson, 1998). Although they did not include VA facilities, Register and Bruning (1987) found similar non-significant differences in efficiency in another two-step DEA and regression analysis cross-sectional study of 457 for-profit, not-for-profit, and government hospitals. Bruning and Register’s (1989) DEA analysis and Vitaliano and Toren’s (1996) cost-based SFA analysis found similar non-significant differences in efficiency across ownership type. All of the studies discussed in the last three sections suggest that property rights and greater efficiency may not be correlated as often hypothesized in the literature.

SFA Estimation Issues and Advances
Some researchers have probed the possibility that pooling of groups facing different true cost/production frontiers into a single frontier estimation creates specification bias. Comparing different ownership types may create this problem. In a pooled estimation of such different groups, parameter estimates, errors, and estimated inefficiencies could be biased. Folland and Hofler (2001) focused on this issue (along with production function specification) in a cost-based SFA cross-sectional study of 1985 U.S. hospitals. The pooled model found for-profit inefficiency to be 16.2% and not-for-profit inefficiency to be 12.6%. Partitioned models found for-profit inefficiency to be 11% and not-for-profit inefficiency to be 8%. The differences between pooled and partitioned models, while not large, were statistically significant. Correlation of efficiency scores among categories (rural/urban, for/non-profit, e.g.) between pooled and partitioned models was high. However, correlation of individual hospital efficiency rankings between the models was low. Zuckerman, Hadley, and Iezzoni (1994) found similar results with respect to pooling and partitioning in an analysis that also included
government hospitals. In their cost SFA study of 1988 U.S. hospitals, public hospital inefficiency was 14.1% in pooled data and 23.3% in partitioned data: for-profit hospital inefficiency was 14.4% in pooled data and 19.5% in partitioned data: not-for-profit hospital inefficiency was 12.9% in pooled data and 11.8% in partitioned data. They found moderate correlation of individual hospital rankings between pooled and partitioned models. The conclusion (informed by Zuckerman, Hadley, and Iezzoni’s (1994) work) that Folland and Hofler drew from their analysis is that estimates of mean group inefficiencies are robust to the issue of partitioning, but individual hospital inefficiency estimates may not be (Folland & Hofler, 2001).

Adaptations of the SFA technique have allowed for consideration of multiple outputs and multiple inputs in studies of technical efficiency (as DEA does), an appealing concept in studying healthcare where there are usually several categories of outputs. Although neither study examined the effects of ownership type, Gerdtham, et al. (1999) and Ferrari (2006) both employed variations of the Shepard’s distance function to study hospital technical efficiency using SFA. Ferrari (2006) used distance functions and data on 52 English hospitals over a 6-year period to evaluate the impact of introducing internal competition. Gerdtham, et al. (1999) used multiple-output stochastic ray analysis and data on 26 Swiss public hospitals over a 7-year period to investigate the effect of reimbursement reform on technical efficiency. In these studies, the distance from an observation to the frontier \( D_0 \) is the measure of technical efficiency. Gerdtham, et al. (1999) employed Euclidean geometry to define polar coordinates that are then included in the production function and allow for measurement of \( D_0 \). Ferrari’s distance function was based on a variation of SFA proposed by Coelli and Perelman in 1996 and applied in 2000 (Coelli & Perelman, 2000). Capitalizing on an assumption that the function is linearly homogeneous in the outputs, \( D_0 \) is expressed as a function of \( M \) outputs, \( K \) inputs, and \( N \) observations and then only requires algebraic manipulation of the model (Ferrari, 2006). However, this requires several econometric assumptions,
including the assumption that the coefficients on the output variables sum to one (Coelli, Prasada Rao, O'Donnell, & Batte, 2005).

Summary
Research focusing on hospital efficiency prior to the mid-eighties supported the established theories on the effects of property rights and related profit incentives: for-profits were found to be more efficient than either not-for-profit or government hospitals in the majority of studies. The introduction of two frontier-based methods, SFA – an econometric technique – and DEA – a linear programming technique – to studies of hospital efficiency occurred in the mid eighties. Both enable evaluation of efficiency based on performance with respect to the production possibilities frontier rather than on average performance or simple ratio comparisons and some have suggested that these frontier methods reflect aspects of efficiency not captured by traditional methods (Fiorentino, Karmann, & Koetter, 2006).

A review of hospital efficiency studies performed after the introduction of these techniques to healthcare indicates that the effects of ownership may not be so straightforward: the seeming certainty of greater efficiency in for-profits disappeared. The performance of not-for-profits and governmental hospitals relative to for-profits appears to have improved in these later studies. The reasons for this are not entirely clear. Perhaps the introduction of frontier techniques influenced the results. Many studies, in particular those using SFA, focused on cost rather than technical efficiency. Perhaps, as Pestieau and Tulkens (1990) theorized, these cost-based studies produced biased results because technical efficiency is the only “fair” way to evaluate governmental organizations, and therefore the significance of property rights persists. However, at about the same time that frontier techniques were introduced to healthcare (the mid-eighties), Medicare’s Prospective Payment System for hospitals was introduced. Perhaps this increased regulation (or other environmental factors) forced
hospitals of all types to function more similarly; diminishing the effects of property rights and their incentives: not-for-profits and governmental organizations truly became more efficient, equaling or surpassing not-for-profits.

Throughout this literature, investigation of differences in performance among hospitals controlled by different levels of government – in particular federal facilities – have been under-represented. With this background of efficiency studies examining hospital ownership, Chapter 4 now turns to an overview of the federally controlled military healthcare system and a discussion of military-specific efficiency studies.
Military Health System Facts

The Military Health System (“MHS”) is a major provider of federally provided healthcare, along with the Veterans’ Administration, the Public Health Service, and several smaller agencies. The mission of the Military Health System is “to enhance the Department of Defense and our nation’s security by providing health support for the full range of military operations and sustaining the health of all those entrusted to our care” (Tricare Management Activity). Tricare – a key component of the Military Health System implemented in 1992 – is the managed care program that integrates military (“direct” or “in-house”) and civilian (“purchased”) health care assets to support this mission. Tricare Management Activity (“TMA”) is the organization that manages the Tricare health program and executes policies issued by the Assistant Secretary of Defense for Health Affairs. Tricare uses civilian organizations via managed care support contracts to integrate both direct and purchased care in three continental U.S. regions and three overseas regions (Tricare Management Activity). The scope of services managed by TMA is vast, including soldiers treated at units near the battlefield and aboard ship as well as retirees treated by civilian providers in their hometown.

Tricare provides health insurance for 9.2 million beneficiaries, of whom only approximately 1.7 million are Active Duty personnel: the bulk of these beneficiaries are family members and retirees. According to the 2008 MHS Stakeholders’ Report, there are currently 63 hospitals, and 413 medical and dental clinics within the Military Health System. The FY07 Defense Health Program (“DHP”) appropriation budget – that covers general operating expenses, wages for civilian providers in military facilities, and purchased civilian care – was $23.7 billion, and the budget for pay of Active-Duty
personnel working in military facilities was $6.9 billion\textsuperscript{2}, for a total of $30.6 billion allocated to the daily provision of healthcare for Tricare beneficiaries. In an average week, the MHS will see 18,500 inpatient admissions (4,800 in military facilities), 2,200 births (1,000 in military facilities), 664,000 direct care outpatient visits, 2.3 million filled prescriptions, and 3.7 million processed claims (Tricare Management Activity, 2008). Clearly, the Military Health System is a non-trivial provider of health care in the U.S. Furthermore, it is not immune to the cost pressures felt in the civilian health care sector.

**MHS Operating Environment**

Direct care provision is supported by considerable infrastructure, personnel, supplies, and equipment and thus entails a high percentage of fixed or semi-fixed costs. Civilian care, purchased for beneficiaries who opt to receive medical services in the private sector and for in-house patients referred out as needed, is comprised of essentially all variable costs. When care is referred out from the direct care system to the civilian sector, the DoD essentially incurs a double bill since the costs of the direct care system are paid regardless of whether care is provided. Thus, efficiency in the direct care system takes on even greater importance.

As the civilian sector introduced managed care and increasingly shifted the burden of healthcare costs back to the individual, the military was constrained by Congress to absorb increases within budget. Minimal cost sharing was successfully introduced with the implementation of TRICARE in 1995, but fees and copayments have not increased since inception. “Sustain the Benefit” is the moniker for an initiative to increase fees and copayments introduced in 2005 that has yet to win Congressional approval (Government Accounting Office, 2007).

\textsuperscript{2} Operation and maintenance of deployable assets such as the Hospital Ships and Expeditionary Medical Facilities are budgeted for separately by the services to which they belong.
Rapidly increasing health care costs have sparked interest in efficiency in all sectors, but perhaps especially in the military. The position of the DHP within the DoD budget\(^3\) places medical care for service members, retirees, and dependents in direct competition for funds with weapons development, ship/aircraft operations, and other direct military expenditures. In 1990, the Defense Health Program accounted for 4.5% of the total Department of Defense budget. By 2015, analysts project this percentage to grow to 12% (Department of Defense Task Force on the Future of Military Health Care, 2007). Given the ongoing global war on terror and the need to maintain state-of-the-art warfighting capabilities, in addition to daily maintenance requirements, absorbing such increases will be increasingly difficult.

**Role of Efficiency in Recent MHS Decision-Making**

Efficiency frequently appears as a goal or a concern in strategic planning and decision-making initiatives for the MHS. Four references to efficiency are discussed here.

1. Based on a Balanced Scorecard approach, the Financial Perspective Strategic Objective of the MHS Strategic Plan is to ensure that “The MHS health care delivery system will be engineered to achieve optimal efficiency and mission effectiveness” (Office of the Assistant Secretary of Defense for Health Affairs, 2007).

2. Efficiency was a key decision point for the Medical Joint Cross-Service Group in drafting the 2005 Base Realignment and Closure Act. While the decision to close Walter Reed Army Medical Center was the most publicized BRAC medical decision, downsizing (including cessation of inpatient care missions) at several other military hospitals was also recommended. Downsizing recommendations were based in part on facility efficiency, as measured by Average Daily Patient Load (Defense Base Closure and Realignment Commission Medical Joint Cross-Service Group, 2005).

\(^3\) The DHP is often referred to as an entitlement lodged in a discretionary appropriation.
3. The growing budgetary pressures that have been discussed previously led to the creation of a taskforce to evaluate the sustainability of the Military Healthcare System in the future. The taskforce was co-chaired by the Vice Chief of Staff for the Air Force and a leading national health economist. It “endeavored to find the right balance between ensuring a cost-effective, efficient, and high-quality health care system for military beneficiaries and managing a system with spiraling costs that, if unchecked, will continue to create an increasing burden on the American taxpayer” (Department of Defense Task Force on the Future of Military Health Care, 2007).

4. In its February 20, 2007 proceedings, the DOD Task Force on the Future of Military Care just discussed heard testimony from the Surgeons General of the Army, Navy, and Air Force on the challenges the services face in providing quality health care to their beneficiaries within their allotted budgets. One of the biggest challenges each Surgeon General discussed was the “Efficiency Wedge” – a lump-sum cut made by TMA in previous budgeting cycles for years beyond the President’s Budget at the time. The cut was not tied to a specific identified excess or change in business practice; it was essentially justified by an assumption that inefficiencies existed and that addressing them would yield savings. Approximately $147M of the wedge was coming to reality in 2007 for the direct care sector (provided in military hospitals) (Department of Defense Task Force on the Future of Military Health Care, 2007).

Clearly, efficiency has been stressed in decision-making, yet analysis of it has been limited. Assessments have been predominantly based on analysis of only a few ratios if any measurement even occurred. In addition, there has been no baseline analysis of efficiency levels by which to measure improvements.

**How Unusual are Military Hospitals?**

Philosophically, military and civilian hospitals should have the same overall goal: to cure the patients they treat. Additionally, military hospitals within the fifty states undergo
the same JCAHO inspections as civilian hospitals, implying that quality should be comparable. However, some unusual characteristics of military operations might influence the behavior of health care providers in achieving their goals, thereby affecting efficiency. Differences from civilian hospitals become apparent when examining data on all U.S. hospitals, as in Table 6.2: General U.S. Short-Term Medical/Surgical Hospital Raw Data. For example, military facilities produce a much higher volume of outpatient workload. Unusual characteristics of military hospitals include:

- **Cost:** Active duty personnel and government employees are paid according to established schedules. Military hospitals may also receive better pricing for supplies via government schedules. Comparing technical efficiency as this dissertation does alleviates concerns about differences in cost.

- **Annual funding:** Military hospitals receive funding annually through Congressional Appropriations. This added level of bureaucracy could alter behavior of health care providers, although fundamentally, any effect should be similar to that of state and local hospitals that receive funding through governmental appropriations.

- **Business model:** In military hospitals, physicians are either employees who receive a salary for their services or contractors also paid directly from the facility’s budget. This is not the case for most civilian hospitals, where physicians are typically credentialed to practice within the facility. This can be explored at least nominally using data from the AHA survey and the fact that essentially 100% of military hospital physicians act as employees, whether active duty, civilian, or contractor.

- **Patient base:** The patient base of military hospitals is likely younger and healthier than the population served by civilian hospitals, and it is likely that the health needs of this patient base are different from patient populations (i.e. more likely to seek care for injuries). Recognizing these differences, TMA calculates its own relative weights for Diagnosis-Related Groups (“DRGs”), rather than merely using Centers for Medicare and Medicaid Services (“CMS”) weights which are based on the Medicare population (Tricare Management Activity, 2008). Yet young healthy patients are not
an absolute. Retirees under 65 remain eligible for care in the MHS, and Medicare-age retirees are eligible for care in military facilities on a space-available basis. Furthermore, case-mix adjustments of outputs and specific inclusion of demographics such as age, gender, and race in modeling can control for such differences.

- **Organizational Mission:** Secondary missions may alter the production process of hospital health care. The MHS has a dual mission: it maintains medical readiness of personnel – including themselves – for war and it provides a “benefit” mission, caring for all its beneficiaries, including family members and retirees. However, other hospitals also have secondary missions of teaching and research, and exclusion of labor related to secondary missions (such as time spent on military exercises) should control for time spent on tasks other than production of healthcare.

- **Malpractice:** Active Duty personnel cannot sue the government, even for cases of medical malpractice, thanks to a 1950 Supreme Court case (now known as the Feres Doctrine) (Pugatch, 2008). Dependents of Active Duty and retirees can sue the government for cases of medical malpractice, but the government – not the individual provider, pays any settlements. This is a result of the Federal Tort Claims Act of 1946 that waives the federal government’s sovereign immunity in certain circumstances, including claims of medical malpractice by federal employees. Federal rules represent a paradigm different from any of the fifty states’ malpractice provisions, and the freedom both from having to maintain malpractice insurance and from worry over personal lawsuits could fundamentally alter how physicians practice medicine. Since military physicians face no risk of financial liability from malpractice, they represent one end of a spectrum of malpractice effects faced by physicians nationwide. The effects of different malpractice legal paradigms on civilian physician behavior are observable – at least at the state level – by using data from the National Practitioner Data Base (Pugatch, 2008). With adequate military
data, further insights on how malpractice affects physician behavior (and efficiency) may be possible.

While it would not be valid to compare efficiency of an overseas military hospital treating soldiers transported from the battlefield to that of a U.S. civilian hospital, a comparison of civilian hospitals to military hospitals within the fifty states treating active duty, family members, and retirees for similar problems seems reasonable to attempt. However, comparisons of military and civilian hospitals have not occurred in the literature, as is discussed in the next section.

Studies of Military Hospital Efficiency

Comparisons of military and civilian health care efficiency are non-existent in academic journals. In the only published application of SFA to federal healthcare I have found, Schmacker and McKay (2008) analyzed technical efficiency of primary care facilities within the Military Health System. The authors found an average efficiency of 82.2% in an unbalanced panel study of 442 observations (both hospital-based and stand-alone departments) over five years. With respect to correlates of inefficiency, the authors found that the percentage of civilian staff was significantly directly correlated with greater efficiency and the clinic size/complexity were directly correlated with greater inefficiency. No significant differences were found between Army, Navy, and Air Force, and physician extenders (i.e. physician assistants and nurse practitioners) had no significant effect on efficiency.

The majority of efficiency studies of military hospitals have been DEA-based. Bannick and Ozcan (1995) compared VA and DoD hospitals using DEA, finding DoD hospitals more efficient (87% vs. 78%). Ozcan and Bannick (1994) examined trends in DoD hospital efficiency. The authors utilized a three-year panel of 124 hospitals to explore
the possibility of inter-service institutional differences in efficiency. Data came solely from the AHA Annual Survey, rather than directly from military sources. The authors used civilian hospital statistics as a benchmark, but again this was not an actual military/civilian comparison. The study found 59.7% of military hospitals operated efficiently, and the mean efficiency score was 95%. There were no significant differences between the Army, Navy, and Air Force. Additionally, the authors found modest correlation between DEA efficiency scores and traditional efficiency ratios.

Finally, in an unpublished dissertation, Van Fulton (2005) thoroughly examined efficiency of Army Medical Facilities using multiple methodological approaches, including DEA and SFA. As a recommendation of avenues for future study, his research actually compared care provided in military facilities (direct care) to care provided in network facilities (care provided to Tricare beneficiaries in private facilities) using DEA. In this comparison, the author found that, in general, civilian facilities were more efficient than military facilities, but that the more-efficient military facilities were generally located in less-efficient networks (Van Fulton, 2005). However, his comparison was of care provided to a particular population – Tricare beneficiaries – in different hospitals, not of care provided to all populations within the same hospital.

While analyses of military hospital efficiency are rare and comparisons to civilian hospitals non-existent, another segment of federal care – the VA – has been used to inform the dialogue on the effects of ownership. Burgess and Wilson (1996) and (1998), discussed in Chapter 2, compared VA hospitals to for-profit, not-for-profit, and state and local government hospitals. In addition to these comparative studies, several analyses have focused exclusively on VA hospital efficiency. Yaisawarng and Burgess (2006) utilized DEA to estimate efficiency of VA hospitals at 94% and then demonstrated how resource allocation based on these results in a performance-based budgeting framework suggest reallocating $267 million in annual funding from lower-performing
hospitals to higher-performing ones. Additionally, at least six other studies analyzed efficiency in VA hospitals, without making any comparisons to civilian hospitals [ (Sexton & et.al., 1989) (Harrison, Coppola, & Wakefield, 2004) (Harrison & Coppola, 2007) (Hao & Pegels, 1994) (Harrison & Ogniewski, 2005) (Burgess & Wilson, 1993)].

Research Questions
As stated at the outset, the primary research question this dissertation explores is, “Are military hospitals inherently more technically efficient (or inefficient) than hospitals controlled by other types of ownership?” In addition, the following are secondary research questions:

- Is ownership a significant variable in estimating technical efficiency once other operational characteristics (such as exposure to competition, size, physician/patient characteristics, and quality) are controlled?
- Does the estimation method affect the results?
- In light of the gap identified in the literature, how does inclusion of military ownership in efficiency studies inform the overall body of ownership research?
CHAPTER 4 - CONCEPTUAL FRAMEWORK

Conceptual Framework

Stochastic Frontier Analysis is the primary measurement method in this dissertation. SFA investigations of technical efficiency typically examine a defined process, of a general form:

\[ y = f(x_1, \ldots, x_n) \]

In words, a given output is a function of a number (n) of given inputs. Commonly defined inputs are capital and labor, although land, raw materials, and machinery may also be inputs, depending on the process.

Defining the production function requires giving f(.) some type of algebraic form based on economic theory. Production functional forms are characterized by several properties. A flexible functional form has “enough free parameters to provide a local second-order approximation to any twice continuously differentiable function” (Barnett & Usui, 2006). Flexible functional forms typically use quadratic terms that are obtained from second-order series expansions. A parsimonious functional form reflects the simplest function that “gets the job done adequately” (Coelli, Prasada Rao, O'Donnell, & Battese, 2005). Parsimony means that inclusion of more parameters or more general functional forms would generate statistically insignificant improvements in sums of squared prediction errors, likelihood, or whatever criterion is employed. Regularity entails maintaining monotonicity and curvature (concavity), as well as non-negative outputs and a requirement that at least one input is necessary for the production process. Attaining all of these properties at the same time is difficult. “Simultaneous imposition of both of these conditions [curvature and monotonicity] on a parsimonious flexible functional form destroys the model’s local flexibility property” (Barnett & Usui, 2006).
Over time, the use of several production functions have become commonplace. These include the quadratic, normalized quadratic, translog, generalized Leontief, and constant elasticity of substitution. The two most common functional forms used in healthcare efficiency research, Cobb-Douglas and Translog, are shown below in a general one-output two-input model.

**Cobb-Douglas**

\[ y = \beta_0 \prod_{n=1}^{N} x_n^{\beta_n} \]

Written in log form:

\[ \ln(\text{Output}) = \alpha + \beta_1 \ln(\text{Input 1}) + \beta_2 \ln(\text{Input 2}) \]

**Translog**

\[ y = \exp(\beta_0 + \sum_{n=1}^{N} \beta_n \ln x_n + \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \beta_{nm} \ln x_n \ln x_m) \]

Written in log form:

\[ \ln(\text{Output}) = \alpha + \beta_1 \ln(\text{Input 1}) + \beta_2 \ln(\text{Input 2}) + \beta_3 \frac{1}{2} [\ln(\text{Input 1})]^2 + \beta_4 \frac{1}{2} [\ln(\text{Input 2})]^2 + \beta_5 \ln(\text{Input 1}) * \ln(\text{Input 2})] \]

The transcendental logarithmic ("translog") model, developed by Christensen, Jorgenson, and Lau in 1971, employs Taylor series expansions in logarithms (Barnett & Usui, 2006) and contains both linear and quadratic terms. The Cobb-Douglas function is essentially a special case of the translog function, where the parameters of squared and
cross products are restricted to zero. Thus, the translog function is more flexible than the Cobb-Douglas function. However, the Cobb-Douglas is more parsimonious as long as it adequately describes the production process (i.e., the quadratic terms have statistically insignificant coefficients tested jointly).

The operations of any hospital can be viewed as a production function. Inputs (labor and capital) combine via medical and surgical care (the production function) to produce outputs. While the ultimate output of healthcare is the marginal change in health status, this is difficult to measure in most data sets, and so intermediate outputs – episodes of care (i.e. inpatient discharges and outpatient visits) – usually become the primary study outputs. Using a Cobb-Douglas functional form for production for discussion purposes, capital (typically beds) and labor (nurses, etc.) combine to produce these episodes of care as follows:

\[
\ln(\text{Episodes of care}) = \alpha \text{Beds}^\alpha + \beta \text{FTEs}, \quad \text{where}
\]

- \(\alpha\) and \(\beta\) represent output elasticities, meaning a 1% increase in Beds would lead to an \(\alpha\)% increase in Episodes of care. A 1% increase in FTEs would lead to a \(\beta\)% increase in Episodes of care, and
- the sum of parameters \(\alpha\) and \(\beta\) indicates returns to scale. \(\alpha + \beta = 1\) indicates constant returns to scale. \(\alpha + \beta > 1\) indicates increasing returns to scale. \(\alpha + \beta < 1\) indicates decreasing returns to scale.

This production process does not occur in a vacuum: other “non-stochastic environmental variables” (Coelli, Prasada Rao, O'Donnell, & Battese, 2005) may influence the process and thus may affect how efficiently it occurs. Often these factors are considered to be uncontrollable by managers of the process. These factors could be theorized either to affect the production process itself or to influence directly the
efficiency of the process (Kumbhakar & Lovell, 2000). The position taken in this dissertation is that these environmental variables affect how efficiently healthcare is delivered and not the actual process itself. Since ownership is a primary theme of this dissertation, it is a key topic of exploration along with other hospital structural characteristics (size, e.g.), environmental characteristics (competition, e.g.), quality, and physician and patient characteristics – all of which were introduced in Chapter 2. The diagram below depicts the conceptual model of the hospital production process and its associated technical efficiency:

Figure 4.1: Conceptual Model of Hospital Technical Efficiency
CHAPTER 5 - EMPIRICAL MODEL

Using the conceptual model developed in the previous chapter, the process of hospital healthcare provision can be specified as a function that quantitatively links inputs and outputs, and the associated efficiency can be modeled based on factors hypothesized to affect it. The general empirical model of the hospital healthcare production function and the related efficiency is depicted as follows:

\[ \text{Hospital Workload} = f(\text{Capital, Labor}); \text{ and} \]

\[ \text{Associated Efficiency} = g(\text{Ownership, Structure, Environment, Quality, MD Characteristics, Quality, Patient Characteristics}) \]

The elements of this model are described below.

**Hospital Workload (Output)**

1. **Total Work**: The primary output of a hospital is its inpatient dispositions, and they are the most common output measure used in efficiency studies (Rosko & Mutter, 2008). Additionally, many hospitals (especially military ones) have large outpatient clinics within their control, so outpatient visits are also a primary output. While other services such as surgeries or bed days may also be considered outputs, they are usually either sub-categories or alternate measures of dispositions and visits. Within these categories, there can be considerable heterogeneity. A tonsillectomy is clearly not the same as an organ transplant, yet a strict count of inpatient dispositions or surgeries would weight them equally. The typical method of accounting for this heterogeneity in inpatient workload is a case-mix adjustment based on Diagnosis-Related Groups (“DRGs”) assigned to each disposition, and often is based solely on Medicare cases (Grosskopf & Valdmanis, 1993). DRGs are codes developed by CMS and are based on the diagnosis, procedure(s), age, sex, and complications associated with a given disposition. An average DRG receives a weight
of 1.0. More resource-intense DRGs are weighted greater than 1.0; less resource-intense DRGs are weighted less than 1.0. Thus, a hospital performing work of average intensity would have an average case mix index of 1.0. Case mix adjusting also indirectly accounts for quality by “giving more credit” for more complex work.

Stochastic frontier technical efficiency studies generally ignore outpatient workload, as in Brown (2003) due to the one-dependent variable methodological limitation. However, outpatient workload is a large piece of military hospital workload, and ignoring it could bias any comparisons with the civilian sector. Case-mix adjustment for outpatient visits is more complicated than it is for inpatient workload because of the wide variation in defining what constitutes a visit. Visits can be as simple as a prescription refill or as complex as an outpatient surgical procedure.

One possible strategy is to adjust outpatient workload based on weighted CMS Ambulatory Payment Classifications (“APCs“) per visit similar to DRG-based adjustments for inpatient workload. CMS created APCs to apply prospective payment methodology to hospital outpatient work, and TMA now uses them as well. Facility average APC weights were obtained from Cleverley and Associates (Cleverley, 2009) and compared to military average APC weights obtained from military-specific data sources, but the data was clearly not comparable. There are several likely reasons. Data for civilian hospitals covers only the Medicare population, and data for military hospitals would likely have much less Medicare-age work. In addition, civilian data included certain drugs and pass-through biologicals in their calculations of averages, while military figures did not. Furthermore, TMA has only recently begun to use APC coding, and therefore there may be considerable measurement error for military facilities.
Alternatively, a broad adjustment for case mix by including a variable for the percentage of outpatient visits in the Emergency Room is a method sometimes used in the literature (Rosko & Mutter, 2008). A third alternative is application of the same case mix index used for inpatient work on the assumption that similar patients generate similar work, regardless of whether it is performed on an inpatient or outpatient basis. This analysis will consider the second and third strategies.

Standard technical efficiency SFA models are limited to only one dependent variable (output). This limitation necessitates aggregation of inpatient and outpatient workload into one variable. In the absence of guidance from literature, aggregation will be based on the ratio of median Medicare cost per outpatient visit to Medicare cost per discharge – a reasonable proxy for the relative resource requirements of inpatient and outpatient care. The ratio – based on 2006 case mix-adjusted Medicare costs per discharge and per inpatient disposition – is 0.012:1 (outpatient: inpatient) (Cleverley, 2009).

Application of a distance function model as in Ferrari (2006) would allow inpatient and outpatient workload to be included in the model separately – without the need for aggregating them into one variable. As discussed in Chapter 2, distance functions allow for inclusion of multiple outputs and inputs. While estimation models of distance functions often do not converge, a model variation using a distance function will be attempted and compared with standard one-output models of technical efficiency in Chapter 8. Distance function methodology will be discussed in Chapter 7.

**Inputs**

- **Labor** – RN FTEs and Other Non-MD FTEs
  
  Like the outputs, hospital labor also consists of several disparate categories – physicians, nurses, aides, administrative, etc. Consistent with other U. S. studies
(Burgess & Wilson, 1996) (Burgess & Wilson, 1998), physician FTEs are not considered in the production function due to data inconsistencies caused by the practice of credentialing. However, the role of physicians in the production function is important, and two aspects of this role are included in the inefficiency variables below. While some studies include an aggregate of all FTEs as the measure of labor (Brown III, 2003), some disaggregation is usually employed, with more labor categories used in DEA, as in Bruning and Register (1989). Registered nurses are a labor category used in the majority of hospital efficiency studies (Bruning & Register, 1989) (Burgess & Wilson, 1996) (Burgess & Wilson, 1998) (Ferrari, 2006). Furthermore, registered nurses actually manage patient care and execute physicians’ orders, making them arguably the most critical member of a healthcare team. Other Non-MD FTEs, while aggregating several categories, represent all other hospital employees including LPNs, paraprofessionals, administrative personnel, and ancillary employees. This conventional approach alleviates concerns over differences in classification of personnel by hospitals, especially between military and civilian facilities.

- **Capital – Hospital Beds**

Beds are a common choice to represent capital [(Brown III, 2003) (Ferrari, 2006) (Bruning & Register, 1989)]. Beds are discrete items and therefore easily counted, thus reducing the possibility of measurement error. The American Hospital Association maintains a National Master Facility Inventory separately from its annual survey process, ensuring availability and increasing the likelihood of accuracy. Square footage, while another possible choice for a capital variable (Schmacker & McKay, 2008), is more susceptible to measurement error when considering the entire hospital as the unit of observation.
Variables Influencing Efficiency

Ownership

The theories that have been put forth as to why ownership might influence efficiency were discussed in detail in Chapter 2. A categorical variable representing for-profit, not-for-profit, state/local government, and military (federal) hospitals will model the effects of property rights.

Structure

- Size: As discussed in Chapter 2, competing theories exist on the effect of size on efficiency. If economies of scale exist (Vitaliano & Toren, 1996) (Register & Bruning, 1987), efficiency and size are likely to be directly correlated. If size complicates the monitoring process (Hao & Pegels, 1994) (Sexton & et.al., 1989), they are likely to be inversely correlated. The number of beds is the typical measure of size. In this analysis, a binary variable, where “1” represents hospitals with less than fifty beds and “0” represents hospitals with fifty or more beds, will address the effects of size on efficiency.

- Scope of Work:
  - Overall hospital complexity (Case Mix Index) – Heterogeneity of work performed, as previously discussed, is a major concern of researchers. The standard method of controlling for this is to use DRGs – codes given to every disposition that reflect the relative complexity of treatment – to calculate an overall measure of a hospital’s work. Some researchers have concluded that case mix does not significantly affect measures of efficiency (Grosskopf & Valdmanis, 1993), while others conclude it does (Rosko & Chilingerian, 1999). Given the added concern in this study that military hospitals treat a very different patient base with different needs, consideration of case mix is important.
  - Major Diagnostic Category – Major Diagnostic Categories (“MDCs”) classify inpatient work into 25 categories based on the system of the body that is
treated, and are a representation of the scope of work a hospital performs. Variation in the range of MDCs a hospital performs could affect efficiency. As will be shown in Chapter 6, two MDCs where military hospitals vary most from civilian hospitals are MDC 14 (Pregnancy, Childbirth, and Puerperium) and MDC 5 (Circulatory System). While all MDCs will be analyzed for descriptive purposes, the effects of volume of work performed in these two MDCs will be included in the efficiency model.

- Percent Surgical DRGs – DRGs are classified as either medical or surgical. Analysis of DRG weight files revealed that on average, length of stay from surgical dispositions is both longer and exhibits greater variability than for medical dispositions. As such, a higher percentage of surgical DRGs may be associated with greater inefficiency.

- Percentage of Emergency Room Visits – In the absence of a separate case mix index for outpatient visits, the percent of total outpatient workload attributed to emergency room visits has been used as a proxy to control for outpatient heterogeneity. Emergency Room visits are assumed to involve greater complexity than standard clinic visits.

**Environment**

- Competition – Several studies have considered the possibility that competition affects a hospital’s performance. The HHI (Burgess & Wilson, 1998) and four-firm concentration ratio (Brown III, 2003) are commonly employed calculations to capture the amount of competition within a given market. The HHI is the sum of squared market shares for all of the hospitals in a given market, typically calculated using dispositions. Higher values represent greater market concentration. More highly sophisticated methods of evaluating competition have been developed, but they require information on patient zip codes (Wong, Zhan, & Mutter, 2005). Generally, the HHI is considered more informative than either simple hospital counts or four-firm concentration ratios. With all of these methods, a key issue is
determination of the relevant geographical market area. The data used in this analysis provides information necessary to calculate HHIs based on county, Metropolitan Statistical Area and Dartmouth Atlas Health Services Areas (“HSA”). Because HSAs were designed to be more relevant to the healthcare market, they will be used here.

- Market Share – Related to competition and market concentration is the market share each hospital has. A greater market share may allow a hospital to worry less about competition, thereby permitting higher levels of inefficiency. However, some research has found the opposite to be true (Register & Bruning, 1987).

Quality
Quality is an elusive concept to measure. While it seems important to include quality in modeling efficiency as discussed in Chapter 2, its effects on efficiency are unclear. Pauly (2004) pointed out that if a healthcare organization is not operating at the frontier (i.e. not technically efficient), there may be no requisite “trade-off” of cost and quality due to greater or lesser amounts of competition. Thus, unless a hospital is operating at the frontier, it is entirely possible that a hospital can improve quality without any increased costs. Likewise, increased costs may have no effect on quality. If these ambiguities are applicable to quality and cost, they might also be applicable to the relationship between quality and efficiency.

In healthcare, quality measures can be classified into three categories according to the Donabedian model (Donabedian, 1980)—structure, process, and outcome. Structural measures, such as services provided or accreditations earned, are easy to measure, but may not necessarily do a good job of explaining quality. Process measures are harder to quantify, but usually have a stronger link to outcomes than structural measures. Without direct observations, researchers must rely on surveys of patients, record
review, or in-depth searches of claims data. However, some processes have been scientifically proven to produce better outcomes. Finally, outcome measures are the most intellectually satisfying, representing the “bottom line”. Typically easy to observe, they are very valuable to patients. Length of stay, mortality, and rates of readmission are three standard outcome measures. Some scholars argue that un-captured severity of patient illness can confound outcome quality measures, especially since administrative data often serves as the only source of information (Romano & Mutter, 2004). However, improvements in capturing patient severity allow for calculating estimates of mortality and length of stay (HCUP Nationwide Inpatient Sample (NIS), 2006). Knowledge of expected measures allows for calculating deviation from expectations – an informative measure more likely to reflect quality. Inclusion of two measures from each category of structure, process, and outcome will consider quality from different perspectives.

- **Structure**: “Structural measures may be viewed as enabling or facilitating factors that make it easier or harder for health professionals to provide high-quality care” (Romano & Mutter, 2004). Whether or not a hospital receives *JCAHO accreditation* is an example of such a structural quality measure. Hospitals that undergo the accreditation process should function at higher levels of quality. *Teaching status* is another structural measure of quality that is commonly used in efficiency studies (Rosko & Mutter, 2008). Teaching hospitals are hypothesized to provide higher-quality care.

- **Process**: Some processes have been proven to produce higher quality results. These processes generally become codified in evidence-based medicine protocols. The Hospital Quality Alliance – a collaboration of key healthcare organizations, including CMS, AHA, and AMA JCAHO, and Blue Cross/Blue Shield – has developed measures of quality with the goal of improving patient care and providing public information for patients’ use (Hospital Quality Alliance).
Reporting is voluntary, but CMS pays hospitals that do report results an incentive payment. Over 4,000 acute care general hospitals currently submit results. The process measures included in the overall set of measures focus on AMI, heart failure, pneumonia, and surgery patients. Two of these processes, chosen based on the high response rate and greatest opportunity for variance, will be included:

- Process 1 – Percentage of surgery patients with recommended venous thromboembolism prophylaxis ordered.
- Process 2 – Percentage of pneumonia patients receiving initial antibiotic within 6 hours of hospital arrival.

(U.S. Department of Health and Human Services)

- **Outcome:** Length of stay and mortality have both been determined to be reliable, sensitive, and valid quality-related indicators of hospital performance (Griffith, Alexander, & Warden, 2002). Different mortality measures have been included in efficiency studies. Rosko and Mutter’s (2008) analysis examined twelve Inpatient and Safety Quality Indicators developed by the Agency for Healthcare Research and Quality as influencers of efficiency. These indicators were primarily outcome measures, such as risk-adjusted mortality rates for specific conditions and risk-adjusted rates of specific complications. However, comparisons of mortality rates do not account for the severity of the patients’ conditions: a higher mortality rate would be expected for hospitals treating the most complex conditions. Data in this analysis allows for calculation of *expected versus actual mortality rate* and *expected versus actual length of stay* as two outcome indicators of quality that take into account patient heterogeneity.
Individual Characteristics

- **Physicians** – It is difficult to control for all characteristics of individual providers in large national studies using administrative data. However, data is available that allows exploration into two “organizational schemes” that may directly affect physicians’ practice behavior – malpractice and terms of employment.
  
  o Malpractice – Previous research has explored the idea that the malpractice environment a physician faces influences her methods of providing care. Some research has found evidence that a high-risk malpractice environment encourages the practice of “defensive medicine” – performing procedures and ordering tests not for the patient’s health, but to safeguard against malpractice (U.S. Congress, Office of Technology Assessment, 1994). This is what Kessler and McClellan (1996) and (2002) found in analyses of acute myocardial infarction in the Medicare population. However, Kim (2007) found no such evidence of defensive medicine in research focusing on obstetrics, a field fraught with malpractice. This study used both the count and dollar value of claims within a state obtained from the National Practitioner Data Bank (“NPDB”) to define the malpractice environment a practitioner faces (Kim, 2007). However, count and value are likely correlated. Thus, this analysis will only use the count of settlements per capita in the state for the period 2000-2005 (six years’ prior to 2006) as provided by the NPDB.

  o Employment arrangement – Every hospital has different predominant physician arrangements. The American Hospital Association (“AHA”) Annual Survey lists eight types of arrangements, and asks hospitals to provide the count of physicians operating in their facility in each category. Yet counts of physicians by category do not adequately address total physician FTEs because a credentialed physician could perform either a handful or hundreds of surgeries in a given hospital. What can be definitively ascertained from the AHA survey is the number of hospitals that report zero physician
employees involved in patient care. Based on all hospitals in the 2006 AHA Annual Survey, thirty-one percent of hospitals overall report no physician employees (i.e. all credentialed physicians) providing patient care. At the other end of the spectrum, military hospitals have practically all physician-employees – whether Active Duty or civilian. Even contract providers (for the most part) are considered employees for instances of malpractice, although the lines between employee and independent contractor blur the issue of control (Shelley, 1998).

- Patients – It can also be difficult to control for the characteristics of individual patients, especially when using administrative databases. Zuckerman, Hadley, and Iezzoni (1994) used patient characteristics derived from Medicare data, but found the benefits this data provided to the model did not outweigh the computational costs. However, both military and civilian sources of data used for this dissertation and detailed in the next chapter provide patient demographics for each disposition allowing aggregation of these characteristics at the hospital level for inclusion in the analysis. Furthermore, the assertion that military patients are too different from civilian patients begs for exploration of patient characteristics as influencers of efficiency. Patient demographics and characteristics to be investigated as independent variables that might influence efficiency are:
  
  o **Average Patient Age** – It would be hard to disagree with the assertion that differences in patient base – especially age – affect care provided. It may be that this difference in care influences efficiency. DRG weighting is typically based on only two or three age groupings and may not adequately capture the effects of age differences.
  
  o **Percent of Female Patients** – Female patients differ from their male counterparts in the medical care they require, and again, this difference could affect efficiency.
o **Percent of Non-White Patients** – Disparities in health care due to race are frequently studied, and any systematic differences could be reflected in efficiency.

o **Percentage of Uninsured/Self Pay Patients** – Several studies have examined hospital efficiency by allocating patients based on type of insurance coverage (Freisner, Roseman, & McPherson, 2008). Military patients are almost all insured through Tricare, while civilian hospitals treat varying percentages of uninsured and self-pay patients. Hospitals may function differently based on the type of insurance their patients carry. Furthermore, patients within these categories may behave in ways that influence efficiency. Certainly, patients with insurance are likely to have fewer physiological effects from stress over financial concerns.

**Proposed Analytical Model**
The previous discussion presented the variables to be explored in this efficiency analysis – output, inputs, and variables that may influence efficiency – and support for their use in the literature. Employing these variables and the general empirical model presented at the beginning of this chapter, a more refined model of hospital healthcare and the related efficiency is presented below:

\[
\text{In- & Outpatient Workload} = f(\text{Beds, RN FTEs, Other Non-MD FTEs}); \text{ and}\\
\text{Efficiency} = g(\text{OWNERSHIP, STRUCTURE, ENVIRONMENT, QUALITY, MD, PATIENT}) \text{ where}\\
\]

- STRUCTURE is a vector consisting of variables for: size, case mix index, percentage of surgical work, scope of MDCs, and percentage of ER visits.
- ENVIRONMENT is a vector consisting of variables for: level of competition and share of the defined market held by the hospital.
• QUALITY is a vector consisting of variables for: JCAHO accreditation, teaching status, percentage of surgery patients with recommended venous thromboembolism prophylaxis ordered, percentage of pneumonia patients receiving an initial antibiotic within 6 hours of hospital arrival, expected vs. actual length of stay, and expected vs. actual mortality rate.

• MD is a vector consisting of variables for: 100% credentialed physicians and count of malpractice claims for the previous six years by state.

• PATIENT is a vector consisting of variables for: average patient age, percentage of female patients, percentage of non-white patients, and percentage of uninsured/self-pay patients.
CHAPTER 6 - DATA

This chapter discusses the sources of data used in this comparison of military hospital technical efficiency to other forms of control and the process of consolidating this data into one set. A discussion of variable creation and descriptive statistics for these variables follow.

Data Sources
Performing this comparative analysis required merging data from several data sources – some civilian and some military.

Civilian Data
As is the case with most efficiency analyses of civilian hospitals, the American Hospital Association ("AHA") Annual Survey of Hospitals is a primary source of data. Conducted since 1946, it is a comprehensive survey of all U.S. hospitals (6,349 in 2006) – including military facilities – meeting ten very basic requirements. Some of these requirements include a minimum of six inpatient beds, an average stay of more than 24 hours, an identifiable governing authority, and an organized medical staff. The survey covers the organizational structure, services offered, utilization, financial, and personnel information of each hospital. The survey’s response rate is typically around 85%. When a hospital fails to respond in full or in part, either regression using past values for that hospital or estimation based on hospitals in the same strata is performed to predict key variables. Facility descriptors and services are not estimated: they come from the AHA master facility inventory (American Hospital Association, 2006). While the AHA Annual Survey is recognized as a legitimate source of information on hospitals, its self-reported nature brings with it the usual concerns regarding accuracy.

In order to obtain more accurate and detailed inpatient workload data, the 2006 Healthcare Cost and Utilization Project ("HCUP") 2006 Nationwide Inpatient Sample
(“NIS”) was also a principal data source. The NIS represents a 20% stratified sample of U.S. hospitals and covers thirty-eight states. It is the largest all-payer inpatient care database in the U.S. The 2006 NIS contains data on over eight million hospital discharges from 1,045 hospitals for the 2006 calendar year (HCUP Nationwide Inpatient Sample (NIS), 2006). Use of the NIS allowed for calculation of an all-payer case mix index rather than the standard Medicare case mix index provided by CMS. It also allowed for computation of facility averages for patient characteristics, length of stay, and deaths. In most – but not all – cases, data can be linked to more detailed facility data in the AHA Annual Survey. Twelve of the thirty-eight states have restrictions that do not permit this linkage (HCUP Nationwide Inpatient Sample (NIS), 2006). Therefore, the civilian hospitals used in the final dataset were limited to twenty-six states.

Military Data
Military data for calendar year 2006 was obtained via a Data Use Agreement with TRICARE Management Activity, Office of the Secretary of Defense for Health Affairs, Department of Defense through the Naval Medical Education and Training Command (“NMETC”). Data was extracted from the Military Health System Management Analysis and Reporting Tool (“M2”) – an ad-hoc data querying system designed to support operations, decision-making, and oversight of the MHS – in aggregated form to protect patient privacy. The M2 receives data extracts from the Military Health System Data Repository, a data warehouse that collects and stores data from multiple primary systems covering MTF workload, patient/beneficiary information, and healthcare services provided. The Medical Expense and Performance Reporting System (“MEPRS”), Standard Ambulatory Data Record (“SADR”), and Standard Inpatient Data Record (“SIDR”) are the primary data systems used for this analysis. Data on inpatient discharges came from the SIDR; data on outpatient encounters came from the SADR; and data on personnel full-time equivalents came from the MEPRS. The SIDR and SADR files received were aggregated based on combinations of patient and care episode descriptor fields by hospital by month. The MEPRS file was aggregated based on unique
combinations of personnel type and time type fields by hospital by month, rather than one record per disposition (visit).

Secondary Data Source
In addition to the data sources described above, a fourth source of data was used to create an independent variable associated with physician malpractice risk. Data on medical malpractice cases was extracted from the National Practitioner Data Bank (“NPDB”) Public Use Data File. The NPDB contains data on malpractice payouts nationwide. Insurance carriers are required by law to submit a report on all payouts. Updated quarterly, it is a cumulative database of malpractice settlements since September, 1990.

For this analysis, year of claim origination, state in which the practitioner was licensed, and total settlement amount were used. Only settlements pertaining to physicians, osteopaths, interns, or residents were used. Because this analysis focuses on 2006, only claims prior to 2006 were used: the six-year period from 2000 to 2006 was chosen. Settlement amounts in the NPDB are recorded in ranges to protect confidentiality, and payments are top-coded at $105 million. Thus, each payment is recorded at the range midpoint. This comprehensive database is not without limitations. Only closed cases with a positive payout are included, yet it is possible that cases with zero payouts entail non-monetary costs that are non-trivial. Data in the NPDB is only provided at the state level, and therefore within-state variation cannot be studied. The NPDB cannot link multiple defendants for a single case when they are reported separately. Settlements involving only a hospital (and not a practitioner) are not included, and settlements involving a hospital and a practitioner only list the practitioner settlement. “Despite these limitations, it is the most accurate source of information for the entire United States over a long period regarding physician malpractice risk” (Kim, 2007).
Steps to create final dataset

1. Data from the 2006 HCUP NIS was converted into Stata-compatible format: Stata/SE 9.2 is the statistical program used for this analysis.

2. Discharge records from hospitals in states that do not allow for linkage to the 2006 AHA Annual Survey database were eliminated, leaving 666 remaining hospitals and 5,659,236 remaining dispositions. The states eliminated were Georgia, Hawaii, Indiana, Kansas, Michigan, Nebraska, Ohio, Oklahoma, South Carolina, South Dakota, Tennessee, and Texas. The remaining states for which linkage between disposition-level data and hospital-level data was possible were: Arkansas, Arizona, California, Colorado, Connecticut, Florida, Iowa, Illinois, Kentucky, Louisiana, Massachusetts, Maryland, Minnesota, Missouri, North Carolina, New Hampshire, New Jersey, Nevada, New York, Oregon, Rhode Island, Utah, Vermont, Washington, Wisconsin, and West Virginia.

3. A case mix index was created by applying DRG weights from CMS DRG Grouper Version 23 for dispositions through September and CMS DRG Grouper Version 24 for dispositions from October to December to each record. In each of these groupers, there are more than 500 DRGs.

4. Variables for hospital statistics were created for: 1) total dispositions, 2) case mix index (average, median, standard deviation), 3) average patient age, 4) percent of female patients, 5) percent of non-Caucasian patients, 6) percent of self-and no-pay patients, 7) percentage of actual over expected mortality rate, 8) percentage of actual over expected length of stay, 9) percent of surgical DRGs, and 10) percent of workload assigned to each of 25 Major Diagnostic Categories.

5. The disposition-level records were then collapsed to the facility level for these statistics, and stored in a separate Stata file.

6. To abstract military data, the SIDR military database – received as a Microsoft Access file – was converted to Stata format. Visual inspection of the data revealed that three hospitals reported less than twelve months of data, and one hospital reported significantly fewer dispositions than estimated on the 2006 AHA Annual Survey.
Further research revealed that the three hospitals with less than a year of data—Naval Hospital Great Lakes, Fort Eustis, and Scott Air Force Base—were converted to ambulatory clinics as part of the 2005 Base Realignment and Closure Act. These conversions occurred during 2006. The hospital with significantly reduced dispositions—Kessler Air Force Base—suffered significant damage from Hurricane Katrina that degraded capabilities for the majority of 2006. Additionally, dispositions occurring in civilian institutions because of resource-sharing agreements that were attributed to the parent military facility were excluded. Excluding these civilian dispositions eliminated Naval Hospital Charleston because all of its inpatient care is performed in civilian hospitals. Adjusting for these cases left 230,998 dispositions from 44 military hospitals.

7. Variables from Step 3 above were created for military hospitals, and the records were collapsed to the facility level and saved in a separate Stata file. Two separate versions were created to investigate potential differences due to use of either CMS or TMA DRG weight schedules for military hospitals. This investigation revealed that the effects of these different weights on the final model were insignificant, so TMA weights were chosen to capture more accurately the true nature of military hospital workload.

8. Data on FTEs from MEPRS was converted to Stata format from Microsoft Access. Annual totals were created for each hospital, and a separate file saved. Data on outpatient visits from the SADR was also converted to Stata format from Microsoft Access. Totals were created for each hospital, and a separate file saved. Both of these files were merged with the inpatient file created in Step 7 above to create one military file.

9. The civilian and military facility-level files were both merged with the AHA 2006 Annual Survey Database by AHA Identification Number. Hospitals not classified as general medical/surgical facilities and hospitals classified as long-term were deleted. All records in the AHA Survey Database without relevant matching disposition-level data from the mergers were deleted.
10. Statewide data on malpractice settlements abstracted from the NPDB was then merged with the current working file. Military facilities received values of zero to reflect the fact that military physicians do not face the personal financial risk associated with malpractice claims.

11. Data on civilian process quality measures from the Hospital Compare website (U.S. Department of Health and Human Services), and military process quality measure from the MHS-QCM website (Tricare Management Activity) was merged with the data set.

12. HHI and market share variables were created by HSA using data from the AHA Annual Survey. Hospitals with no HSA code were deleted.

13. A binary variable capturing differences in physician-hospital agreements was created based on AHA Annual Survey data. Hospitals with zero physician employees were coded “1”, and those with any physician employees were coded “0”.

14. Case Mix-adjusted dispositions were calculated by multiplying hospital case mix indices created in steps 3 and 6 above by total dispositions.

15. Because the log of zero is not defined, and the production function – whether Cobb-Douglas or translog – utilizes logged input values, three hospitals with zero values in hospital beds, dispositions, nurse FTEs, or other FTEs, were excluded.

The final data set contains 668 hospitals (624 civilian and 44 military).

Variable Definitions
This section defines the variables used for the analysis conducted in Chapter 8.

OUTPUT VARIABLES

TOTAL_DI: A continuous variable, total hospital dispositions come from NIS data for civilian hospitals. It is the total number of records for each facility. For the military, the
variable representing total dispositions is the sum of \textit{dispositions\_raw} from the SIDR\(^4\). No exclusions for transfers or deaths were made.

**VTOT**: A continuous variable, total outpatient visits for all facilities were provided in the AHA Survey. However, military data is available and allows for a more accurate count since it comes directly from the data source. Thus, for military hospitals, the variable representing total outpatient visits is the sum of \textit{encounters} from the SADR for all records.

**TOTWORK**: As discussed in the previous chapter, this continuous variable is an aggregation of inpatient and outpatient work based on the ratio of average Medicare outpatient visit cost to average Medicare inpatient disposition cost as follows:

\[
\text{TOTWORK} = \text{TOTAL\_DI} + (\text{VTOT} \times 0.012)
\]

Different model specifications will apply the case mix index to this variable in different ways to consider heterogeneity of work performed.

**INPUT VARIABLES**

**HOSPBD**: The number of hospital beds is a continuous variable taken from the 2006 AHA Annual Survey for all hospitals. This survey item is never estimated. If not provided by the hospital, the value reported in the survey is taken from the independently-maintained AHA master facility inventory.

**FTERN**: RN Full-time Equivalents is a continuous variable taken from the 2006 AHA Annual Survey for civilian hospitals and from total \textit{available ftes} in personnel category 2R in the MEPRS file for military hospitals, excluding RN FTEs assigned to military-unique (MEPRS codes F & G) tasks.

\(^4\) This variable was chosen instead of another variable in the SIDR that calculates total dispositions by crediting facilities for workload related to patients still in the facility at year-end.
**FTEOTH**: Other Full-time Equivalents is a continuous variable calculated using items from the AHA Survey for civilian hospitals and from the MEPRS file for military hospitals, excluding FTEs assigned to military-unique tasks. It is calculated as follows:

\[
\text{Other FTEs} = \text{Total hospital FTEs} - \text{RN FTEs} - \text{MD/Dentist FTEs}
\]

- Civilian: FTEH-FTERN-FTEMD (Variables from AHA Survey)
- Military: \([\left(\sum \text{Available FTEs for MEPRS codes A-D}\right) - (\sum \text{FTEs for skill type } 3R \text{ (RNS)}) - (\sum \text{FTEs for skill type } 1P \text{ and } 1D \text{ (MDs and Dentists)})] \times (\text{excludes FTEs attributed to military-unique tasks MEPRS Codes F and G})\) (Variables from MEPRS)

**IDENTIFICATION VARIABLE**

**AHAID**: This seven-digit identifier code is the American Hospital Association ID. It is taken directly from the 2006 AHA Annual Survey for all hospitals.

**EFFICIENCY VARIABLES**

**CONTROLCAT**: This categorical variable captures the four different ownership types. It aggregates values from the AHA Annual Survey item \(cntrl\) as follows:

- CNTRL codes less than 20 = CONTROLCAT 1 – State/Local government
- CNTRL codes ≥ 20 and <30 = CONTROLCAT 2 – Not-For-Profit
- CNTRL codes ≥ 30 and <40 = CONTROLCAT 3 – For Profit
- CNTRL codes 41-43 = CONTROLCAT 4 – Military

**SMALLHOSP**: Based on hospital beds, this binary variable is coded “1” if hospital beds are <50 (bed size categories (bsc) 1 and 2 from the AHA Survey).
**JCAHO:** Accreditation by the Joint Commission for the Accreditation of Healthcare Organizations is indicated by this binary variable. The AHA survey item *mapp1* was re-coded so that “1” represents accreditation and “0” represents no accreditation.

**TEACH:** Membership in the Council of Teaching Hospitals is indicated by this binary variable. The AHA survey item *mapp8* was re-coded so that “1” represents members and “0” represents non-members.

**CRED:** This is a binary variable with “1” representing hospitals with 100% Credentialied MDs (0 Employees) providing care, based on AHA survey item *ftemd*, and “0” otherwise. Military hospitals are all coded “0” since all physicians are considered employees.

**SHAREHSA:** Based on AHA data for all hospitals, this continuous variable represents each hospital’s market share based on its share of total admissions in the Hospital Service Area, as defined by the Dartmouth Index.

**HSAHHI:** Based on AHA data for all hospitals, this continuous variable is the Herfindahl-Hirschman Index calculated based on Dartmouth Atlas-defined Health Service Areas (“HSAs”) and total admissions as follows: 

\[ HHI = \sum_{i=1}^{N} SHAREHSA_i^2 \]  

where \( N \) = number of hospitals in a given HSA.

**CLAIMSPERCAP2000:** Based on National Practitioner Databank information, this continuous variable represents the total number of medical malpractice claims per capita (in thousands) for physicians in a given state from 2000 – 2005. Values for all military facilities, although included in the NPDB as state “AA” or “AE”, were re-coded to “0” to reflect the relative protection of military physicians from state malpractice laws.

**ERPERCENT:** This continuous variable represents the percentage of outpatient visits occurring in the Emergency Room.

- **Civilian:** ER visits \( \text{vem} \)/Total Outpatient Visits \( \text{vtot} \) (Source: AHA Survey)
- **Military:** encounters coded *product_line* “ER”/\( \Sigma \) encounters (Source: SADR)
**PNEUMONIA**: This continuous variable represents the percentage of pneumonia patients receiving their first antibiotic within 6 hours of hospital arrival. Civilian figures were taken from data available at the Hospital Compare website (U.S. Department of Health and Human Services), and military figures were taken from data available at the MHS-QCM (Tricare Management Activity).

**SURGERY**: Percent of surgical patients receiving the recommended venous thromboembolism prophylaxis ordered. Civilian figures were taken from data available at the Hospital Compare website (U.S. Department of Health and Human Services), and military figures were taken from data available at the MHS-QCM (Tricare Management Activity).

Each of the following continuous variables was created by aggregating disposition (patient)-level data from the NIS for civilian hospitals and from the SIDR for military hospitals. Each variable represents a hospital-level calculation.

**HOSPAVGAGE**: Hospital Average Patient Age for all Dispositions.

Civilian and Military: \( \frac{\text{Patient Age}}{\text{TOTAL_Di}} \)

**PCTFEMALE**: Percentage of Female Patients.

Civilian:

\[
\sum \text{Dispositions where female} = 1 \\
\text{TOTAL_DI}
\]

Military:

\[
\sum \text{Dispositions where gender} = F \\
\sum \text{dispositions_raw}
\]
**PCTNONWHITE**: Percentage of Non-white Patients.

Civilian:
\[
\frac{\sum \text{Dispositions where race } \neq 1}{\text{TOTAL\_DI}}
\]

Military:
\[
\frac{\sum \text{Dispositions where race } \neq C}{\sum \text{dispositions\_raw}}
\]

**PCTSELFNOPAY**: Percent of Self or No-Pay Dispositions

Civilian:
\[
\frac{\sum \text{Dispositions where PAY1 = 4 or 5}}{\text{TOTAL\_DI}}
\]

Military: 0

**PCTSURG**: Percent of Dispositions identified as being surgical in nature based on DRG

Civilian:
\[
\frac{\sum \text{Dispositions coded SURG = 1}}{\text{TOTAL\_DI}}
\]

Military:
\[
\frac{\sum \text{Dispositions coded S in medical\_surgical\_indicator}}{\sum \text{dispositions\_raw}}
\]

**PCTOVERMORT**: Percentage that Actual Mortality Rate Differs from Expected Mortality Rate. The NIS provides Ingenix All Payer Severity DRG scales that allow for estimation of expected mortality rate for a given disposition based on patient demographics, diagnoses, and procedures performed (HCUP Nationwide Inpatient Sample (NIS), 2006). The SIDR provides a similar estimate for military dispositions. Calculating the percentage that the overall actual mortality rate is above (or below) the expected overall mortality rate measures how well a hospital performs compared to expectations and is a better measure of quality than a direct comparison of actual mortality rates.
Civilian: Average of All-payer Severity DRG * Mean Mortality – Actual Mortality

$$\sum_{\text{DIED}} = 1 - \frac{\sum \text{APSDRG Mortality Weight} \times .02131}{\sum \text{APSDRG Mortality Weight} \times .02131}$$

Military: Average of Civilian Norm Death Rate - Actual Mortality

$$\sum \frac{\text{disposition code} = 20 - \sum \text{deaths civ norm raw}}{\sum \text{dispositions raw}} - \frac{\sum \text{deaths civ norm raw}}{\sum \text{dispositions raw}}$$

**PCTOVERLOS:** Percentage that Actual Length of Stay Differs from Expected Length of Stay. The NIS provides Ingenix APS DRG scales that allow for estimation of expected length of stay for a given disposition based on patient demographics, diagnoses, and procedures performed (HCUP Nationwide Inpatient Sample (NIS), 2006). The SIDR provides a similar estimate for military dispositions. Calculating the percentage that the actual length of stay is above (or below) the expected length of stay measures how well a hospital performs compared to expectations and is a better measure of quality than a direct comparison of actual lengths of stay.

Civilian: Average of All-payer Severity & Mean Length of Stay

$$\frac{\left( \sum \frac{\text{LOS}}{\text{TOTAL DI}} - \frac{\sum \text{APSDRG LOS Weight} \times 4.5427}{\sum \text{APSDRG LOS Weight} \times 4.5427} \right)}{\sum \text{APSDRG LOS Weight} \times 4.5427}$$

Military: Average of Civilian Norm Length of Stay

$$\frac{\left( \sum \frac{\text{bed days raw}}{\text{dispositions raw}} - \frac{\sum \text{bed days civ norm raw}}{\text{dispositions raw}} \right)}{\sum \text{bed days civ norm raw}}$$
HOSP: Case Mix Index. For both military and civilian hospitals, weights for each disposition’s DRG from either CMS or TMA versions of DRG Groupers 23 and 24 (depending on the month in which the disposition occurred) were averaged.

CMS publishes overall facility “Case Mix Indexes” for hospitals treating Medicare beneficiaries. However, this index is based on the DRGs for the dispositions of only Medicare patients, and therefore facility average DRG weights calculated for all patients based on either NIS or military disposition data are preferable. Tricare Management Activity “TMA” follows CMS Prospective Payment System procedures to the greatest extent practicable, and the vast majority of TMA DRGs are identical to CMS DRGs. The most significant differences in TMA and CMS DRGs are the 34 additional codes that TMA uses for neonates, since Medicare’s requirement for neonatal codes is negligible (Tricare Management Activity, 2008). However, the relative weights of each DRG differ: TMA weights more accurately reflect the characteristics of the patient base of military facilities (i.e. typically younger and healthier) (Tricare Management Activity, 2002). Both TMA and CMS relative weights were applied to military dispositions in separate models, with non-significant differences in efficiency between the two. Thus, TMA weights are used in the remaining analysis. Finally, the set of DRG weights for both CMS and TMA changed on October 1, 2006. The appropriate set of weights was used based on date of disposition to address this.

Civilian:

\[
\sum_{DISPO} \frac{CMS \ DRG \ Relative \ Weights}{TOTAL\ DI}
\]

Military:

\[
\sum_{DISPO} \frac{TMA \ DRG \ Relative \ Weights}{\sum dispositions_{raw}}
\]
HOSPMEDCMI and HOSPSDCMI: It is possible that the shape of the distribution of HOSPCMI affects efficiency. The average case mix index is likely to be positively skewed due to a zero lower bound and the possibility that outliers of extreme severity will stretch the upper tail of the distribution. Furthermore, variability in a hospital’s case mix index could affect efficiency: greater variability may correlate with higher inefficiency. Therefore, the effects of using median rather than average CMI and inclusion of the standard deviation of the CMI will be considered.

PCTMDC (1 – 25): For each of the 25 Major Diagnostic Categories, a variable was created for the percentage of total dispositions assigned to DRGs in that category.

Civilian:
\[
\frac{\sum_{n=1}^{25} MDC_n}{TOTAL\_DI}
\]

Military:
\[
\frac{\sum_{n=1}^{25} mdc_n}{\sum \text{dispositions}_\text{raw}}
\]

As mentioned previously, the MDCs where military hospitals differ the most from civilian ones are MDC 14 (childbirth) and MDC 5 (circulatory). These two variables are included in the model, but all MDCs are explored as descriptive statistics.
Final Data Set Variable List

The final consolidated file contains the variables listed below.

Table 6.1: Final Data Set Variables

<table>
<thead>
<tr>
<th>NAME</th>
<th>DEFINITION</th>
<th>TYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHAID</td>
<td>AHA ID Number</td>
<td>String</td>
</tr>
<tr>
<td>CONTROLCAT</td>
<td>Ownership Category</td>
<td>Categorical</td>
</tr>
<tr>
<td>TOTAL_DI</td>
<td>Total Inpatient Dispositions</td>
<td>Continuous</td>
</tr>
<tr>
<td>VTOT</td>
<td>Total Outpatient Visits</td>
<td>Continuous</td>
</tr>
<tr>
<td>TOTWORK</td>
<td>Dispositions + Visits*0.12</td>
<td>Continuous</td>
</tr>
<tr>
<td>HOSPBD</td>
<td>Total Non-Nursing Home Beds</td>
<td>Continuous</td>
</tr>
<tr>
<td>FTEOTH</td>
<td>Other Non-MD FTEs</td>
<td>Continuous</td>
</tr>
<tr>
<td>HOSPNAVAGE</td>
<td>Average Patient Age</td>
<td>Continuous</td>
</tr>
<tr>
<td>PCTFEMALE</td>
<td>Percent of Female Patients</td>
<td>Continuous</td>
</tr>
<tr>
<td>PCTNONWHITE</td>
<td>Percent of Non-White Patients</td>
<td>Continuous</td>
</tr>
<tr>
<td>PCTSELFNOPAY</td>
<td>Percent of Self and No Pay Patients</td>
<td>Continuous</td>
</tr>
<tr>
<td>PCTSURG</td>
<td>Percent Surgical DRGs</td>
<td>Continuous</td>
</tr>
<tr>
<td>ERPERCENT</td>
<td>Percent of ER Outpatient Visits</td>
<td>Continuous</td>
</tr>
<tr>
<td>HOSPCMI</td>
<td>Case Mix Index Average</td>
<td>Continuous</td>
</tr>
<tr>
<td>HOSMEDCMI</td>
<td>Case Mix Index Median</td>
<td>Continuous</td>
</tr>
<tr>
<td>HOSPSDCMI</td>
<td>Case Mix Index Standard Deviation</td>
<td>Continuous</td>
</tr>
<tr>
<td>PCTMDCXX</td>
<td>Percent of Dispositions MDCs 1-25</td>
<td>Continuous</td>
</tr>
<tr>
<td>CLAIMSPERCAP2000</td>
<td>Malpractice Claims per cap 2000-2005</td>
<td>Continuous</td>
</tr>
<tr>
<td>HSAHII</td>
<td>HSA Herfindahl-Hirschman Index</td>
<td>Continuous</td>
</tr>
<tr>
<td>SHAREHSA</td>
<td>Hospital HSA Market Share</td>
<td>Continuous</td>
</tr>
<tr>
<td>PNEUMONIA</td>
<td>Pneumonia Process</td>
<td>Continuous</td>
</tr>
<tr>
<td>SURGERY</td>
<td>Surgery Process</td>
<td>Continuous</td>
</tr>
<tr>
<td>PCTOVERMORT</td>
<td>% Actual Over Expected Mortality</td>
<td>Continuous</td>
</tr>
<tr>
<td>PCTOVERLOS</td>
<td>% Actual Over Expected LOS</td>
<td>Continuous</td>
</tr>
<tr>
<td>SMALLHOSP</td>
<td>Beds &lt;50 (Bed Size Category 1 &amp; 2)</td>
<td>Binary</td>
</tr>
<tr>
<td>JCAHO</td>
<td>JCAHO Accreditation (Y/N)</td>
<td>Binary</td>
</tr>
<tr>
<td>TEACH</td>
<td>Council of Teaching Hospital Member</td>
<td>Binary</td>
</tr>
<tr>
<td>CRED</td>
<td>100% Credentialied Physicians (Y/N)</td>
<td>Binary</td>
</tr>
</tbody>
</table>
Descriptive Statistics

In 2006, there were 4,750 short-term general medical/surgical hospitals within the fifty states, categorized by ownership type as follows:

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>State/Local Government</td>
<td>1,111</td>
<td>23.4%</td>
</tr>
<tr>
<td>Not-For-Profit</td>
<td>2,738</td>
<td>57.6%</td>
</tr>
<tr>
<td>For-Profit</td>
<td>705</td>
<td>14.8%</td>
</tr>
<tr>
<td>Military</td>
<td>49</td>
<td>1.0%</td>
</tr>
<tr>
<td>Other Federal</td>
<td>147</td>
<td>3.1%</td>
</tr>
</tbody>
</table>

After following the steps described earlier to create the final data set, the final sample for this analysis contains 668 short-term general medical/surgical hospitals within the fifty states, broken down below. Other Federal hospitals are excluded because the HCUP NIS does not include them in their sample, and no other means of accessing disposition-level data are available. Furthermore, they are not the focus of this dissertation. The proportion of military hospitals in the sample is naturally higher than in the total population due to sample design. The proportion of state/local government and for-profits is slightly lower than in the total population, while the proportion of not-for-profit hospitals is slightly higher than in the total population. Overall, however, the sample is comparable to the population.

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>State/Local Government</td>
<td>126</td>
<td>18.9%</td>
</tr>
<tr>
<td>Not-For-Profit</td>
<td>413</td>
<td>61.8%</td>
</tr>
<tr>
<td>For-Profit</td>
<td>85</td>
<td>12.7%</td>
</tr>
<tr>
<td>Military</td>
<td>44</td>
<td>6.6%</td>
</tr>
</tbody>
</table>

Table 6.2 below provides descriptive statistics on output, beds, and FTEs for the overall population of 4,750 hospitals from the 2006 AHA Annual Survey:
Table 6.2: General U.S. Short-Term Medical/Surgical Hospital Raw Data

<table>
<thead>
<tr>
<th>Mean (SD)</th>
<th>Military (N=49)</th>
<th>Non-Military (N=4,701)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inpatient Admissions</td>
<td>4,310.08 (3,604.10)</td>
<td>7,446.83 (9,144.92)</td>
</tr>
<tr>
<td>Outpatient Visits (000)</td>
<td>420.92 (239.71)</td>
<td>133.54 (193.23)</td>
</tr>
<tr>
<td>Hospital Beds</td>
<td>98.11 (79.07)</td>
<td>170.06 (186.27)</td>
</tr>
<tr>
<td>MD FTEs</td>
<td>174.00 (160.42)</td>
<td>19.15 (74.45)</td>
</tr>
<tr>
<td>RN FTEs</td>
<td>236.92 (232.24)</td>
<td>238.81 (337.47)</td>
</tr>
<tr>
<td>LPN FTEs</td>
<td>69.33 (43.90)</td>
<td>24.96 (30.34)</td>
</tr>
<tr>
<td>Other FTEs (incl LPNs)</td>
<td>958.04 (929.02)</td>
<td>509.15 (706.07)</td>
</tr>
</tbody>
</table>

Admissions (rather than dispositions) are reported in the AHA Annual Survey and are thus enumerated above. Military hospitals, on average, perform less inpatient work (4,311 vs. 7,447 admissions), but significantly more outpatient work than civilian hospitals (420.1K vs. 133.5K visits), and they are considerably smaller (98 vs. 170 beds). Military hospitals appear to have many more MDs than civilian ones (174 vs. 19 FTEs), but this is in part due to differences in business models: all MDs are considered employees in military hospitals, while civilian hospitals merely credential physicians to work in their facility. While the number of RN FTEs is similar (237 vs. 239 FTEs), all other labor categories are also significantly higher for military hospitals (958 vs. 509 FTEs). It is possible that administrative positions not directly associated with the hospital such as dental, environmental and occupational health personnel (Ozcan & Bannick, 1994), as well as positions dedicated to operational tasks have been included in the military figures. Data from MEPRS allows adjustment for effort spent on military-unique tasks that would not occur in civilian hospitals so that comparisons are more relevant.
Table 6.3 below provides descriptive statistics for the entire sample:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispositions</td>
<td>8,676.58</td>
<td>10,687.59</td>
<td>18</td>
<td>115,262</td>
</tr>
<tr>
<td>Outpatient Visits</td>
<td>167,458.30</td>
<td>233,841.10</td>
<td>2</td>
<td>2,166,327</td>
</tr>
<tr>
<td>Beds</td>
<td>173.50</td>
<td>191.33</td>
<td>6</td>
<td>1,834</td>
</tr>
<tr>
<td>RN FTEs</td>
<td>252.88</td>
<td>334.55</td>
<td>6</td>
<td>3,195</td>
</tr>
<tr>
<td>Other FTEs</td>
<td>739.66</td>
<td>953.43</td>
<td>7</td>
<td>10,573</td>
</tr>
<tr>
<td>Case Mix Index – Avg</td>
<td>1.03</td>
<td>.24</td>
<td>.45</td>
<td>2.36</td>
</tr>
<tr>
<td>Case Mix Index – Med</td>
<td>.85</td>
<td>.14</td>
<td>.39</td>
<td>1.91</td>
</tr>
<tr>
<td>Case Mix Index- SD</td>
<td>.85</td>
<td>.41</td>
<td>.09</td>
<td>2.59</td>
</tr>
<tr>
<td>% Surgery DRGs</td>
<td>19.9%</td>
<td>12.0%</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>% ER Outpatient</td>
<td>24.6%</td>
<td>17.0%</td>
<td>0</td>
<td>96.9%</td>
</tr>
<tr>
<td>% MDC 14 Work</td>
<td>10.6%</td>
<td>8.7%</td>
<td>0.0%</td>
<td>43.4%</td>
</tr>
<tr>
<td>% MDC 5 Work</td>
<td>14.7%</td>
<td>7.0%</td>
<td>0.0%</td>
<td>87.6%</td>
</tr>
<tr>
<td>&lt;50 Beds</td>
<td>.2764</td>
<td>.4475</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Average Age(yrs)</td>
<td>50.79</td>
<td>11.72</td>
<td>15.36</td>
<td>78.27</td>
</tr>
<tr>
<td>% Female</td>
<td>58.5%</td>
<td>5.2%</td>
<td>28.4%</td>
<td>72.7%</td>
</tr>
<tr>
<td>% Self/No Pay</td>
<td>5.1%</td>
<td>5.7%</td>
<td>0.0%</td>
<td>65.2%</td>
</tr>
<tr>
<td>% Non-White</td>
<td>18.4%</td>
<td>23.4%</td>
<td>0.0%</td>
<td>99.8%</td>
</tr>
<tr>
<td>JCAH0 Accreditation</td>
<td>74.6%</td>
<td>43.57</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>COTH Member</td>
<td>7.43%</td>
<td>26.24</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>% Over Exp Mortality</td>
<td>-1.04%</td>
<td>79.3%</td>
<td>-100%</td>
<td>1407%</td>
</tr>
<tr>
<td>% Over Exp LOS</td>
<td>-.59%</td>
<td>43.6%</td>
<td>-59.1%</td>
<td>786%</td>
</tr>
<tr>
<td>Pneumonia Process</td>
<td>93.15%</td>
<td>95%</td>
<td>49%</td>
<td>100%</td>
</tr>
<tr>
<td>Surgery Process</td>
<td>82.34%</td>
<td>86%</td>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>&lt;50 Beds</td>
<td>27.6%</td>
<td>44.8%</td>
<td>0.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Malpr Claims per cap (000s)</td>
<td>.355</td>
<td>.173</td>
<td>0</td>
<td>.704</td>
</tr>
<tr>
<td>100% Credentialed MDs</td>
<td>23.9%</td>
<td>42.7%</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HSA Market Share</td>
<td>66.8%</td>
<td>40.0%</td>
<td>14.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>HSA HHI</td>
<td>7,105</td>
<td>3,458</td>
<td>421</td>
<td>10,000</td>
</tr>
</tbody>
</table>
The average hospital performed almost 8,700 inpatient dispositions and just under 167.5K outpatient visits in 2006. With respect to inputs, the average hospital had 173.5 beds. The average hospital utilized 252.9 RN FTEs and 739.7 Other FTEs.

**Environment:** The mean HHI based on Dartmouth Atlas Health Service Areas reflects the typically high concentration characteristic of hospital markets. HHI values over 1,800 represent highly concentrated markets. The average HHI in this sample is 7,105, with a minimum of 421 and a maximum of 10,000. The average market share was 66.8%.

**Structure:** 27.6% of the sample hospitals had less than 50 beds. The average hospital’s case mix index was 1.03, with a median value of .85 and a standard deviation of .85. A skew test revealed that this positive skew was significant at the p<.001 level. 10.6% of the average hospital’s inpatient workload was related to pregnancy and childbirth (MDC 14). 14.7% of the average hospital’s workload was related to the circulatory system (MDC 5). Slightly less than 20% of dispositions were surgical in nature and just under 25% of outpatient visits occurred in the Emergency Room.

**Quality:** 74.6% of hospitals received JCAHO accreditation, and 7.4% of hospitals were members of the Council of Teaching Hospitals. The average hospital has a .009% lower mortality rate than expected and a 31.9% longer length of stay than expected. Of surgery patients, 82.3% received the recommended venous thromboembolism prophylaxis ordered, and 93.2% of pneumonia patients received their first antibiotic within 6 hours of hospital arrival.
Physician Characteristics: The average number of malpractice claims per capita in the state where the hospital is located is .35 per 1,000, and 23.9% of hospitals had no physician employees.

Patient Characteristics: Patients of the average hospital were 50.29 years old. 58.5% were female, 5.1% were either self-insured or uninsured, and 18.4% were non-white.
Table 6.4 below provides mean descriptive statistics broken down by ownership and a comparison of military to overall average civilian hospitals:

Table 6.4: Sample Mean Descriptive Statistics By Ownership

<table>
<thead>
<tr>
<th>Variable</th>
<th>State/Loc Gov</th>
<th>Not-for-Profit</th>
<th>For-Profit</th>
<th>Military</th>
<th>Sig.</th>
<th>Civ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Production Function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dispositions</td>
<td>5,649.5</td>
<td>10,181.2</td>
<td>7,538.5</td>
<td>5,250.0</td>
<td>**</td>
<td>8,905.5</td>
</tr>
<tr>
<td>Outpatient Visits</td>
<td>114.8K</td>
<td>166.3K</td>
<td>59.1K</td>
<td>538.8K</td>
<td>***</td>
<td>141.7K</td>
</tr>
<tr>
<td>Beds</td>
<td>129.1</td>
<td>200</td>
<td>150.2</td>
<td>93.6</td>
<td>***</td>
<td>179.2</td>
</tr>
<tr>
<td>RN FTEs</td>
<td>195.5</td>
<td>292.3</td>
<td>182.3</td>
<td>179.3</td>
<td></td>
<td>258.4</td>
</tr>
<tr>
<td>Other FTEs</td>
<td>562.6</td>
<td>810.3</td>
<td>381.8</td>
<td>1,266.60</td>
<td>***</td>
<td>702.9</td>
</tr>
<tr>
<td><strong>Efficiency Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case Mix Index-Avg</td>
<td>0.954</td>
<td>1.062</td>
<td>1.097</td>
<td>0.777</td>
<td>***</td>
<td>1.05</td>
</tr>
<tr>
<td>Case Mix Index-Med</td>
<td>0.835</td>
<td>0.866</td>
<td>0.887</td>
<td>0.603</td>
<td>***</td>
<td>0.863</td>
</tr>
<tr>
<td>Case Mix Index-SD</td>
<td>0.668</td>
<td>0.894</td>
<td>0.921</td>
<td>0.794</td>
<td></td>
<td>0.853</td>
</tr>
<tr>
<td>% Surgery DRGs</td>
<td>13.69%</td>
<td>20.57%</td>
<td>23.94%</td>
<td>24.10%</td>
<td>**</td>
<td>19.66%</td>
</tr>
<tr>
<td>% ER Outpatient</td>
<td>21.96%</td>
<td>24.43%</td>
<td>39.18%</td>
<td>6.25%</td>
<td>***</td>
<td>25.80%</td>
</tr>
<tr>
<td>% MDC 14 Work</td>
<td>8.74%</td>
<td>10.14%</td>
<td>9.43%</td>
<td>23.19%</td>
<td>***</td>
<td>9.70%</td>
</tr>
<tr>
<td>% MDC 5 Work</td>
<td>14.26%</td>
<td>15.29%</td>
<td>16.59%</td>
<td>7.38%</td>
<td>***</td>
<td>15.30%</td>
</tr>
<tr>
<td>Average Age</td>
<td>54.09</td>
<td>52.02</td>
<td>51.3</td>
<td>28.7</td>
<td>***</td>
<td>52.3</td>
</tr>
<tr>
<td>% Female</td>
<td>58.50%</td>
<td>58.60%</td>
<td>58.03%</td>
<td>57.82%</td>
<td></td>
<td>58.49%</td>
</tr>
<tr>
<td>% Self/No Pay</td>
<td>7.63%</td>
<td>4.63%</td>
<td>6.04%</td>
<td>0.00%</td>
<td>***</td>
<td>5.42%</td>
</tr>
<tr>
<td>% Non-White</td>
<td>11.96%</td>
<td>17.15%</td>
<td>24.48%</td>
<td>36.37%</td>
<td>***</td>
<td>17.11%</td>
</tr>
<tr>
<td>Length of Stay</td>
<td>4.75</td>
<td>4.09</td>
<td>4.28</td>
<td>2.82</td>
<td>***</td>
<td>4.25</td>
</tr>
<tr>
<td>Mortality Rate</td>
<td>2.08%</td>
<td>2.20%</td>
<td>2.22%</td>
<td>0.66%</td>
<td>***</td>
<td>2.21%</td>
</tr>
<tr>
<td>JCAHO Accreditation</td>
<td>48.41%</td>
<td>77.27%</td>
<td>88.24%</td>
<td>97.73%</td>
<td>***</td>
<td>73.24%</td>
</tr>
<tr>
<td>COTH Member</td>
<td>9.52%</td>
<td>8.61%</td>
<td>2.35%</td>
<td>0.00%</td>
<td>***</td>
<td>8.01%</td>
</tr>
<tr>
<td>Process Pneumonia</td>
<td>91.89%</td>
<td>94.02%</td>
<td>92.76%</td>
<td>88.22%</td>
<td>***</td>
<td>93.56%</td>
</tr>
<tr>
<td>Process Surgery</td>
<td>80.50%</td>
<td>83.33%</td>
<td>76.14%</td>
<td>86.32%</td>
<td></td>
<td>81.90%</td>
</tr>
<tr>
<td>% Over Exp Mort</td>
<td>32.19%</td>
<td>-2.22%</td>
<td>-13.67%</td>
<td>-59.65%</td>
<td>***</td>
<td>3.09%</td>
</tr>
<tr>
<td>% Over Exp LOS</td>
<td>4.21%</td>
<td>-8.02%</td>
<td>-6.91%</td>
<td>-13.62%</td>
<td>-5.94%</td>
<td></td>
</tr>
<tr>
<td>&lt;50 Beds</td>
<td>46.83%</td>
<td>22.49%</td>
<td>21.18%</td>
<td>34.09%</td>
<td></td>
<td>27.24%</td>
</tr>
<tr>
<td>Malpr Claims per cap (000s)</td>
<td>0.35</td>
<td>0.39</td>
<td>0.38</td>
<td>0.00</td>
<td>***</td>
<td>0.38</td>
</tr>
<tr>
<td>100% Credentialed MDs</td>
<td>27.78%</td>
<td>23.44%</td>
<td>32.94%</td>
<td>0.00%</td>
<td>***</td>
<td>25.64%</td>
</tr>
<tr>
<td>HSA Market Share</td>
<td>79.87%</td>
<td>71.31%</td>
<td>49.87%</td>
<td>20.39%</td>
<td>***</td>
<td>70.12%</td>
</tr>
<tr>
<td>HSA HHI</td>
<td>8,375.8</td>
<td>7,275.9</td>
<td>5,755.8</td>
<td>4,467.7</td>
<td>***</td>
<td>7,290.9</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, and * p<0.10.
In addition to a breakout of sample descriptive statistics, Table 6.4 above provides results of tests for significance of differences between military and civilian (average of state/local government, not-for-profit, and for-profit) hospitals. With respect to the production function variables, military hospitals produce significantly fewer dispositions (5,250 vs. 8,905.5; p<.05) than civilian hospitals, although they are comparable to other government hospitals. Military hospitals produce significantly more outpatient visits (538K vs. 141.7K; p<.01), and report significantly more Non-MD/Non-RN FTEs (1,266 vs. 703; p<.01). Numbers of RN FTEs are similar across ownership type.

**Structure:** The scope of work performed in military hospitals is considerably less resource intensive than that performed in civilian hospitals. Both the average case mix index (HOSPCMI) and the median case mix index (HOSPMEDCMI) are significantly less for military hospitals. The standard deviation, while slightly higher than that of state and local government hospitals, is also less than the civilian average. Table 6.7 will describe this difference in scope in more detail. A higher percentage of dispositions are surgical in nature in military hospitals than in civilian ones, likely reflective of a younger patient base with fewer chronic maladies. The percentage of outpatient work performed in the ER is significantly less than it is in civilian hospitals. Military hospitals perform significantly more work in MDC 14 (Pregnancy, Childbirth, and Puerperium) and significantly less in MDC 5 (Circulatory System), differences enumerated in more detail in Table 6.8.

**Environment:** While military hospitals do not compete on the same basis as civilian hospitals, in essence, they do compete for Tricare beneficiaries in markets where civilian hospitals are available. Family members and retirees can choose to receive care in the civilian sector. Thus, analysis of the effects of competition are possible, but should be made with caution. On average, military hospitals occupy a smaller portion of their relative market, and are located in markets with greater competition than civilian
hospitals. For-profits own a smaller share of the market and are located in less concentrated markets than state/local government and not-for-profit hospitals.

Quality: With respect to quality, military hospitals compare favorably to civilian hospitals. Military hospitals performed better on average on the surgical process measure than any of the other three ownership types. This could be attributed in part to the greater percentage of surgical work performed in military hospitals (i.e. “practice makes perfect”). Although pneumonia process scores were significantly worse (p<.01) for military hospitals, the difference was not large (88.2% vs. 93.6%). Similarly, this could be attributed in part to the fact that pneumonia cases represent a smaller percentage of MDC 4 workload in military hospitals than in civilian ones. Military hospitals outperformed their expected mortality rate by 60% and their expected length of stay by 13.6%, although this measure was not significantly different from that of civilian hospitals. Finally, almost all military hospitals (97.7%) undergo JCAHO accreditation, compared to 73.24% for civilian hospitals, although no military hospitals are members of the COTH.

Physician Characteristics: As discussed in Chapter 5, the working relationship between hospitals and doctors differs for military hospitals. Essentially all military physicians perform as employees, and essentially none of them faces the financial risk associated with malpractice. Malpractice claims per capita for the other three ownership types are quite similar, as might be expected since this is based on the state in which the hospital is located. The number of hospitals with zero physician employees is greater for for-profit facilities (32.94%) than for either state/ local government (27.78%) or not-for-profit ones (23.44%).
**Patient Characteristics:** Patients in military hospitals differ significantly from those seen in the other three ownership types, except in the percentage of females. Their patients have shorter lengths of stay (2.82 days vs. 4.3 days; p<.01) and are younger (28.7 vs. 52.3; p<.01). A higher proportion of military patients are non-white (36.4% vs. 17.1%; p<.01). Finally, Tricare provides insurance for all beneficiaries.
Table 6.5: Top Ten Military Inpatient DRGs

<table>
<thead>
<tr>
<th>RANK</th>
<th>DRG</th>
<th>FREQ</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>373</td>
<td>31,162</td>
<td>Vaginal Delivery W/O Complicating Dx</td>
</tr>
<tr>
<td>2</td>
<td>391</td>
<td>25,880</td>
<td>Normal Newborn</td>
</tr>
<tr>
<td>3</td>
<td>630</td>
<td>11,469</td>
<td>Neonate, Birthwt &gt;2499g, Other Prob</td>
</tr>
<tr>
<td>4</td>
<td>371</td>
<td>9,710</td>
<td>Cesarean Section W/O Cc</td>
</tr>
<tr>
<td>5</td>
<td>143</td>
<td>6,852</td>
<td>Chest Pain</td>
</tr>
<tr>
<td>6</td>
<td>372</td>
<td>6,560</td>
<td>Vaginal Delivery W Complicating Diagnoses</td>
</tr>
<tr>
<td>7</td>
<td>359</td>
<td>3,978</td>
<td>Uterine &amp; Adnexa Proc For Non-Malignancy W/O Cc</td>
</tr>
<tr>
<td>8</td>
<td>430</td>
<td>3,111</td>
<td>Psychoses</td>
</tr>
<tr>
<td>9</td>
<td>183</td>
<td>2,977</td>
<td>Esophagitis, Gastroent. &amp; Digest. Disorders &gt;17 W/O Cc</td>
</tr>
<tr>
<td>10</td>
<td>370</td>
<td>2,856</td>
<td>Cesarean Section W Cc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>104,555</td>
<td>TOTAL TOP 10 Drgs</td>
</tr>
</tbody>
</table>

Total Military Dispositions: 255,726
40.9% % Dispositions Accounted For

Table 6.6: Top Ten Civilian DRGs

<table>
<thead>
<tr>
<th>RANK</th>
<th>DRG</th>
<th>FREQ</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>391</td>
<td>447,819</td>
<td>Normal Newborn</td>
</tr>
<tr>
<td>2</td>
<td>373</td>
<td>348,984</td>
<td>Vaginal Delivery W/O Complicating Dx</td>
</tr>
<tr>
<td>3</td>
<td>430</td>
<td>153,057</td>
<td>Psychoses</td>
</tr>
<tr>
<td>4</td>
<td>371</td>
<td>144,007</td>
<td>Cesarean Section W/O Cc</td>
</tr>
<tr>
<td>5</td>
<td>127</td>
<td>140,469</td>
<td>Heart Failure &amp; Shock</td>
</tr>
<tr>
<td>6</td>
<td>544</td>
<td>117,740</td>
<td>Major Joint Replacement Or Reattachment Of Lower Extremity</td>
</tr>
<tr>
<td>7</td>
<td>89</td>
<td>115,736</td>
<td>Simple Pneumonia &amp; Pleurisy Age &gt;17 W Cc</td>
</tr>
<tr>
<td>8</td>
<td>143</td>
<td>115,568</td>
<td>Chest Pain</td>
</tr>
<tr>
<td>9</td>
<td>182</td>
<td>101,619</td>
<td>Esophagitis, Gastroent. &amp; Digest. Disorders &gt;17 W Cc</td>
</tr>
<tr>
<td>10</td>
<td>88</td>
<td>95,793</td>
<td>Chronic Obstructive Pulmonary Disease</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1,780,792</td>
<td>TOTAL TOP 10 Drgs</td>
</tr>
</tbody>
</table>

Total Civilian Dispositions: 5,725,536
31.1% % Dispositions Accounted For
Tables 6.5 and 6.6 provide the top ten DRGs for military and civilian hospital dispositions in 2006. Seven DRGs appear in both lists, indicating some broad similarity. The top ten DRGs account for 40.9% of military inpatient dispositions, yet only account for 31.1% of civilian inpatient dispositions, indicating on average, a narrower range of services provided in military facilities. While dispositions related to childbirth (italicized in the tables above) are a large percentage of both military and civilian hospitals, they comprise a more substantial proportion of military workload. Six of the top ten DRGs are pregnancy-related for military hospitals, accounting for 34.3% of dispositions, while only three are pregnancy-related for civilian hospitals, accounting for 16.4% of dispositions.

<table>
<thead>
<tr>
<th>DRG Weight (Case Mix)</th>
<th>State/Local</th>
<th>Not-For-Profit</th>
<th>For Profit</th>
<th>Military</th>
<th>Sig.</th>
<th>Civ.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRG≤1</td>
<td>64.8%</td>
<td>59.8%</td>
<td>59.6%</td>
<td>75.6%</td>
<td>***</td>
<td>60.75%</td>
</tr>
<tr>
<td>1&lt;DRG≤2</td>
<td>30.1%</td>
<td>31.2%</td>
<td>30.6%</td>
<td>16.3%</td>
<td>***</td>
<td>30.92%</td>
</tr>
<tr>
<td>DRG&gt;2</td>
<td>4.2%</td>
<td>7.8%</td>
<td>8.5%</td>
<td>4.7%</td>
<td>***</td>
<td>7.68%</td>
</tr>
</tbody>
</table>

Analyzing the complexity of inpatient work performed even further, table 6.7 above shows the breakdown of DRG weights into three categories: 1) below average (≤1), 2) average to above average (1<DRGs≤2), and 3) much greater than average (>2), by ownership type. Over ¾ of military inpatient healthcare is considered to be of below-average intensity. Only 16.3% of their inpatient work is classified as average or above average, compared to approximately 31% for civilian hospitals. The amount of high-complexity cases is similar for both government categories, but almost ½ of that performed by for-profits and not-for-profits. For all three categories, the difference between military and civilian hospitals is significant at the $p<.01$ level.
## Table 6.8: Inpatient Work Profile

<table>
<thead>
<tr>
<th>MDC</th>
<th>Govt</th>
<th>NFP</th>
<th>FP</th>
<th>Mil</th>
<th>Sig</th>
<th>Civ</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCTMDC1</td>
<td>Nervous Sys</td>
<td>4.74%</td>
<td>4.81%</td>
<td>4.61%</td>
<td>2.38%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC2</td>
<td>Eye</td>
<td>0.10%</td>
<td>0.10%</td>
<td>0.09%</td>
<td>0.23%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC3</td>
<td>Ear/Nose/Mouth/Throat</td>
<td>1.28%</td>
<td>1.03%</td>
<td>0.93%</td>
<td>2.87%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC4</td>
<td>Respiratory</td>
<td>15.70%</td>
<td>12.43%</td>
<td>11.97%</td>
<td>5.91%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC5</td>
<td>Circulatory</td>
<td>14.26%</td>
<td>15.29%</td>
<td>16.59%</td>
<td>8.52%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC6</td>
<td>Digestive</td>
<td>10.54%</td>
<td>10.00%</td>
<td>9.05%</td>
<td>8.47%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC7</td>
<td>Hepatobiliary/Pancreas</td>
<td>3.07%</td>
<td>2.94%</td>
<td>3.02%</td>
<td>2.07%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC8</td>
<td>Musculoskeletal/Connective Tissue</td>
<td>6.32%</td>
<td>7.93%</td>
<td>8.66%</td>
<td>7.28%</td>
<td>7.71%</td>
</tr>
<tr>
<td>PCTMDC9</td>
<td>Skin/Subcutaneous/Breast Endocrine/Metabolic</td>
<td>3.05%</td>
<td>2.68%</td>
<td>3.38%</td>
<td>3.71%</td>
<td>*</td>
</tr>
<tr>
<td>PCTMDC10</td>
<td>Kidney/Urinary Tract</td>
<td>4.28%</td>
<td>3.83%</td>
<td>3.88%</td>
<td>2.89%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC11</td>
<td>Pregnancy/Childbirth/Puerperium</td>
<td>4.42%</td>
<td>4.26%</td>
<td>4.25%</td>
<td>2.41%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC12</td>
<td>Male Reproduction</td>
<td>0.40%</td>
<td>0.45%</td>
<td>0.51%</td>
<td>0.46%</td>
<td>0.45%</td>
</tr>
<tr>
<td>PCTMDC13</td>
<td>Female Reproduction</td>
<td>1.69%</td>
<td>1.93%</td>
<td>2.48%</td>
<td>4.41%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC14</td>
<td>Newborn/Perinatal</td>
<td>8.74%</td>
<td>10.10%</td>
<td>9.43%</td>
<td>24.78%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC15</td>
<td>Blood/Immune</td>
<td>8.04%</td>
<td>9.41%</td>
<td>8.53%</td>
<td>15.33%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC16</td>
<td>Myeloproliferative DDs</td>
<td>1.26%</td>
<td>1.09%</td>
<td>0.92%</td>
<td>1.02%</td>
<td>1.10%</td>
</tr>
<tr>
<td>PCTMDC17</td>
<td>Infectious/Parasitic</td>
<td>0.50%</td>
<td>0.61%</td>
<td>0.29%</td>
<td>0.35%</td>
<td>0.54%</td>
</tr>
<tr>
<td>PCTMDC18</td>
<td>Mental</td>
<td>1.94%</td>
<td>1.98%</td>
<td>1.99%</td>
<td>1.12%</td>
<td>***</td>
</tr>
<tr>
<td>PCTMDC19</td>
<td>Alcohol/Drug/Mental</td>
<td>2.37%</td>
<td>3.14%</td>
<td>3.17%</td>
<td>2.50%</td>
<td>2.99%</td>
</tr>
<tr>
<td>PCTMDC20</td>
<td>Injuries/Poison/Drugs</td>
<td>0.94%</td>
<td>1.33%</td>
<td>1.24%</td>
<td>0.18%</td>
<td>1.24%</td>
</tr>
<tr>
<td>PCTMDC21</td>
<td>Burns</td>
<td>1.45%</td>
<td>1.34%</td>
<td>1.51%</td>
<td>1.76%</td>
<td>*</td>
</tr>
<tr>
<td>PCTMDC22</td>
<td>Factors Influencing Health Status</td>
<td>0.08%</td>
<td>0.04%</td>
<td>0.03%</td>
<td>0.10%</td>
<td>*</td>
</tr>
<tr>
<td>PCTMDC23</td>
<td>Multiple Significant Trauma</td>
<td>3.64%</td>
<td>1.92%</td>
<td>1.97%</td>
<td>1.12%</td>
<td>2.27%</td>
</tr>
<tr>
<td>PCTMDC24</td>
<td>HIV</td>
<td>0.15%</td>
<td>0.14%</td>
<td>0.15%</td>
<td>0.03%</td>
<td>*</td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, and * p<0.10.
Table 6-8 above compares inpatient work done by MDC across ownership types, and tests for significant differences between military and non-military (the average of for-profit, not-for-profit, and other government) facilities. Castenada and Falaschetti (2008) performed a similar analysis of 1997 HCUP NIS data to determine if scope of operations varied due to ownership type using the probability that a given hospital treats a given condition or performs a given procedure in a particular MDC as the dependent variable. However, Castenada and Falaschetti’s (2008) main interest was whether hospitals of different ownership chose to perform different scopes of medical care because of different incentives attributed to ownership type. The focus here is whether differences in scope are associated with efficiency.

Based on the number of significant differences shown in the table above, it is evident that the type of work done in military facilities does differ on average from that performed in civilian facilities. Military hospitals do significantly less work (p<0.01) in MDCs 1, 4, 5, 6, 7, 10, 11, and 18. Military hospitals do significantly more work (p<0.01) in MDCs 2, 3, 13, 14, and 15. The differences in MDCs 13 (Female Reproductive Disorders), 14 (Pregnancy, Childbirth, and Puerperium), and 15 (Newborn, and Perinatal) most likely reflect the fact that the patients treated in military hospitals (both Active Duty and family members) are on average younger (and therefore of childbearing age) than patients treated in civilian hospitals. Military hospitals also perform significantly fewer procedures in MDCs 4 (respiratory) and 5 (circulatory). While coronary artery bypass grafting and percutaneous transluminal coronary angioplasty (balloon and/or stent procedures) are the most heavily researched procedures in MDC 5, the top two diagnoses within this category are chest pain and heart failure. With respect to MDC 4, 40% of civilian workload in this category is attributed to two DRGs – chronic obstructive pulmonary disease and pneumonia with complications. These two DRGs, with the addition of bronchitis also comprise 40% of military workload in this DRG. In all of these cases, but especially obstetrics (childbirth), it is feasible that high volumes of a given
type of work improve efficiency due to learning by doing ("practice makes perfect") and specificity of work – creating an "assembly line". This possibility will be further explored by including the percentage of MDC 14 and MDC 5 as potential "influencers" of efficiency.

Summary
This chapter has described the sources of data used for this analysis, explained how this data was manipulated to create the actual data set, and presented descriptive statistics of this data. Based on the observations in this chapter, military hospitals differ significantly in many aspects from civilian hospitals. Military hospitals treat less complex cases: their patients are younger, relative DRG weights are less, length of stay is shorter, and mortality is lower. Military hospitals perform a much greater proportion of obstetric work than civilian hospitals. Finally, they perform considerably more outpatient work than civilian hospitals. Controlling for these differences, estimation of the proposed model can continue. The next chapter will describe in detail the statistical methods employed in the analysis.
CHAPTER 7 - STATISTICAL METHODS

Study Design
This study will be cross-sectional, therefore estimating long-run relationships between efficiency and ownership (Kennedy, 2008). Cross-sectional studies are not without criticism. They cannot address trends, and generally, they estimate the inefficiency term less consistently than panel data efficiency studies (Linna, 1998). Dor (1994) said that panel estimators are preferred to cross-sectional estimators because they are “less likely to yield biased estimates of the $\beta$s due to omitted variables, and because they require fewer distributional assumptions about the deterministic error ($u_i$).” However, Rosko and Mutter (2008) point out that for truly consistent estimation, the number of observations per unit of observation (hospital) must approach $\infty$, and since most panels contain fewer than ten observations per hospital, assuming a given panel will produce consistent estimation of inefficiency may be unjustifiable. Given this viewpoint, a cross-sectional design has been determined to be reasonable.

Stochastic Frontier Analysis
Although Data Envelopment Analysis has been more popular than Stochastic Frontier Analysis in studies of hospital technical efficiency, SFA serves as the primary measurement technique here due to its previously discussed econometric advantages. Stata/SE 9.2, the statistical software used for all SFA models, performs stochastic frontier analysis and allows for one-stage exploration of various assumptions associated with the distribution of the estimated inefficiency. DEA was used to validate the SFA results, and this validation is discussed in Chapter 9.

The basic Stochastic Frontier Analysis model developed in 1977 concurrently by two groups of researchers – 1) Aigner, Lovell, and Schmidt and 2) Meeusen and van den Broeck is as follows:
\[ \ln q_i = x_i' \beta + v_i - u_i, \] where

- \( q_i \) is the output of the \( i \)-th firm,
- \( x_i \) is a \( K \times 1 \) vector of the logarithms of the inputs,
- \( \beta \) is a parameter vector,
- \( u_i \) is a non-negative random variable representing technical inefficiency, and
- \( v_i \) is a symmetrically distributed random error term (Coelli, Prasada Rao, O'Donnell, & Battese, 2005).

SFA requires assumptions, some of which are the same as those required in ordinary least squares regression.

- \( v_i \) is assumed to have zero mean;
- \( v_i \) is assumed to have constant variance (homoscedasticity);
- \( v_i \) is assumed to be independently and identically distributed across all observations;
- \( u_i \) is assumed to have a given distribution, either half-normal, truncated-normal, gamma, or exponential;
- \( u_i \) is assumed to be independently and identically distributed across all observations; and
- \( u_i \) and \( v_i \) are distributed independently of each other and neither is correlated with the explanatory variables.

(Kumbhakar & Lovell, 2000)

To review the discussion of frontier methods in Chapter 2, SFA is based on the theory that a production function frontier represents the maximum output possible, given a set of inputs. Since the frontier represents an upper bound of production levels, the resulting error due to inefficiency (\( u_i \)) must be a subtraction from the production frontier. Once the inefficiency is accounted for, the remaining error (\( v_i \)) is assumed to be random noise, distributed as \( N(0, \sigma_v^2) \) (Coelli, Prasada Rao, O'Donnell, & Battese,
Thus, a fundamental purpose of SFA is to divide the traditional OLS error term into two pieces – inefficiency and random noise – and it is the inefficiency term in the model (rather than the parameter coefficients) in which researchers are usually most interested. SFA requires two key assumptions: one regarding the functional form that models the underlying production function and the other regarding the distribution of the inefficiency error term.

**Distance Functions**

In the basic SFA model just discussed, \( u_i = -\ln(D_0) \) and represents distance from the frontier (inefficiency). This distance from the frontier is a function of \( M \) outputs and \( N \) inputs. Cobb-Douglas is used here for convenience.

\[
\ln(D_{0i}) = \beta_0 + \sum_{m=1}^{M} \beta_m \ln y_{mi} + \sum_{n=1}^{N} \beta_n \ln x_{ni} + v_i
\]

Capitalizing on an assumption of linear homogeneity in the outputs (\( \sum_{n=1}^{N} \beta_n = 1 \)) and on the fact that \( u_i = -\ln(D_0) \), and re-arranging terms yields the following:

\[
\ln(y_{Mi}) = \beta_0 + \sum_{m=1}^{M-1} \beta_m \ln \left( \frac{y_{mi}}{y_{Mi}} \right) + \sum_{n=1}^{N} \beta_n \ln x_{ni} + v_i - u_i
\]


Coelli and Perelman note that Cobb-Douglas may not be suitable in distance functions because it is not concave in the output dimensions, and it is less flexible (Coelli & Perelman, 2000). This issue may be moot, however, depending on the final functional form indicated as optimal by likelihood ratio hypothesis testing discussed later.
Regardless of the final form, using a distance function will allow the specified model of hospital production and inefficiency to be explored without weighting and aggregating inpatient and outpatient workload.

Model Building
The assumptions just discussed, as well as a few others, are now considered in enumerating the final model. This process begins with assuming the most basic alternatives and then testing for the necessity of more complex modeling. Hypotheses testing of the stochastic frontier parameters can be performed using the generalized likelihood ratio statistic. The likelihood test statistic is:

$$\lambda = -2\ln(L(H_0)) - \ln(L(H_1)),$$

where $L(H_0)$ and $L(H_1)$ are the values of the likelihood function under null and alternative hypotheses (i.e., restricted and unrestricted versions of the model), respectively. With one exception, the resulting test statistic has an approximate chi-square distribution with degrees of freedom equal to the difference in the number of parameters in the null and alternative hypotheses. The resulting test statistic is compared to a chi-square critical value at $p<.05$. The test for the presence of inefficiency, however, follows a mixed chi-distribution (Rosko & Mutter, 2008).

Given the variables listed in Chapter 6 and the proposed analytical model in Chapter 5, hypothesis testing is based on the following:

*Weighted Case-Mix Adjusted Total Output = f(Capital, RN FTEs, Other FTEs), and

*Efficiency = f(CONTROLCAT, HOSPCMI ( or HOSPMEDCMI), HOSPSDCMI, PCTSURGERY, ERPERCENT, PCTMDC14, PCTMDC5, HOSPAVGAGE, PCTFEMALE, PCTNONWHITE, PCTSELFNOPAY, SHAREHSA, HSAHII, COTH, JCAHO, PCTOVERLOS, PCTOVERMORT, PNEUMONIA, SURGERY, CRED, CLAIMSPERCAP2000, SMALLHOSP)*
Appropriateness of SFA

Intuitively, it seems reasonable to assume that appreciable inefficiency exists in hospital healthcare provision. However, if this were not the case — or if the influence of the stochastic error overshadowed the influence of deterministic error — inefficiency would not be a statistically significant portion of the total error. The use of SFA would not be indicated, and OLS regression would be sufficient for analysis. The test for statistical significance of the deterministic inefficiency portion of total error involves computation of a statistic — $\gamma$. $\gamma$ is defined as the proportion of total $\sigma^2$ attributed to inefficiency.

$$\gamma = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$$

If $\gamma$ approaches 1, then nearly all of the error is due to (deterministic) inefficiency (Fiorentino, Karmann, & Koetter, 2006). If $\gamma$ approaches 0, then nearly all of the error is due to random error, and dividing the error term into two parts would not provide meaningful information regarding inefficiency. The test is:

$H_0$: $\gamma = 0$  

$H_1$: $\gamma \neq 0$

Using the most basic production function of Cobb-Douglas functional form along with a half-normal deterministic inefficiency error, $\gamma$ is calculated to be .76. The likelihood ratio test statistic for $\sigma_u$ based on a mixed chi$^2$ distribution (Rosko, 2004) was 39.56 ($p \leq 0.000$) (Kodde & Palm, 1986). This supports rejection of the null hypothesis. Inefficiency is a significant portion of the total error, and SFA is appropriate for the analysis.

Production Function

The parametric nature of SFA requires specification of a frontier functional form for the model. As discussed in Chapter 4, common functional forms used to model production functions include linear, log-linear, Cobb-Douglas, and transcendental logarithmic (“translog”), but the predominant functions assumed in healthcare efficiency studies are
Cobb-Douglas and Translog. The production function specified in this analysis is provided below for each functional form assumption.

Cobb-Douglas:

\[
\ln(\text{TotWork}) = \alpha + \beta_1 \ln(\text{HospBd}) + \beta_2 \ln(\text{FTERN}) + \beta_3 \ln(\text{FTEOTH}) + v_i - u_i
\]

Translog:

\[
\ln(\text{TotWork}) = \alpha + \beta_1 \ln(\text{HospBd}) + \beta_2 \ln(\text{FTERN}) + \beta_3 \ln(\text{FTEOTH}) + \beta_4 \frac{1}{2} [\ln(\text{HospBd})]^2 + \beta_5 \frac{1}{2} [\ln(\text{FTERN})]^2 + \beta_6 \frac{1}{2} [\ln(\text{FTEOTH})]^2 + \beta_7 [\ln(\text{HospBd}) \times \ln(\text{FTERN})] + \beta_8 [\ln(\text{HospBd}) \times \ln(\text{FTEOTH})] + \beta_9 [\ln(\text{FTEOTH}) \times \ln(\text{FTERN})] + v_i - u_i
\]

Cobb-Douglas is a special case of the Translog, where coefficients of square and cross-products are zero, and is therefore a nested model within translog, allowing for statistical testing of fit using a likelihood ratio test. Following the principle of parsimony, if the more restrictive Cobb-Douglas adequately represents the production function, it would be preferred to translog for further modeling. The likelihood ratio test performed follows:

\[H_0: \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0 \quad H_1: \text{At least one parameter } \neq 0\]

The resulting likelihood ratio test of the two models produced a \(\lambda = 52.70\), supporting rejection of the null hypothesis when compared with the \(\chi^2(6)\) value at an \(\alpha = .05\) of 12.59 (\(p \leq z = .0000\)). Thus, Cobb-Douglas does not adequately represent the production function. Translog better models the production of hospital care in this study.
Distribution of inefficiency error term

SFA also requires an assumption regarding the distribution of the inefficiency term. The inefficiency component \( u_i \) is always assumed to be strictly positive (to be subtracted from the production frontier). Common distributional assumptions of the error term used in the literature include half-normal, truncated-normal, gamma and exponential. The half-normal is a special case of the truncated normal distribution with the truncation at the mean of the normal distribution. The exponential is a one-parameter special case of the two-parameter gamma distribution. Stata allows for half-normal, truncated normal and exponential distributions. Diagrams of the half-normal, truncated normal and exponential distributions are shown below. The gamma distribution differs from these distributions in the pdf near zero, but the truncated normal can approximate it.

Figure 7.1: Error Term Distributions

![Half-Normal, Truncated-Normal, Exponential Distributions](image)

(Pascoe, Kirkley, Greboval, & Morrison-Paul, 2003)

Many researchers feel that the half-normal and exponential distributions are inappropriate because the mode or point of greatest density of the PDF is at zero – meaning the greatest density of observations (firms, hospitals etc.) experience near 100% efficiency. This excludes the possibility of a small probability of very efficient firms with most concentrated at some greater inefficiency. No firm is ever exactly 100%
efficient in SFA, and this is a major difference from DEA\(^5\). While the debate about which distribution best models the inefficiency term is important, several researchers point out that past SFA results have been fairly robust to the choice of error term distribution, and therefore the more restrictive half-normal error distribution assumption is adequate (Coelli, Prasada Rao, O'Donnell, & Battese, 2005) (Rosko & Mutter, 2008). However, since the half-normal represents a special case of the truncated-normal, a likelihood ratio test comparing restricted and unrestricted log likelihoods of each is both possible and appropriate. The test is:

\[ H_0: \ u \text{ distributed half-normal } (\mu_u=0) \quad H_1: \ u \text{ not distributed half normal } (\mu_u \neq 0) \]

In performing this test, the truncated normal model (using either Cobb-Douglas or Translog) failed to converge. However, this may not be problematic because the addition of specific variables thought to affect efficiency in one-stage estimation may improve model fit, allowing convergence to occur. Additionally, because the half-normal distribution is a special case of the truncated-normal distribution and a truncated-normal assumption allows for further investigation of factors influencing efficiency, the final model incorporates a truncated-normal inefficiency error distribution. Thus, \( \mu_u \) will be allowed to be located somewhere other than zero.

Inefficiency Variables
As just mentioned, the truncated normal model without any efficiency variables failed to converge. Adding the variables enumerated in the previous chapter thought to influence efficiency did allow convergence, with two exceptions. The process quality variables – PNEUMONIA and SURGERY – presented problems. Approximately 100 observations have no data for these variables. Furthermore, when included, \( \gamma \) – the statistic that indicates the degree of inefficiency – becomes very small, indicating inefficiency cannot be detected. Since reporting of these variables is optional for

\(^5\) DEA guarantees at least one 100% efficient observation occurs.
hospitals, higher quality hospitals may be self-selecting to report these measures. Because of these problems, these variables have been excluded from further analysis. The usefulness of the remaining variables in explaining efficiency can be tested. The null hypothesis is that they do not significantly affect $\mu_u$.

\[ H_0: \delta_1 = \delta_2 = \delta_3 = ... = \delta_{22} = 0 \quad H_1: \text{At least one parameter } \neq 0 \]

The resulting $\chi^2(22)$ test statistic was 83.12 ($p>\chi^2 = .0000$), indicating that the null hypothesis can be rejected. This indicates that the explanatory variables for inefficiency do improve the model fit, and should be included.

The results of the hypotheses testing just discussed are presented in the following table:

<table>
<thead>
<tr>
<th>Null</th>
<th>Test Statistic $\lambda$</th>
<th>$\chi^2_{0.95}$ Value</th>
<th>Decision</th>
<th>Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma = 0$</td>
<td>39.56</td>
<td>2.706</td>
<td>Reject</td>
<td>Inefficiency Exists</td>
</tr>
<tr>
<td>$\beta_4 = \beta_5 = ... = \beta_9 = 0$</td>
<td>52.70</td>
<td>12.59</td>
<td>Reject</td>
<td>Translog Production Function</td>
</tr>
<tr>
<td>$\delta_1 = \delta_2 = ... = \delta_{22} = 0$</td>
<td>83.12</td>
<td>33.92</td>
<td>Reject</td>
<td>Include inefficiency terms</td>
</tr>
<tr>
<td>$\mu = 0$</td>
<td>No Convergence</td>
<td>N/A</td>
<td>N/A</td>
<td>Truncated Normal Distribution of $u_i$</td>
</tr>
</tbody>
</table>

One vs. Two Stage Models

SFA can be performed using one-stage and two-stage variations. In a two-stage approach, SFA is used to produce inefficiency estimates for each observation. These estimates are then regressed against proposed explanatory variables of the inefficiency, generally using OLS or Tobit (which would account for censoring at zero associated with non-negativity). Kumbhakar and Lovell (2000) point out that a two-stage model requires assuming that the explanatory variables of the inefficiency and the independent variables (i.e. inputs) in the production function are uncorrelated. If they are correlated, then estimates of $\beta$, $\sigma_u$, and $\sigma_v$ are biased because of the omission of the
analyzed inefficiency explanatory variables from the production function. Furthermore, the second-stage efficiency estimates would be biased, too. Another major problem is that the first stage requires an assumption that \( u_i \) be distributed identically and independently across all observations, yet the second stage assumes that a functional relationship exists between the inefficiency represented by \( u_i \) in the first stage and the explanatory variables.

An early approach, seemingly more straightforward, included efficiency variables and their parameters in the model as follows:

\[
\ln(y_i) = \ln f(x_i; z_i; \beta) + v_i - u_i
\]

However, while this specification may provide additional information as to the production function, it does not specifically address the issue of identifying factors associated with being more or less efficient (Coelli, Prasada Rao, O'Donnell, & Battese, 2005). A more recent strategy, as specified by Kumbhakar, Ghosh, and McGuckin (1991), significantly advanced the one-stage approach so that estimation of efficiency and the parameters of proposed inefficiency explanatory variables is accomplished simultaneously, thereby avoiding problems associated with the two-stage approach. Its form is as follows:

\[
\ln(y_i) = f(x_i; \beta) + v_i - u_i, \quad u_i \sim N^*(z_i \delta, \sigma_u^2)
\]

Several researchers have empirically shown that this one-stage procedure leads to less biased and more efficient results (Rosko & Mutter, 2008). Given its econometric advantages and the goals of this analysis, the final model of hospital efficiency used in this analysis employs this one-stage approach. Stata allows use of the one-stage model
Heteroscedasticity

Heteroscedasticity refers to variance of a random variable that is not constant. If the model employed does not address heteroscedasticity, several consequences may occur. If $v_i$ is heteroscedastic and not modeled, estimates of technical efficiency will be biased, but parameters of the production function will not. If $u_i$ is heteroscedastic and not modeled, both parameters of the production function and estimates of technical efficiency will be biased, but in the opposite direction as bias associated with $v_i$. Therefore, if both $u_i$ and $v_i$ possess un-modeled heteroscedasticity, it is possible that overall bias is small (Kumbhakar & Lovell, 2000). Stata/SE 9.2 allows for modeling of heteroscedasticity of $u_i$ and $v_i$ by including variables thought to influence the variance of each term. However, since these models require an assumption of a half-normally distributed inefficiency term, they will be only used to validate results obtained in the one-stage translog, truncated-normal model that follows.
Final Proposed Model
Given the discussion presented in this chapter, the results of the hypotheses tests performed, and the variables defined in Chapter 6, the final one-stage, translog, truncated-normal model proposed for evaluation of hospital inefficiency across ownership types follows:

\[
\ln(\text{TotWork}) = \alpha + \beta_1 \ln(\text{HospBd}) + \beta_2 \ln(\text{FTERN}) + \beta_3 \ln(\text{FTEOTH}) + \beta_4 \frac{1}{2} \ln(\text{HospBd})^2 + \beta_5 \frac{1}{2} \ln(\text{FTERN})^2 + \beta_6 \frac{1}{2} \ln(\text{FTEOTH})^2 + \beta_7 [\ln(\text{HospBd}) \times \ln(\text{FTERN})] + \beta_8 [\ln(\text{HospBd}) \times \ln(\text{FTEOTH})] + \beta_9 [\ln(\text{FTEOTH}) \times \ln(\text{FTERN})] + v_i - u_i
\]

And, \(u_i \sim N^+(\theta, \sigma_u^2)\), where

\[
\theta = \delta_1(\text{CONTROLCAT NFP}) + \delta_2(\text{CONTROLCAT FP}) + \delta_3(\text{CONTROLCAT MIL}) + \\
\delta_4(\text{HOSPMEDCMI}) + \delta_5(\text{HOSPSDCMI}) + \delta_6(\text{PCTSURG}) + \delta_7(\text{ERPERCENT}) + \\
\delta_8(\text{PCTMDC14}) + \delta_9(\text{PCTMDC5}) + \delta_{10}(\text{HOSPAVGAGE}) + \delta_{11}(\text{PCTFEMALE}) + \\
\delta_{12}(\text{PCTNONWHITE}) + \delta_{13}(\text{PCTSELFNOPAY}) + \delta_{14}(\text{PCTOVERMORT}) + \delta_{15}(\text{PCTOVERLOS}) + \\
\delta_{16}(\text{JCAHO}) + \delta_{17}(\text{COTH}) + \delta_{18}(\text{SMALLHOSP}) + \delta_{19}(\text{HSAHHI}) + \delta_{20}(\text{SHAREHSA}) + \delta_{21} \\
(\text{CLAIMSPERCAP2000}) + \delta_{22}(\text{CRED}) + \varepsilon,
\]

where

\(u_i = \text{non-negative inefficiency term and } v_i = \text{random error (iid)}\)

However, four different model specifications – three exploring different methods of case-mix adjusting outputs and one exploring use of a distance function – will be considered to evaluate overall robustness of results.

Copyright © Linda Gail Kimsey 2009
CHAPTER 8 - RESULTS

Given the large volume of outpatient workload that military hospitals perform in relation to other ownership types, four specifications of the general model just presented in Chapter 7 will be explored to assess different alternatives for handling patient heterogeneity in outpatient visits. The first three specifications vary only with respect to TOTWORK, the dependent variable. Inpatient and outpatient workload are aggregated as discussed in Chapter 6. Median case mix index is an explanatory variable for $u_i$, the inefficiency term. The fourth specification employs a distance function methodology to alleviate the need for aggregation of inpatient and outpatient workload.

- Model 1 – (CMI): Both inpatient and outpatient workload are adjusted for severity using the median inpatient case mix index. It is assumed that this case mix is reflective of the complexity of outpatient work in addition to that of inpatient dispositions.
- Model 2 – (Raw): No case mix indexing is used for either inpatient or outpatient workload.
- Model 3 – (CMI_In): Only inpatient workload is adjusted using the median case mix index.
- Model 4 (Dist): Model 3 is modified to incorporate a distance function approach. This obviates the need for aggregating inpatient and outpatient workload. Average case mix index is used to control for overall work complexity due to convergence problems when employing the median value.

After describing the model results for both the production function and the corresponding inefficiency obtained by maximum likelihood estimation, further analysis of relative efficiency percentages will be performed. Pooled and partitioned results will be compared to determine how different military and civilian hospitals are. Finally, the possibility of heteroscedasticity and its effects will be considered.
Final Results

The results of the main production function of the final specifications are shown below:

Table 8.1: Production Function

<table>
<thead>
<tr>
<th>Variable</th>
<th>CMI</th>
<th>Raw</th>
<th>CMI_In</th>
<th>Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(totalwork)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Hospital Beds)</td>
<td>0.2253</td>
<td>0.2796</td>
<td>0.3215</td>
<td>*</td>
</tr>
<tr>
<td>ln(RN FTEs)</td>
<td>0.9333</td>
<td>**</td>
<td>**</td>
<td>0.4069</td>
</tr>
<tr>
<td>ln(Other FTEs)</td>
<td>0.2117</td>
<td>0.2861</td>
<td>0.2566</td>
<td>0.4607</td>
</tr>
<tr>
<td>ln(Beds)^2</td>
<td>-0.0876</td>
<td>-0.0593</td>
<td>-0.1016</td>
<td>-0.1123</td>
</tr>
<tr>
<td>ln(RN FTEs)^2</td>
<td>0.1930</td>
<td>0.1356</td>
<td>0.1747</td>
<td>0.1345</td>
</tr>
<tr>
<td>ln(Other FTEs)^2</td>
<td>0.2973</td>
<td>*</td>
<td>**</td>
<td>0.1770</td>
</tr>
<tr>
<td>ln(Beds) x ln(Other FTEs)</td>
<td>-0.0142</td>
<td>-0.0335</td>
<td>-0.0162</td>
<td>-0.0074</td>
</tr>
<tr>
<td>ln(RN FTEs) x ln(Beds)</td>
<td>-0.3251</td>
<td>**</td>
<td>**</td>
<td>-0.2196</td>
</tr>
<tr>
<td>ln(y*)</td>
<td>0.1043</td>
<td>0.0919</td>
<td>0.1014</td>
<td></td>
</tr>
<tr>
<td>ln(y*) x ln(Beds)</td>
<td></td>
<td></td>
<td></td>
<td>-0.3057***</td>
</tr>
<tr>
<td>ln(y*) x ln(RN FTEs)</td>
<td></td>
<td></td>
<td>0.0699</td>
<td></td>
</tr>
<tr>
<td>ln(y*) x ln(Other FTEs)</td>
<td>0.1043</td>
<td>0.0919</td>
<td>0.1014</td>
<td></td>
</tr>
<tr>
<td>ln(y*)^2</td>
<td>-0.0374</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2.7578</td>
<td>***</td>
<td>2.5344</td>
<td>***</td>
</tr>
</tbody>
</table>

*** p<0.001, ** p<0.01, and * p<0.05.

Production Function

There is broad agreement between the models with respect to the production function, and in all specifications, the production function is well-behaved (i.e. the coefficients on the first-order terms are positive). The sum of the first-order coefficients for hospital beds, RNs, and other labor in all four models is greater than one, suggesting increasing
returns to scale, and in each case, Wald statistics for the constraint $FTERN + FTEOTH + HOSPBD=1$ supports this hypothesis at the $p<.05$ level. The coefficients of these variables indicate a positive elasticity in total work with respect to changes in hospital beds (.23 - .39), RNs (.41 - .93), and other labor (.21 - .46), holding the other inputs constant. Thus, in specification $CMI$, a one percent increase in RNs produces a .93% increase in total work, holding hospital beds and other FTEs constant. RNs are significant ($p<.001$) in all but the distance function specification. Beds are significant ($p<.05$) in specifications $CMI_IN$ and $Dist$. The coefficients on the squared terms reveal that investment in beds (-.05 - .11) yields decreasing returns to scale while investment in nurses (.13 - .19) and other personnel (.18 - .30) yield increasing returns to scale. Of the squared terms, only $fteoth^2$ is significant in three of the specifications.

The coefficients on the cross-product variables are indicators of input complementarity/substitutability. The negative sign on $\ln(Other FTEs) \times \ln(RN FTEs)$ indicates that RNs and other FTEs are substitutes, suggesting that the production of patient care in the hospital involves many other tasks than just the direct interaction of nurses and patients, and perhaps reflects the importance of administrative personnel. This variable is significant ($p<.01$) in the first three model specifications and remains negative in the distance function specification. In specification $CMI$, this coefficient indicates that a one percent increase in other FTEs should reduce the number of RNs required by .33%. The negative sign on $\ln(beds) \times \ln(Other FTEs)$ indicates some substitutability of other FTEs and hospital beds. Beds and RNs are slightly complementary, as indicated by the positive sign of this coefficient in all specifications. Neither of these last two cross-product variables is significant in any model, suggesting that inclusion of outpatient workload as an output reduces the importance of hospital beds in the production of a hospital’s total work.

In the distance function, the dependent variable is the natural log of case mix adjusted dispositions. The coefficient of $\ln(y^*)$ (Outpatient Dispositions/Inpatient Visits) is
negative and significant ($p<.001$). This is merely a function of the design of this variable, providing no real information. If total output includes inpatient and outpatient workload, the elasticity of each of the components will be inversely related: more outpatient work means less inpatient work. The only other distance function-specific variable that is significant is the squared term – $ln(y^*)^2$ – and its importance is again design-related.

The results of the analysis on efficiency-related factors of the final specifications are shown in Table 8.2 below.
<table>
<thead>
<tr>
<th></th>
<th>CMI</th>
<th>Raw</th>
<th>CMI_In</th>
<th>Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not-For-Profit</td>
<td>-0.1352</td>
<td>-0.1352</td>
<td>-0.2037</td>
<td>-0.2618 *</td>
</tr>
<tr>
<td>For-Profit</td>
<td>-0.0047</td>
<td>0.0397</td>
<td>0.0295</td>
<td>0.0263</td>
</tr>
<tr>
<td>Military</td>
<td>0.3017</td>
<td>-0.2456</td>
<td>-0.7422</td>
<td>-0.4383</td>
</tr>
<tr>
<td>Case Mix Index - Med</td>
<td>-1.5229 *</td>
<td>1.6599 *</td>
<td>-0.7101</td>
<td>1.0936</td>
</tr>
<tr>
<td>Case Mix Index- SD</td>
<td>-0.3502</td>
<td>-0.4422</td>
<td>-0.5191</td>
<td>-1.0927 **</td>
</tr>
<tr>
<td>% Surgery DRGs</td>
<td>-1.5581 *</td>
<td>-1.5355 *</td>
<td>-1.8838 *</td>
<td>-2.5704 *</td>
</tr>
<tr>
<td>% ER Outpatient</td>
<td>0.2266</td>
<td>0.0833</td>
<td>0.3240</td>
<td>0.9050</td>
</tr>
<tr>
<td>% MDC 14 Work</td>
<td>1.2877</td>
<td>-0.2597</td>
<td>0.7582</td>
<td>1.1570 **</td>
</tr>
<tr>
<td>% MDC 5 Work</td>
<td>-5.9287 ***</td>
<td>-6.0140 ***</td>
<td>-7.7408 ***</td>
<td>-8.1101 *</td>
</tr>
<tr>
<td>Average Age</td>
<td>0.0396 ***</td>
<td>0.0390 ***</td>
<td>0.0423 ***</td>
<td>0.0291 **</td>
</tr>
<tr>
<td>% Female</td>
<td>-1.6670 *</td>
<td>-2.5905 *</td>
<td>-2.5348 *</td>
<td>-2.3345 *</td>
</tr>
<tr>
<td>% Self/No Pay</td>
<td>0.9042</td>
<td>1.3081 *</td>
<td>0.9992</td>
<td>0.6574</td>
</tr>
<tr>
<td>% Non-White</td>
<td>0.3283</td>
<td>0.3425</td>
<td>0.5547</td>
<td>0.7988</td>
</tr>
<tr>
<td>JCAHO</td>
<td>-0.1774 *</td>
<td>-0.2725 *</td>
<td>-0.2666 *</td>
<td>-0.3140</td>
</tr>
<tr>
<td>COTH</td>
<td>0.0858</td>
<td>0.2943</td>
<td>0.0611</td>
<td>0.1452 **</td>
</tr>
<tr>
<td>Exp-Actual Mortality</td>
<td>-0.0515</td>
<td>-0.0308</td>
<td>-0.0743</td>
<td>-0.0737</td>
</tr>
<tr>
<td>Exp-Actual LOS</td>
<td>0.0825</td>
<td>0.0466</td>
<td>0.0889</td>
<td>0.0913 *</td>
</tr>
<tr>
<td>&lt;50 Beds</td>
<td>-0.2257</td>
<td>-0.4524 *</td>
<td>-0.4540 **</td>
<td>-0.5433</td>
</tr>
<tr>
<td>Malpr Claims per cap</td>
<td>-0.3108</td>
<td>-0.6397</td>
<td>-0.5677</td>
<td>-0.7924</td>
</tr>
<tr>
<td>100% Credntld MDs</td>
<td>-0.2936 **</td>
<td>-0.3440 *</td>
<td>-0.4045 **</td>
<td>-0.3928</td>
</tr>
<tr>
<td>HSA Market Share</td>
<td>-0.0025</td>
<td>-0.0026</td>
<td>-0.0017</td>
<td>-0.0005</td>
</tr>
<tr>
<td>HSA HHI</td>
<td>0.0001 *</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0000</td>
</tr>
<tr>
<td>Constant</td>
<td>1.4841 *</td>
<td>-0.4686</td>
<td>1.5312</td>
<td>0.9649</td>
</tr>
</tbody>
</table>

**Ilgtgamma**

Constant 1.0243 *** 1.1838 *** 1.3180 *** 1.5650 **

**Insigma2**

Constant -1.8155 *** -1.5843 *** -1.5512 *** -1.5254 **

**Gamma**

.7358 .7656 .7888 .8271

*** p<0.001, ** p<0.01, and * p<0.05.
**Inefficiency Variables**

In Table 8.2, coefficients reflect the relationship between the variable and inefficiency. Therefore, positive coefficients are associated with increased inefficiency (and decreased efficiency). The portion of total error attributed to inefficiency – $\gamma$ (discussed in Chapter 7 and presented in Table 8.2) is high in all four specifications, ranging from 73.6% to 82.7%.

**Ownership:** In all models, the only instance where the correlation between ownership and efficiency is found to be significantly different from the reference group is in specification Dist – the distance function: not-for-profit ownership was significantly ($p<.05$) correlated with less inefficiency than other government hospitals – the reference group. However, assessment of whether significant differences exist between other categories requires further testing as presented in Table 8.3 below.

<table>
<thead>
<tr>
<th>Table 8.3: Tests of Significance for Differences By Ownership Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi$^2$(1) Scores</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Military vs. Not-For-Profit</td>
</tr>
<tr>
<td>Military vs. For-Profit</td>
</tr>
<tr>
<td>For-Profit vs. Not-For-Profit</td>
</tr>
</tbody>
</table>

Table 8.3 reports results of Wald tests for equality between other ownership types as listed for comparisons not involving state and local hospitals – the model reference category. A $\chi^2(1)$ statistic of 3.841 or greater is required for significance at the $p<.05$ level. When compared to this critical value, the resulting $\chi^2$ statistics with one degree of freedom in the table above show that there are no other significant differences in efficiency among any pair of hospital types.
**Environment:** The coefficient for the variable representing the HHI a hospital faces is slightly above zero in all specifications. Although it is significant at the $p<.05$ level in only the first model, the consistent positive sign of this coefficient suggests that increased market concentration may be correlated with greater inefficiency. The coefficient on the variable representing market share is slightly negative. Although insignificant in all specifications, the consistent negative sign suggests that greater market share may be correlated with less inefficiency. Overall, no specification provides strong support for either Leibenstein’s (1966) or Carroll’s (1990) theories on the relationship between competition and efficiency.

**Structure:** 1) Size: Having fewer than fifty beds is associated with less inefficiency and is significant in two specifications at varying levels of significance. This supports the theories of Spann (1977) and Liebenstein (1966) as well as research by Sexton and colleagues (1989), which found smaller VA hospitals to be more efficient than large ones. Thus, economies of scale do not appear to be relevant. 2) Scope: When TOTWORK – both inpatient and outpatient workload – is adjusted for case mix (Model CMI), a higher median index (more complex work) is correlated with less inefficiency. In the Raw specification, a lower median index (less complex work) is correlated with less inefficiency. In both cases, this correlation is significant at the $p<.05$ level. This swing in associations seems counterintuitive initially. However, if the case mix index is reflective of the experience required to perform the associated output, the results make more sense. When raw workload numbers are used, complex procedures requiring substantial experience to perform receive the same credit as simpler procedures requiring less experience, and experience is actually associated with greater inefficiency because a hospital receives less credit for complex cases. The opposite is true for case mix adjusted workload. Thus, experience matters more when case mix adjusting is used and less when raw output figures are used, and the results for the coefficients associated with median case mix index are indicative of a “learning by doing” effect.
Case mix index standard deviation is only significant (p<.01) in Dist. However, in all specifications, the negative sign indicates that greater variation of case mix index is correlated with less inefficiency. This result also seems counterintuitive, but may also be related to the correlation of higher median case mix index with greater efficiency, since higher standard deviations are associated with higher case mix indexes (correlation = .3789). Experience matters.

In all four specifications, the percentage of surgical DRGs is inversely correlated with inefficiency, and this relationship is significant at the p<.05 level. Perhaps this reflects a greater certainty associated with diagnosis and treatment of surgical cases over medical cases. Finally, a higher percentage of MDC 5 work is associated with less inefficiency (p<.001). The top two DRGs in MDC 5 are 143 – Congestive Heart Failure and 127 – Chest Pain. They represent 40% of military MDC 5 workload and 27.6% of civilian workload. Even though bypass surgery is included in MDC 5, it appears that the large proportion of less complex DRGs 143 and 127 dominate and create an overall positive association with efficiency for MDC 5 workload. The results for MDC 14 were only significant in the Dist specification, but the relationship was not what was expected: a greater percentage of MDC 14 was hypothesized to correlate with less inefficiency, as a result of “learning by doing”. However, it is likely that much of the effects of MDC 14 workload volume have been captured in the variables for age and gender.

**Quality:** JCAHO accreditation is significantly correlated (p<.05) with greater efficiency in three specifications. Membership in the Council of Teaching Hospitals is significantly correlated (p<.01) with greater inefficiency in only specification Dist. The only time an outcome quality measure is significantly correlated with inefficiency is the percentage of actual over expected length of stay in specification Dist: greater inefficiency is significantly (p<.05) correlated with a greater (positive) percentage difference between actual and expected length of stay. Process variables, as mentioned in Chapter 7, were
excluded from the final model. The lack of significant findings may reflect the multidimensional nature of quality and difficulty in measuring it. It may also support the idea that Pauly’s (2004) theories on the effects of competition on quality and price in healthcare also apply to quality and efficiency. Tradeoffs occur when a firm operates at the production possibilities frontier. For firms not operating at the frontier (i.e. less than 100% efficient), a consistent relationship between quality and efficiency cannot exist.

**Physician Characteristics:** Having 100%-credentialed physicians (i.e. no physician employees) is associated with less inefficiency. This correlation is significant in three specifications, although at varying levels. This finding suggests that the incentives of individual credentialed physicians to work efficiently translate into greater efficiency for a hospital than a more traditional employee relationship between hospital and doctor. The number of malpractice claims per capita in the state where the hospital is located had no significant association with inefficiency in any model. The coefficient sign was consistently negative in all specifications, indicating the possibility that riskier malpractice environments encourage efficiency. However, the state-level identification of this variable may not sufficiently reflect the actual relationship between malpractice risk faced by physicians and efficiency.

**Patient Characteristics:** In all four specifications, significant correlations are present between efficiency and both average patient age and percentage of female patients. An older patient base is associated with greater inefficiency, and a higher percentage of female patients is associated with less inefficiency. These two associations may help explain why the percentage of MDC 14 (obstetric) workload discussed earlier was not found to be significant. Positive signs for both the percentage of nonwhite patients and the percentage of self/no-pay patients indicate a direct relationship between each of them and inefficiency. Each reaches a significant level (p<.05) in one of the four
specifications. The underlying meaning of these associations requires further analysis and may have substantial policy considerations.

**Technical Efficiency Scores**

Technical efficiency is defined as:

\[ TE_i = \frac{y_i}{y_i^*} = \exp(-u_i) \]

Where \( y_i^* \) is the production frontier – the maximum output given the inputs for each hospital. In the first three specifications, \( TE_i \) represents aggregated total workload (inpatient dispositions and outpatient visits) achieved vs. total workload possible. In DIST, the fourth specification,

\[ TE_i = D_0 = \exp(-u_i) \]

In all specifications, total average technical efficiency would be:

\[ \overline{TE} = \frac{1}{I} \sum_{i=1}^{I} TE_i \] for each hospital \( i, i = 1,2 \ldots I \) (Coelli et al., 2005)

And, Average Inefficiency = 1 - \( \overline{TE} \).
Table 8.4: Average Technical Inefficiency by Ownership Type

<table>
<thead>
<tr>
<th>Ownership Type</th>
<th>Control CMI</th>
<th>Raw CMI</th>
<th>CMI In Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>State/Local Government</td>
<td>Mean 34.47%</td>
<td>29.66%</td>
<td>30.15% 29.41%</td>
</tr>
<tr>
<td></td>
<td>S.D. 19.33%</td>
<td>18.82%</td>
<td>18.50% 18.54%</td>
</tr>
<tr>
<td>Not-For-Profit</td>
<td>Mean 20.26%</td>
<td>18.70%</td>
<td>17.72% 17.49%</td>
</tr>
<tr>
<td></td>
<td>S.D. 15.12%</td>
<td>13.98%</td>
<td>13.83% 13.80%</td>
</tr>
<tr>
<td>For-Profit</td>
<td>Mean 20.01%</td>
<td>19.60%</td>
<td>18.83% 20.90%</td>
</tr>
<tr>
<td></td>
<td>S.D. 17.15%</td>
<td>17.84%</td>
<td>16.64% 17.34%</td>
</tr>
<tr>
<td>Military</td>
<td>Mean 37.10%</td>
<td>10.03%</td>
<td>11.69% 14.39%</td>
</tr>
<tr>
<td></td>
<td>S.D. 15.71%</td>
<td>4.43%</td>
<td>4.55%  6.15%</td>
</tr>
<tr>
<td>Civilian</td>
<td>Mean 23.09%</td>
<td>21.04%</td>
<td>20.38% 20.36%</td>
</tr>
<tr>
<td></td>
<td>S.D. 17.28%</td>
<td>16.19%</td>
<td>16.02% 16.05%</td>
</tr>
<tr>
<td>Total</td>
<td>Mean 24.02%</td>
<td>20.31%</td>
<td>19.81% 19.97%</td>
</tr>
<tr>
<td></td>
<td>S.D. 17.52%</td>
<td>15.93%</td>
<td>15.68% 15.66%</td>
</tr>
</tbody>
</table>

As shown in Table 8.4 above, total average technical inefficiency ranges from 19.8% to 24.0% across the four models. For military hospitals, the average inefficiency varies widely, from 10.0% in specification Raw to 37.1% in specification CMI. Inefficiency of state and local hospitals varies across specifications from 29.4% to 34.5%. Inefficiency of not-for-profits varies across specifications from 17.5% to 20.3%. Inefficiency of for-profits varies across specifications from 18.8% to 20.9%. Thus, the large variation for military ownership does not occur with any of the other ownership types. This reflects the influence of the much larger volume of outpatient work for military hospitals in conjunction with the importance of weighting method chosen for it. In CMI, outpatient workload is weighted based on the median of the hospital’s inpatient case mix index. In the other models, outpatient workload is not weighted. The large volume of outpatient workload performed by military facilities makes estimation of their overall efficiency
sensitive to the weighting method. The comparatively small number of military hospitals likely amplifies this effect.

Differences between military and civilian hospital inefficiency are significant (p<.05) in all specifications: military hospitals are less efficient than civilian hospitals on average in CMI, but more efficient in Raw, CMI_In, and Dist. For civilian hospitals overall, the average inefficiency varies from 20.4% to 23.1% across specifications. Recall, however, that the coefficients of the variables representing different ownership types are not statistically significant in the SFA estimation discussed previously. This means that factors other than ownership – such as complexity and scope of work performed and patient demographics – are the sources of these differences in overall inefficiency.

<table>
<thead>
<tr>
<th>Size</th>
<th>State/Local Government</th>
<th>Not-For-Profit</th>
<th>For-Profit</th>
<th>Military</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥50 Beds</td>
<td>24.92%</td>
<td>16.36%</td>
<td>15.88%</td>
<td>12.04%</td>
<td>17.22%</td>
</tr>
<tr>
<td>&lt;50 Beds</td>
<td>36.10%</td>
<td>22.41%</td>
<td>29.81%</td>
<td>11.03%</td>
<td>26.57%</td>
</tr>
<tr>
<td>Total</td>
<td>30.15%</td>
<td>17.72%</td>
<td>18.83%</td>
<td>11.69%</td>
<td>19.81%</td>
</tr>
</tbody>
</table>

Table 8.5 uses CMI_In, the specification with the highest log likelihood. It indicates that large hospitals are more efficient than small ones (with the exception of military hospitals), yet the coefficient for the variable representing hospitals with less than 50 beds is significant and negative in the SFA estimation results of Table 8.2, indicating that smaller size is correlated with greater efficiency. Other variables, such as case mix index must be creating the observed differences. For example, t-tests for equality show that on average, small hospitals have significantly older patients (55.1 vs. 49.1) and perform significantly fewer MDC 5 procedures (15.5% vs. 12.9%). Both of these characteristics
were shown to be associated with greater inefficiency in the SFA estimation results of Table 8.2.

Table 8.6: Average Technical Inefficiency by Military Service

<table>
<thead>
<tr>
<th></th>
<th>CMI</th>
<th>Raw</th>
<th>CMI_In</th>
<th>Dist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Army</td>
<td>33.14%</td>
<td>8.93%</td>
<td>9.60%</td>
<td>11.85%</td>
</tr>
<tr>
<td>Navy</td>
<td>43.41%</td>
<td>9.62%</td>
<td>12.63%</td>
<td>16.46%</td>
</tr>
<tr>
<td>Air Force</td>
<td>37.69%</td>
<td>13.33%</td>
<td>15.47%</td>
<td>17.63%</td>
</tr>
</tbody>
</table>

Past research on efficiency of military hospitals tested the hypothesis that institutional differences between services might affect performance. Table 8.6 above displays average technical inefficiency by service for the four specifications. T-tests for significance of each service compared to the average of the remaining two found inefficiency of Air Force hospitals to be significantly higher (either p<.05 or p<.01) than Navy and Army hospitals combined in the Raw, CMI_In and Dist specifications. These tests also found inefficiency of Army hospitals to be significantly lower (either p<.05 or p<.01) than Navy and Air Force hospitals combined in the CMI, CMI_In and Dist specifications. Efficiency of Navy hospitals was not significantly different from the average efficiency of Army and Air Force hospitals in any specification. Given these results, it appears that Army hospitals are most efficient. However, this analysis cannot conclude that Army control is more efficient, ceteris paribus. Investigation of possible variables that might explain these differences would require additional SFA estimation, most likely using a panel design to obtain more observations.
Table 8.7: Average Technical Inefficiency by Ownership and Median Case Mix Index

<table>
<thead>
<tr>
<th>CMI</th>
<th>State/Local Government</th>
<th>Not-For-Profit</th>
<th>For-Profit</th>
<th>Military</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMI&lt;1</td>
<td>31.05%</td>
<td>18.72%</td>
<td>20.14%</td>
<td>11.89%</td>
<td>20.84%</td>
</tr>
<tr>
<td>CMI≥1</td>
<td>18.45%</td>
<td>10.15%</td>
<td>13.56%</td>
<td>7.49%</td>
<td>11.83%</td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30.15%</td>
<td>17.72%</td>
<td>18.83%</td>
<td>11.69%</td>
<td>19.81%</td>
</tr>
</tbody>
</table>

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 8.7 examines differences in technical efficiency based on median values of hospital case mix index using specification CMI_In. For all ownership categories, median case mix indices greater than 1 are correlated with lower levels of inefficiency. (Table 8.2 indicates similar results based on the coefficient for median case mix index (-0.7101), although the results were not significant.) Differences between the two categories are significant for state/local government and not-for-profit hospitals. Learning by doing, in which experience brings about improvements in performance, may be the explanation for this finding. Although not presented, specification Raw, where output is unadjusted, produces contradictory results. Hospitals with median case mix indexes less than one are either more efficient or are no less efficient than those with median case mix indexes greater than or equal to one.

Table 8.8: Average Technical Inefficiency by Volume of MDC 14 Workload

<table>
<thead>
<tr>
<th></th>
<th>State/Local Government</th>
<th>Not-For-Profit</th>
<th>For-Profit</th>
<th>Military</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥Average</td>
<td>25.58%</td>
<td>14.79%</td>
<td>13.60%</td>
<td>12.01%</td>
<td>15.90%</td>
</tr>
<tr>
<td>&lt;Average</td>
<td>33.16%</td>
<td>21.07%</td>
<td>23.26%</td>
<td>9.69%</td>
<td>24.03%</td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>***</td>
<td>**</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>30.15%</td>
<td>17.72%</td>
<td>18.83%</td>
<td>11.69%</td>
<td>19.81%</td>
</tr>
</tbody>
</table>

legend: * p<0.05; ** p<0.01; *** p<0.001
Table 8.8 reflects differences in average technical inefficiency based on whether a hospital performs either less or average or greater percentages of MDC 14 workload using the non-distance specification with the highest log-likelihood, $CMI_{In}$. It was hypothesized that higher percentages of obstetric work might correlate with greater overall technical efficiency due to an “assembly line” effect of procedural familiarity from high volume and “learning by doing”. (The related coefficients in Table 8.2 indicated a significant correlation between volume of MDC 14 work and efficiency was only present in specification $Dist$, and the correlation was opposite – greater obstetric volume was correlated with more inefficiency.) However, greater proportions of both young and female patients are also significantly correlated with greater efficiency, and these relationships are likely the reason for the results in Table 8.8 above. It indicates that hospitals with higher percentages of MDC 14 workload are indeed more efficient, and the difference in above and below average obstetric-volume workloads is significant ($p<.001$) in total. Results of $t$-tests for significance are reflected for each ownership category in the table. Military hospitals are the exception, but the difference is not significant. Furthermore, only five of the 44 total military hospitals actually have lower-than-average obstetric workload.

**Partitioning of Results**

Zuckerman, Hadley, and Iezzoni (1994) and Folland and Hofler (2001) investigated the possibility that different types of hospitals might face different production possibility frontiers, which could lessen the relevance of comparisons between those types. Because the primary research question of this dissertation concerns differences in efficiency by ownership category, and some researchers contend that military and civilian hospitals are too different for meaningful comparisons, an analysis of pooled vs. partitioned data seems appropriate. The small number of military hospitals necessitated a reduction in the total number of variables to reach convergence. Convergence problems also required TOTWORK to be case-mix-weighted using average (as opposed to the median) case mix index. Use of a panel data design would have
increased the number of observations and possibly alleviated some of these problems. Table 8.9 below shows the civilian/military partitioned results for a more parsimonious model that uses a Cobb-Douglas production function and fewer inefficiency variables.

### Table 8.9: Partitioned Results

<table>
<thead>
<tr>
<th></th>
<th>CDAll</th>
<th>CDMil</th>
<th>CDCiv</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>668</td>
<td>44</td>
<td>624</td>
</tr>
<tr>
<td><strong>Log Likelihood</strong></td>
<td>-149.065</td>
<td>33.286</td>
<td>-151.852</td>
</tr>
</tbody>
</table>

**Intotwork**

<table>
<thead>
<tr>
<th></th>
<th>CDAll</th>
<th>CDMil</th>
<th>CDCiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Hospital Beds)</td>
<td>0.263</td>
<td>0.0015</td>
<td>0.2626</td>
</tr>
<tr>
<td>ln(RN FTEs)</td>
<td>0.5017</td>
<td>0.5333</td>
<td>0.49</td>
</tr>
<tr>
<td>ln(Other FTEs)</td>
<td>0.2775</td>
<td>0.5255</td>
<td>0.2888</td>
</tr>
<tr>
<td>Constant</td>
<td>3.6446</td>
<td>2.911</td>
<td>3.6446</td>
</tr>
</tbody>
</table>

**mu**

<table>
<thead>
<tr>
<th></th>
<th>CDAll</th>
<th>CDMil</th>
<th>CDCiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>JCAHO</td>
<td>-0.2109</td>
<td>-0.2452</td>
<td>-0.2226</td>
</tr>
<tr>
<td>&lt;50 Beds</td>
<td>-0.2441</td>
<td>-0.0021</td>
<td>-0.2565</td>
</tr>
<tr>
<td>MDC 14 % of Workload</td>
<td>-0.3147</td>
<td>1.3426</td>
<td>-1.1631</td>
</tr>
<tr>
<td>MDC 5 % of Workload</td>
<td>-6.0889</td>
<td>-4.4693</td>
<td>-6.084</td>
</tr>
<tr>
<td>PCTSURGERY</td>
<td>-2.2822</td>
<td>-2.235</td>
<td>-2.0649</td>
</tr>
<tr>
<td>Average Patient Age (mean)</td>
<td>0.0332</td>
<td>0.0283</td>
<td>0.029</td>
</tr>
<tr>
<td>PCTNONWHITE</td>
<td>0.2636</td>
<td>0.0173</td>
<td>0.2319</td>
</tr>
<tr>
<td>% Diff from Expected Mort</td>
<td>-0.0328</td>
<td>-0.0975</td>
<td>-0.0267</td>
</tr>
<tr>
<td>% Diff from Expected LOS</td>
<td>0.112</td>
<td>-1.286</td>
<td>0.1124</td>
</tr>
<tr>
<td>Mean DRG Weight</td>
<td>-1.669</td>
<td>0.2038</td>
<td>-1.6713</td>
</tr>
<tr>
<td>Constant</td>
<td>1.4351</td>
<td>-0.157</td>
<td>1.7395</td>
</tr>
</tbody>
</table>

**ilgtgamma**

<table>
<thead>
<tr>
<th></th>
<th>CDAll</th>
<th>CDMil</th>
<th>CDCiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.1549</td>
<td>1.9646</td>
<td>1.1389</td>
</tr>
</tbody>
</table>

**Insigma2**

<table>
<thead>
<tr>
<th></th>
<th>CDAll</th>
<th>CDMil</th>
<th>CDCiv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-1.6602</td>
<td>-3.9527</td>
<td>-1.6365</td>
</tr>
</tbody>
</table>

**legend:** * p<0.05; ** p<0.01; *** p<0.001
The results above indicate that there are, in fact, some differences in the production function of military and civilian hospitals. Beds are not a significant input in the production function of military hospitals, and Other FTEs, while significant (p<.001) in the pooled and partitioned models, has a much larger coefficient for military facilities. Coefficients of inefficiency variables are, for the most part, similar. The most significant differences occur for the mean DRG weight (case mix index), percentage difference in actual and estimated length of stay, and percentage of MDC 14 workload. Average case mix index (mean DRG weight) is not significantly correlated with inefficiency for military hospitals, yet a higher case mix index is significantly correlated with less inefficiency for civilian hospitals. For military hospitals, the more that actual length of stay exceeds estimated length of stay, the greater the efficiency. For civilian hospitals, this relationship is reversed, and in both cases, the results were significant (p<.05).

Interestingly, for military hospitals, a higher percentage of MDC 14 (obstetric) workload is significantly correlated with greater inefficiency. For civilian hospitals, a higher volume of MDC 14 workload, while not significant, was associated with less inefficiency. These results provide little support for “learning by doing” with respect to obstetrics.

What about estimated efficiency? Average military inefficiency is estimated to be 15.7% in the partitioned model and 23.2% in the pooled model. Average civilian inefficiency is estimated to be 26.7% in the partitioned model and 26.6% in the pooled model. Therefore, pooling does appear to affect the results. However, regardless of pooled or partitioned results, one can conclude that military hospitals are at worst, no less technically efficient than civilian hospitals.

**Heteroscedasticity**

Problems with efficiency of estimation can arise when the variance of the dependent variable varies across the data. Heteroscedasticity affects standard errors, and thus determinations of significance of a given variable. Standard tests for heteroscedasticity
following a linear regression are not available for frontier maximum likelihood estimation. However, one can visually inspect scatter plots of the residuals and fitted values (y-hat) for patterns in the data. Because the distribution of u_i has been modeled to be dependent on the explanatory variables, heteroscedasticity of u_i would be expected. However, it is not expected for v_i. Figure 8.10 depicts a scatter plot of predicted values of TOTWORK against values of v_i for CMI_In, the specification with the highest log-likelihood value.

![Figure 8.10: Scatter Plot of Residuals](image)

While some outliers appear to affect variance in the mid-range of y-hat, no strong pattern of heteroscedasticity is apparent.
Stata allows for explicit modeling of variables thought to influence the variance of both \( u_i \) and \( v_i \), but an assumption of a half-normal inefficiency error term is required, and thus were not used in the primary model specifications. However, for \( RAW \) and \( CMI\_In \), models using the same explanatory variables discussed in this chapter as explanatory variables for the variance of \( u_i \) and size (i.e. HOSPBD) as the sole explanatory variable for \( v_i \) were run. While estimates of overall efficiency were lower, the significance of variable coefficients did not change markedly, especially with respect to ownership. This result, when considered with the scatter plot above, suggests that heteroscedasticity is not a significant problem.

**Summary**

This chapter presented results of the Stochastic Frontier Analysis of hospital technical efficiency. Four models were presented. \( CMI \) applied the median case mix index derived from actual inpatient dispositions to both inpatient dispositions and outpatient visits to account for heterogeneity in work performed. \( Raw \) applied no case mix index at all. \( CMI\_In \) indexed inpatient dispositions only using the median case mix index. Of these three specifications, \( CMI\_IN \) had the highest log likelihood, indicating better fit. This is congruent with both the notion that heterogeneity of work affects output, and the notion that adjusting outpatient work using inpatient case mix index may not be optimal, especially when the relative volume of outpatient work is high. \( Dist \) applied distance function methods to the model indexing for inpatient dispositions only using the average case mix index, and this eliminated the requirement for aggregation of outputs. Ownership, the focus of the dissertation, was not significantly correlated with efficiency in any specification. Estimates of inefficiency were consistent both in total (~20% - 24%) and for each ownership category across the four specifications, with one key exception. Inefficiency of military hospitals was substantially higher in specification \( CMI \) than in any of the other three. This reflects the importance of controlling for heterogeneity in outpatient workload. Assuming a hospital’s outpatient visits are similar in complexity to its inpatient dispositions (as in \( CMI \)) may be a somewhat
stringent assumption for military hospitals, given their relatively low inpatient complexity when compared to the other three types of ownership. However, even in this model, there are no significant correlations between any ownership category and efficiency. Other factors found to be significantly related to hospital efficiency in at least three of the four specifications included average patient age, gender, percentage of surgical inpatient workload, percentage of MDC 5 work, 100% credentialed MDs, and JCAHO accreditation. The effects of these other factors influenced total inefficiency by ownership so that in specifications RAW, CMI_In, and Dist, military hospitals exhibit significantly greater efficiency than civilian hospitals. However, in specification CMI, military hospitals exhibit significantly less efficiency than civilian hospitals. Analysis of pooled and partitioned results indicated that military hospitals are, in fact, different from civilian ones. However, these differences are not insurmountable for comparison purposes. In no way can mean efficiency scores be interpreted as indicating military hospitals are less efficient than civilian ones, and in some cases they could be considered more efficient. Chapter 9 will now provide validation of these results using the other key frontier method discussed in Chapter 2, Data Envelopment Analysis.
While the primary results of this dissertation presented in Chapter 8 were developed using SFA, validation using another method will increase their reliability. Use of both DEA and SFA has been recommended by Coelli and colleagues (2005) and implemented by Jacobs (2001) and Linna (1998). Validation with DEA allows for exploitation of its advantages – in particular, the ability to accommodate multiple outputs without loss of information due to aggregation – while tempering results with the key disadvantages of the method – i.e. no recognition of stochastic error and the inability to generate standard errors for estimates.

Data Envelopment Analysis

As previously mentioned, DEA is attributed to Charnes, Cooper, and Rhodes’ 1978 work. Development of the general Data Envelopment Analysis (“DEA”) mathematical formulation, is depicted as follows:

Given $n$ DMUs, the $j$th DMU uses $m$ inputs as represented by:

Input vector $x_j = (x_{ij},...,x_{mj})$;

and produces $s$ outputs as represented by:

Output vector $y_j=(y_{ij},...,y_{sj})$

The comparison group (linear combination) identified (as discussed in Chapter 2) for a given DMU (DMU 0) is represented by:

Weight vector $\lambda = (\lambda_i,...,\lambda_n)$, where $\lambda_j$ is the weight of DMU $j$.

The comparison group must produce at least as much output as DMU 0:

$$\sum_{j=1}^{n} y_{rj} \lambda_j \geq y_{r0} \quad r = 1,...,s$$

The weighted comparison group can use no more than a fraction ($h_0$) of the inputs that DMU 0 uses, and $0 \leq h_0 \leq 1$:

$$\sum_{j=1}^{n} x_{ij} \lambda_j \geq h_0 x_{i0} \quad i = 1,...,m$$
The minimized value of $h_0 = \text{DMU 0's relative technical efficiency}$. The efficient comparison group minimizes $h_0$. The vector of optimal weights is comprised of the weights applied to each DMU in the comparison group. The linear program that executes these steps finds the optimal values of $h_0$ and $\lambda$ (Hollingsworth & Peacock, 2008)

Variables
The outputs used are the same as those used in the SFA model, without aggregation:

- Case-mix adjusted Inpatient Dispositions
- Case-mix adjusted Outpatient Visits

The inputs used are the same as used on the SFA model as well:

- Hospital Beds
- RN FTEs
- Other Non-MD FTEs

Descriptive statistics for these variables were discussed in Chapter 6.

Methods
All outputs and inputs were logged as in the SFA model. Three DEA specifications were completed – one with case mix adjusted inpatient dispositions and case mix adjusted outpatient visits ($DEACMI$), one with raw inpatient dispositions and raw outpatient visits ($DEARAW$), and one with only case mix adjusted dispositions ($DEACMI\_IN$). All three models used the terms derived in the translog SFA models – i.e. they included square- and cross-products of the inputs.
The software used to perform this DEA validation of the SFA results in Chapter 8 was Efficiency Measurement System (“EMS”) 1.3 developed by Holger Scheel. It is available at no cost to academic researchers at http://www.wiso.uni-dortmund.de/lsfg/or/scheel/ems/. EMS allows for the calculation of super-efficiency, a technique that calculates relative efficiency scores (over 1.0) for hospitals deemed 100% efficient. For purposes of deriving correlations between SFA and DEA specifications, super-efficiency was allowed to create a more granular efficiency ranking of all hospitals, not those deemed inefficient. Variable returns to scale were also permitted in all DEA specifications.

As presented in Chapter 7, the γ statistic may provide an estimate of how closely correlated efficiency scores obtained using SFA and DEA will be. If inefficiency is a very high percentage of total error, correlation with efficiency scores using DEA should be higher. As indicated in Chapter 8, γ ranged from 73.6% to 82.7%. While neither at nor very near 100%, these high values for γ suggest it is reasonable to expect moderately high correlation.

Results
To calculate average technical efficiency, each DEA model was re-run without super-efficiency so that no hospital achieved a score greater than 1. Average technical efficiency in specification DEARAW was 87.4%. In DEACMI, average technical efficiency was 87.2%. In DEACMI-IN, it was 87.9%. These results are both similar to each other and slightly higher than results obtained using SFA (76% - 80%), yet comparable. To assess correlation between methods, Pearson correlation coefficients were calculated based on rankings of each hospital’s efficiency score for SFA specifications MED, RAW, CMI-IN, and DIST, as well as for DEA specifications DEARAW and DEACMI, and DEACMI-IN. The results of these correlations are presented in Table 9.1.
Table 9.1: Ranking Correlation Matrix of SFA and DEA Specifications

<table>
<thead>
<tr>
<th></th>
<th>RAW</th>
<th>CMI-IN</th>
<th>CMI</th>
<th>DIST</th>
<th>DEACMI</th>
<th>DEARAW</th>
<th>DEACMI-IN</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAW</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMI-IN</td>
<td>0.8507</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMI</td>
<td>0.6801</td>
<td>0.9038</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIST</td>
<td>0.7981</td>
<td>0.9557</td>
<td>0.8741</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEACMI</td>
<td>0.4619</td>
<td>0.6535</td>
<td>0.5805</td>
<td>0.6856</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEARAW</td>
<td>0.6039</td>
<td>0.5654</td>
<td>0.3620</td>
<td>0.5824</td>
<td>0.8482</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>DEACMI-IN</td>
<td>0.5561</td>
<td>0.6479</td>
<td>0.4714</td>
<td>0.6643</td>
<td>0.9404</td>
<td>0.9459</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

The area shaded above represents correlations between SFA specifications and DEA specifications. Overall, these scores indicate only moderate positive correlation between DEA efficiency scores and SFA efficiency scores. However, correlations between similar models (i.e. DEARAW and RAW; CMI and DEACMI; CMI-IN and DEACMI-IN) are approximately .6 in each case. Recall that specification DIST allowed for explicit modeling of multiple outputs without aggregation, which is not true of standard SFA technical efficiency models and is usually considered a key disadvantage of them. Thus, one would expect somewhat higher correlations between DIST and the three DEA models, and the table above confirms this expectation. Correlations between DIST and DEACMI, as well as between DIST and DEACMI-IN are higher than for the other SFA models. Correlation between DIST and DEARAW is only second to the correlation between DEARAW and RAW. Correlations among DEA specifications are high, as indicated by the results in the last three columns (.84 - .95). Correlations marked SFA specifications range from marked to high, as indicated by the results in the top four rows (.68 - .96).

Correlation with Traditional Ratio Measures
As previously discussed, the preponderance of pre-frontier hospital efficiency studies reached the conclusion that for-profit organizations were more efficient than either not-
for-profit or government organizations. Typically, these conclusions were based on ratio analyses or standard OLS (Clarkson, 1972) (Herzlinger & Krasker, 1987).

Since the development of frontier techniques, efficiency scores are often obtained using only one method (Nayar & Ozcan, 2008) (Burgess & Wilson, 1998) (Rosko, 2001) (Rosko & Mutter, 2008) (Brown III, 2003). Some studies do compare results from DEA and SFA, however (Chirikos & Sear, 2000) (Linna, 1998) (Jacobs, 2001) (Van Fulton, 2005). Even fewer studies compare results obtained using a frontier technique (DEA or SFA) to results using a non-frontier technique (Hao & Pegels, 1994) (Thanassoulis, Boussofiane, & Dyson, 1996). Yet, such a comparison might yield information about whether frontier techniques provide additional information on efficiency, as posited by Fiorentino, et al., (2006) in their study of banking efficiency. These authors found moderate correlation between SFA and DEA results, but low correlation with accounting ratios of return on equity and return on assets, leading them to conclude that frontier techniques consider “alternative drivers of success and failure, such as market power or economic value maximization” better than standard accounting ratios (Fiorentino, Karmann, & Koetter, 2006). In a healthcare setting, Thanassoulis, Boussofiane, and Dyson (1996) found similar results in an analysis of perinatal care in England. Comparing results obtained from DEA and from analysis of 25 productivity ratios, these authors found: 1) low correlation of rankings based on each individual ratio; 2) low (but positive) correlation between DEA hospital rankings and rankings based on each productivity ratio; and 3) moderate correlation when DEA rankings were compared to aggregate ratio performance. Furthermore, correlations of DEA rankings and aggregate ratio rankings were sensitive to the method chosen to aggregate the ratios (i.e., mean vs. minimum rank achieved on a given ratio). The authors attributed these correlation results to the fact that DEA considers all modeled inputs and outputs simultaneously (Thanassoulis, Boussofiane, & Dyson, 1996).
To investigate how ownership might be evaluated using standard productivity ratios, two ratios – one for output per unit of capital (bed) and one for output per unit of labor (FTE) – will be analyzed. Linear regression with the ratio as the dependent variable and 1) the other inputs, and 2) the explanatory variables previously used as influencers of efficiency as dependent variables will be performed. Inpatient and outpatient workload are aggregated as in two of the previous specifications – *CMI* (both case mix adjusted) and *RAW* (neither case mix adjusted). The linear regression models are as follows:

1) \[
\text{Work/FTE} = \alpha + \beta_1(\text{HOSPBD}) + \beta_2(\text{CONTROLCAT NFP}) + \beta_3(\text{CONTROLCAT FP}) + \beta_4(\text{CONTROLCAT MIL}) + \beta_5(\text{HOSPMEDCMI}) + \beta_6(\text{HOSPSDCMI}) + \beta_7(\text{PCTSUG}) + \beta_9(\text{ERPERCENT}) + \beta_9(\text{PCTMD14}) + \beta_{10}(\text{PCTMD5}) + \beta_{11}(\text{HOSP AVERAGE}) + \beta_{12}(\text{PCT FEMALE}) + \beta_{13}(\text{PCT NONWHITE}) + \beta_{14}(\text{PCT SELF NOPAY}) + \beta_{15}(\text{PCT OVER MORT}) + \beta_{16}(\text{PCT OVER LOS}) + \beta_{17}(\text{JCAHO}) + \beta_{18}(\text{COTH}) + \beta_{19}(\text{SMALLHOSP}) + \beta_{20}(\text{HSA HHI}) + \beta_{21}(\text{SHARE HSA}) + \beta_{22}(\text{CLAIMS PER CAP2000}) + \beta_{23}(\text{CRED}) + \epsilon
\]

and

2) \[
\text{Work/Bed} = \alpha + \beta_1(\text{FTERN}) + \beta_2(\text{FTEOTH}) + \beta_3(\text{CONTROLCAT NFP}) + \beta_4(\text{CONTROLCAT FP}) + \beta_5(\text{CONTROLCAT MIL}) + \beta_6(\text{HOSPMEDCMI}) + \beta_7(\text{HOSPSDCMI}) + \beta_8(\text{PCTSUG}) + \beta_9(\text{ERPERCENT}) + \beta_{10}(\text{PCTMD14}) + \beta_{11}(\text{PCTMD5}) + \beta_{12}(\text{HOSP AVERAGE}) + \beta_{13}(\text{PCT FEMALE}) + \beta_{14}(\text{PCT NONWHITE}) + \beta_{15}(\text{PCT SELF NOPAY}) + \beta_{16}(\text{PCT OVER MORT}) + \beta_{17}(\text{PCT OVER LOS}) + \beta_{18}(\text{JCAHO}) + \beta_{19}(\text{COTH}) + \beta_{20}(\text{SMALLHOSP}) + \beta_{21}(\text{HSA HHI}) + \beta_{22}(\text{SHARE HSA}) + \beta_{23}(\text{CLAIMS PER CAP2000}) + \beta_{24}(\text{CRED}) + \epsilon
\]

Table 9.2 below provides results for the ownership coefficients from these regressions of each ratio for the different aggregation methods – *CMI* and *RAW*. 
Table 9.2: Ownership Coefficients Based on Regression Using Ratios

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Not-For-Profit</th>
<th>For-Profit</th>
<th>Military</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMI</td>
<td>Work/FTE</td>
<td>.86</td>
<td>2.37</td>
</tr>
<tr>
<td></td>
<td>Work/Bed</td>
<td>6.21</td>
<td>0.47</td>
</tr>
<tr>
<td>RAW</td>
<td>Work/FTE</td>
<td>1.08</td>
<td>3.15</td>
</tr>
<tr>
<td></td>
<td>Work/Bed</td>
<td>7.54</td>
<td>3.25</td>
</tr>
</tbody>
</table>

Legend: * p<0.05; ** p<0.01; *** p<0.001

Table 9.2 differs significantly from the relevant ownership coefficients in the overall SFA estimation results in Table 8.2. Recall that in both tables, state/local government hospitals are the reference group. Conclusions about overall efficiency based on either of these two productivity ratios would be different from conclusions based on SFA. Table 8.2 found only one significant difference between not-for-profit and state/local government ownership in specification DIST. In Table 9.2, however, military ownership (when compared to state/local government ownership) is significantly correlated with higher productivity (i.e. greater efficiency) based on the linear regression with total work per bed as the dependent variable, yet significantly correlated with lower efficiency based on the linear regression with total work per FTE as the dependent variable. Significant differences occur for all comparisons to state/local government with the exception of for-profit work/bed ratios. As in Chapter 8, assessment of whether significant differences exist between other categories requires further testing.
Table 9.3: Significance Tests for Differences By Ownership Category – Ratio Analysis

<table>
<thead>
<tr>
<th>Chi^{2}(1) Scores</th>
<th>CMI Bed</th>
<th>CMI FTE</th>
<th>Raw Bed</th>
<th>Raw FTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Military vs. For-Profit</td>
<td>10.55</td>
<td>46.21</td>
<td>24.62</td>
<td>20.03</td>
</tr>
<tr>
<td>For-Profit vs. Not-For-Profit</td>
<td>4.37</td>
<td>14.90</td>
<td>1.51</td>
<td>64.82</td>
</tr>
</tbody>
</table>

Table 9.3 reports results of Wald tests for equality between other ownership types as listed for comparisons not involving state and local hospitals. A chi^{2}(1) statistic of 3.841 or greater is required to be significant at the p<.05 level. When compared to this critical value, the resulting chi^{2} statistics with one degree of freedom in the table above show significant differences among all ownership combinations except the for-profit/not-for-profit comparison in the regression of raw output/bed. In either specification RAW or CMI, military hospitals are significantly more efficient in terms of output/bed and significantly less efficient in terms of output/FTE than for-profit, not-for-profit, or state/local government hospitals.

Table 9.4 provides Pearson correlations for individual hospital rankings based on the two ratios with SFA efficiency rankings and DEA rankings obtained using super-efficiency.
Table 9.4: Correlation of SFA, DEA, and Ratio Rankings – CMI and RAW

<table>
<thead>
<tr>
<th></th>
<th>CMI FTE Ratio</th>
<th>CMI Bed Ratio</th>
<th>SFACMI</th>
<th>DEACMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMI FTE Ratio</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMI Bed Ratio</td>
<td>0.342</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFACMI</td>
<td>0.814</td>
<td>0.623</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>DEACMI</td>
<td>0.450</td>
<td>0.547</td>
<td>0.581</td>
<td>1.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Raw FTE Ratio</th>
<th>Raw Bed Ratio</th>
<th>SFARAW</th>
<th>DEARAW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw FTE Ratio</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw Bed Ratio</td>
<td>0.313</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFARAW</td>
<td>0.705</td>
<td>0.706</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>DEARAW</td>
<td>0.364</td>
<td>0.517</td>
<td>0.604</td>
<td>1.000</td>
</tr>
</tbody>
</table>

In both cases, the correlation between the two ratios is lower than any correlation with SFA or DEA (.313 and .342). This is consistent with the findings of Thanassoulis, et al. (1996) discussed above. Individual ratios assess performance from a very narrow perspective, and as such, low correlations are expected. Correlations of DEA rankings to ratios are primarily low and moderate at best (.364 - .547). Interestingly, correlations of ratio rankings and SFA rankings are marked (.623 - .814).

Summary

This chapter attempted to validate the findings of Chapter 8 with respect to overall efficiency and the systematic differences attributed to ownership type using DEA. It also explored how conclusions regarding efficiency might differ when using standard productivity ratios as the basis for comparison. The analysis found overall efficiency scores using DEA (~87%) to be similar to those obtained using SFA (76% - 80%). Correlations of individual hospital rankings between methods were generally moderate. Of these moderate correlations, those between DIST – the SFA specification using a distance function – and all DEA specifications were the highest, reflecting the fact that none of these methods aggregated inpatient and outpatient workload.
Conclusions about the effects of ownership based on standard productivity ratios differed significantly from those based on frontier estimation. Ownership differences were significant in almost all cases. This finding is consistent with observations made in Chapter 2 regarding how conclusions regarding overall efficiency of not-for-profits and government hospitals when compared to for-profits changed in the literature as it progressed. Previously significant differences lessened with the introduction of frontier techniques.
Summary

This dissertation used Stochastic Frontier Analysis, a cross-sectional design, and four model specifications to investigate possible effects of ownership on hospital technical efficiency, with a special interest in military (federal) control. Results were validated using Data Envelopment Analysis. This dissertation discovered that, on average, hospital ownership – including that by the military – is not significantly correlated with either more or less technical efficiency in producing inpatient and outpatient workload. Results for the effects of case mix index were sensitive to the weighting method used for the dependent variable representing total output. When workload was not weighted based on complexity, case mix index was significantly correlated with greater inefficiency: when both inpatient and outpatient workload were weighted, case mix index was significantly correlated with less inefficiency. These results likely reflect the importance of skill and experience of personnel in the production process.

Factors found to have a significant correlation with efficiency in at least two of four specifications include 1) average patient age; 2) percentage of female patients; 3) having all credentialed (i.e. no employee) physicians; 4) percentage of MDC 5 work; 5) percentage of surgical DRGs; 6) JCAHO accreditation, and 7) having fewer than fifty beds. Characteristics of a hospital’s patient base do affect efficiency. Greater proportions of both young patients and female patients are significantly correlated with greater efficiency. The employment arrangement between hospitals and physicians affects efficiency. Facilities with all credentialed physicians exhibited significantly less technical efficiency than facilities with any employed physicians. Scope of work performed affects efficiency. Higher surgical workload and higher circulatory system workload are significantly correlated with greater efficiency. Accreditation by JCAHO – a structural measure of quality – was also correlated with greater efficiency. It was the only quality measure to produce statistically significant results. Finally, small hospitals
(i.e. fewer than fifty beds) were significantly more efficient than were those with at least fifty beds. Finally, this analysis found that both competition and volume of obstetric workload, factors hypothesized to have a significant correlation with inefficiency, were not statistically significant.

The results for the categorical ownership variable supports the theory that external factors affecting all hospitals (regulatory environment, e.g.) are more important to efficiency than are property rights and residual earnings associated with different types of ownership. The findings with respect to size indicate relative unimportance of economies of scale and support viewpoints put forth by Niskanen (1975), Hannsmann (1980), and Tullock (1977). The non-significant results for competition provide no support for the importance of either price (Leibenstein, 1966) (Spann, 1977) or non-price competition (Carroll, 1990) (Harris, 1977) in determining efficiency.

The presence of multiple missions in military hospitals (wartime readiness, education, and research) is one frequently cited reason why military hospitals might face a different production possibilities frontier, rendering comparisons with civilian hospitals inappropriate. In theory, multiple, often competing missions should lessen efficiency of patient care due to fragmented allocation of resources and time. Partitioning analysis, estimating civilian and military hospital efficiency separately and comparing those results to the combined efficiency estimation, revealed that military hospitals are indeed different. However, overall inefficiency scores were comparable to those derived in the four pooled specifications, and results of significance for variable coefficients were similar. These partitioned findings, when added to the non-significant findings on ownership, indicate that multiple missions, as well as other structural military/civilian differences, might not be critical in assessing overall technical efficiency.

---

6 Other reasons were discussed in Chapter 3.
Finally, even though average military efficiency was substantially lower in the most conservative specification where outpatient visits were weighted using the inpatient median case mix index, significant differences in the ownership efficiency coefficients still did not materialize.

Limitations
As with any study, there are limitations with this dissertation that should be noted.

1. Aggregation: The output of a hospital is heterogeneous, yet stochastic frontier analyses of technical efficiency generally allow for only one dependent variable/output. Distance functions, as employed in Ferrari (2006), alleviate some concerns about aggregation, but it can still be problematic. As Bradford et al. (2001) pointed out, national hospital-level studies necessitate some degree of aggregation due to the complexity of services provided and this can result in bias. Dor (1994) noted that SFA studies view the firm as one cohesive unit and thereby ignore the fact that firms are comprised of people with different incentives and/or disincentives for making decisions. This analysis addressed potential aggregation problems by weighting outpatient and inpatient work using their relative average unit costs, by considering case-mix adjusted inpatient and outpatient workload in different ways, and by examining a distance function in one specification.

2. Cross-Sectional Design: The cross-sectional design used here does not allow analysis of trends over time. It may be that the increased wartime operational tempo for military personnel has forced greater efficiency in caring for dependents of all types. A panel study might have been able to detect such a trend. A panel design would also allow for more observations and thus would contend with the criticisms of cross-sectional design, namely that it estimates inefficiency more inconsistently (Linna, 1998), that it is more likely to produce biased coefficient estimates, and that it requires stricter assumptions than a panel design (Dor, 1994). More observations
would also improve estimation of partitioned military data, which only had 44 observations in this analysis.

3. Data Comparability: This analysis merged well-established sources of data for civilian hospitals (AHA Survey, HCUP NIS, e.g.) with data on equivalent data elements (dispositions, DRGs, etc.) from military sources. However, it is possible that there are systematic military/civilian differences in the recording of data for these administrative databases. For example, accuracy of monthly timesheets filled out by military hospital staff has been a problem since neither hospital reimbursement nor individual salary is dependent on this information. The extent of this problem in civilian facilities is unknown. Additionally, this analysis combined data on a calendar year basis from the NIS and SIDR with data on a fiscal year basis from the AHA Annual Survey. Not all hospitals provided the fiscal year basis for their survey responses. However, many hospitals have fiscal years that are calendar years, and others have fiscal years beginning in October – only three months’ deviation from a calendar year.

4. Lack of outpatient workload descriptors: Descriptors for inpatients (age, gender, race, and insurance) taken from the HCUP NIS were assumed to be appropriate for outpatients as well. Coefficient estimates for these variables may be biased to the extent that average outpatients differ significantly from average inpatients. In addition, as discussed in Chapter 5 and applied in Chapter 8, model specification CMI used a broad assumption that the overall complexity of outpatient workload was comparable to that of inpatient workload due to military/civilian data discrepancies. Other specifications made no case mix adjustment to outpatient workload. The results in Chapter 8 found that overall results with respect to the relative efficiency of military hospitals were sensitive to changes in this assumption. Coefficient estimates for case mix variables may be biased to the extent that the complexity of outpatient work differs significantly from that of inpatient work, and that this complexity varies by ownership.
5. Omitted Variables: A production function has relatively few independent (input) variables (typically capital and labor). For an operation as complex as healthcare, the specified production function might omit important sub-categories of input, such as the use of paraprofessionals. Omitted variables result in specification errors that are likely to confound efficiency estimates (Rosko & Mutter, 2008). Dor (1994) discussed the fact that omitted variables might not only bias the coefficients of model variables, but also the estimated inefficiency scores in SFA. This analysis used production function variables to represent capital and labor that were consistent with previous literature.

6. Endogeneity: Pauly (1987) discussed the possibility that the decision to operate on a for-profit basis is endogenous. For-profit hospitals may choose the local market in which they will operate based on likelihood of profitability, and this could affect efficiency. While this could bias comparisons between for-profit and not-for-profit hospitals, it does not lessen the value of any comparisons to military facilities. Furthermore, the direction of likely bias from such endogeneity is clear and thus can be acknowledged.

7. Selection Bias: It is possible – but not likely – that patients choose hospitals based on their ownership type in the civilian world. This possibility is less likely with military hospitals since non-beneficiaries cannot choose to receive care in a military hospital, and active duty personnel only receive care in civilian facilities for emergencies or special cases. Thus, the relevance of military/civilian comparisons should not be weakened by selection bias.

8. Incomplete modeling of physician malpractice: The malpractice variable created using the National Practitioner Database only captured information at the state level. Furthermore, it only captured cases with a monetary settlement with respect
to the physician: hospital settlements and those without a monetary component are excluded.

**Recommendations for Future Research**

This dissertation breaks new ground in efficiency research by directly comparing efficiency of military and civilian hospitals. It can serve as the foundation for further research in the following areas:

1. **Effects of personnel rotation policies (turnover):** One of the somewhat unusual characteristics of active duty personnel management is the standard rotation policy. Military personnel change duty stations (i.e. hospitals) on average every three years. Personnel rotation represents the equivalent of civilian personnel turnover in a hospital. Frequent turnover is thought to have deleterious effects on organizational performance. A frontier study designed specifically to investigate the effects of this policy would be valuable for policymakers of military and civilian hospitals alike.

2. **Efficiency as an independent variable:** This dissertation investigated overall hospital efficiency levels and the factors thought to influence them. However, efficiency is really a means to an end: hospitals desire improved efficiency in order to reduce costs, improve quality, etc. While use of efficiency as an independent variable was an approach infrequently found in the literature, McKay and Deily (2008) provided a noteworthy exception. They investigated the potential effects of inefficiency on outcome measures of quality. Assuming adequate definition and measurement of quality, a similarly designed study would be a logical follow-on study to this dissertation, which quantified efficiency.

3. **Clinic-level unit of observation:** Hospital-level analyses must deal with an extremely complex production process of inpatient and outpatient work, and heterogeneity further complicates this process. Given the previously discussed concerns of military-to-civilian comparisons, focusing analysis on a specific clinical area of medical care (such as obstetrics) with potentially more homogeneous patients and
procedures across ownership types might serve as a means of validating the results of this dissertation.

4. Comparison of performance differences between levels of government: Literature revealed well-established theories on how public/private and not-for-profit/for-profit comparisons might turn out. Theories on how control by different levels of government might affect performance are much less prevalent. A direct comparison of federal, state, local, and district control could provide useful information for government healthcare policymakers and add to the body of literature on how federalism operates in healthcare.

5. Re-evaluation using APCs to weight outpatient visits: The divergent results produced by different workload weighting assumptions on efficiency scores of military hospitals, which performed significantly more outpatient workload, confirm the importance of controlling for complexity of outpatient visits. APCs, codes for resource intensity of outpatient workload that are similar to inpatient DRGs, may improve comparability of efficiency as their use in military facilities becomes more widespread.

6. More thorough modeling of quality: As discussed, quality in healthcare is an elusive concept to define and measure. This study included structural, process, and outcome measures of efficiency, but process measures proved troublesome – primarily due to reduction in sample size and the fact that hospitals still self-select whether to report quality data. A panel design might be one partial remedy, as reporting of quality becomes more widespread (and even required) in both military and civilian settings. Revisiting this study with more extensive modeling of quality could produce definitive results regarding the relationship of quality and efficiency.
Conclusions

Rapidly increasing healthcare costs continue to generate growing interest in the measurement and improvement of efficiency. Furthermore, the ever-increasing share of government’s role in healthcare engenders concern as to whether publicly managed healthcare facilities can function as efficiently (and effectively) as private ones. This dissertation examined technical efficiency – the most elemental type of economic efficiency – across ownership types, focusing on military hospitals as the representative of federally controlled facilities, an approach not found elsewhere in the literature. Controlling for characteristics of the patient base, characteristics of physicians’ practice environment, quality, level of competition, and scope of work, this dissertation found no significant relationship between ownership and efficiency. Military medical leaders face the difficult position of preparing military personnel for deployment, caring for wounded warriors returning from battle, and providing day-to-day medical care for active duty personnel, their family members, and retirees – all while being morally responsible to citizens and fiscally responsible to taxpayers. These results should be both informative and reassuring to military leaders and citizens alike.


Tricare Management Activity. (2008). *TRICARE; Civilian Health and Medical Program of the Uniformed Services (CHAMPUS); Fiscal Year 2009 Diagnosis Related Group (DRG) Updates*. 73 Federal Register 212.


VITA
Linda Gail Kimsey

Date of Birth: August 19, 1964

Place of Birth: Cleveland, TN

Degrees: B.S. Accounting, University of Tennessee, 1986

Positions Held: Action Officer for Manpower and Medical Programs, Office of the Chief of Naval Operations N80 Navy Programming Division

Medical Programs Analyst, Office of the Chief of Naval Operations N931 Navy Medical Resources, Plans and Policy Division

Comptroller, Naval Dental Center Camp Lejeune

Scholastic Honors: Top Business School Graduate, University of TN, Summer 1986

Top M.H.A Graduate, University of KY, 2001