University of Fort Hare
Together in Excellence

PROFIT INCENTIVES AND TECHNICAL EFFICIENCY IN THE PROVISION OF HEALTH CARE IN ZIMBABWE: AN APPLICATION OF DATA ENVELOPMENT ANALYSIS & ECONOMETRIC METHODS.

BY

MAREDZA ANDREW

A DISSERTATION SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS OF THE MASTER OF COMMERCE DEGREE IN ECONOMICS
NOVEMBER 2009

DEPARTMENT OF ECONOMICS

FACULTY OF MANAGEMENT & COMMERCE

UNIVERSITY OF FORT HARE

SOUTH AFRICA

DECLARATION AND COPYRIGHT

I, the undersigned Andrew Maredza student number 200808288, do hereby declare that this dissertation is my own original work and that it has never been presented to another institution for the award of a degree.

Signature

……………………………..

Date

…………/…………/……….

DEDICATION

This dissertation is dedicated to my beloved wife Vongai.

ACKNOWLEDGEMENTS

First and foremost, I am very grateful to the Almighty God for his grace and strength that has made all things possible. I would like to express my profound gratitude to my supervisor, Professor. M. Ncube whose incisive analysis and enlightening guidance was of paramount value.

The unquantifiable financial and moral support of my sponsor-The Zimbabwean Presidential Scholarship cannot go without mention, not forgetting the valuable comments and suggestions from many students at University of Fort Hare.

I would like also to thank the administrative support of the Department of Economics for their support in various capacities.

My family and friends also contributed to the successful completion of this dissertation in more ways than would be appropriate to list here. Special mention goes to the saints in Alice for their unceasing prayers and love for me, without whose spiritual support I would not have attempted or completed this dissertation.

Thank you to everyone who has helped to make this dissertation a success.

ABSTRACT

This study examines issues surrounding efficiency in the Zimbabwean health sector with specific emphasis on for-profit hospitals in order to find out whether they are significantly more efficient than non-profit hospitals. The study attempts to explore the significance of profit incentives on efficiency. This study uses the Data Envelopment Analysis (DEA) methodology to examine hospital efficiency scores for the 100 hospitals in the sample classified as for-profit, mission and public. Outputs of the study include inpatient days and outpatient visits. The number of beds, doctors and nurses were used to capture hospital inputs. The findings indicated that there was a marked deviation of efficiency scores from the best practice frontier with for-profit hospitals having the highest mean PTE of 71.1%. The mean PTE scores for mission and public hospitals were 64.8% and 62.6% respectively. About 85 %, 83 % and 91 % of the for-profit, mission and public hospitals were found to be operating below their average PTE. More than half of the hospitals are being run inefficiently. Of more importance to this study is the fact that the hypothesis of for-profit hospital superiority was accepted implying that for profit hospitals are significantly more efficient than the non-profit category. The study indicated that the amount of inputs being used could be decreased substantially without decreasing the quantity of outputs achieved. In each of the hospitals included in the study, the total input reductions needed to make inefficient hospitals efficient are more than 50%. These input savings could go a long way in achieving other health concerns without mobilizing additional resources in the sector.

LIST OF FIGURES AND TABLES

LIST OF ACRONYMS

AE	Allocative Efficiency
BCC	Banker, Charnes and Cooper
CCR	Charnes, Cooper and Rhodes
CRS	Constant Returns to Scale
CSO	Central Statistical Office
DEA	Data Envelopment Analysis
DMU	Decision Making Unit
DRS	Diminishing Returns to Scale
FP	For Profit
IRS	Increasing Returns to Scale
MOH	Ministry of Health
MOH&CW	Ministry of Health and Child Welfare
NFP	Not For Profit
NGOs	Non-Governmental Organizations
NHS	National Health Strategy
OTE	Overall Technical Efficiency
PTE	Pure Technical Efficiency
SE	Scale Efficiency
SFA	Stochastic Frontier Approach
TE	Technical Efficiency
VRS	Variable Returns to Scale

ZACH	Zimbabwe Association of Church-related Hospitals
ZNHP	Zimbabwe National Health Profile

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

Health care systems within Sub-Saharan African countries are faced with increasingly scarce resources, which constrain their ability to extend health services of acceptable quality to the vast majority of their people. This severe shortage of health care resources is accounted for by a host of factors, the most important of which include poor macroeconomic performance, cutbacks in public spending, rapid population growth, the AIDS epidemic and the resurgence of diseases such as malaria (Zere. et al, 2001). The constrained ability to adequately meet health care needs is exacerbated by the extensive inefficiency in the health care systems of developing countries, especially within hospitals. The World Bank's study on Financing Health Services in Developing countries indicates that one of the major problems in African health care systems is inefficiency of government health programs, with the others being problems of allocation and inequity (Akin et al, 1987). Empirical evidence emerging from various studies: [e.g. South Africa (Zere, 2000), Mozambique (Mbalame, 2001), Kenya (Kirigia, 2002) and Ghana (Osei et al, 2005)] also indicates the wide prevalence of technical inefficiency in hospital care provision.

Barnum and Kutzin (1993) in their study found that hospitals in developing countries absorb an average of 50 – 80 percent of public sector health resources and that percentage varies from country to country. It is because of this enormous consumption of resources that the efficiency of the hospital sector merits close attention and scrutiny. In view of the deteriorating macroeconomic conditions plaguing Zimbabwe coupled with dwindling donor support, the health sector delivery system has been on a down spiral despite government efforts to revive the health sector (National Health Strategy for Zimbabwe, 1997 – 2007, 1998). There is also an increasing recognition that improved health status contribute significantly to economic development. At the Millennium Summit in 2000, Member States of the United Nations reaffirmed their commitment to eradicate world poverty and improve the health and welfare of the world's poorest by 2015 (WHO, 2005). Thus improvement in health is at the centre of the Millennium Development Goals (MDGs). The achievement of the health access and care among other things requires the availability of adequate resources for the health sector to improve access and quality of care. Therefore issues of efficiency need to be looked at before strategies to mobilize additional resources are considered. Good health is important because it is an intrinsic element of human well-being. As a component of human capital, health is a key factor in the creation of wealth. In a study of the connection between health and wealth Pritchett and Summers (1996) conclude that wealthier nations are healthier nations. A stressed healthy system and an unhealthy labor force do not augur well for any meaningful economic activity. As health is a form of human capital, its disruptions inevitably dislocate all the fundamental links this sector has with the rest of the economy.

Efficiency is a way of generating more resources without necessarily looking for additional investment. The fact that resources are wasted in an inefficient system means that an improvement in efficiency is similar to an increase in resources that can be used in the system. Although health resources are always scarce, inefficient use of those resources severely restricts the ability of health planners and policy makers to extend health services of acceptable quality to the general public. The hospital sector in Zimbabwe has been undergoing significant structural changes in the way in which hospital care is provided for the past two decades. One important structural change has been the growth of the private sector. The role of the private sector (for-profit and not for-profit) in health is becoming very significant in Zimbabwe.

Many health care analysts have suggested that a continual movement in the direction of profit incentives in the production of hospital care might significantly improve the industry's uninspiring economic performance. These individuals arguing from the theory of property rights believe that the non-profit hospital is inherently inefficient because no individuals` income is tied to economic performance (Nyman et al, 1989). Furthermore, it is believed that the non-profit hospital administrator may cause the hospital to pursue goals other than strict cost minimization. While the theoretical case to argue for for-profit hospitals is relatively strong, however, empirically the case for superior for-profit hospital performance is unclear. The purpose of this study is therefore to examine efficiency and to evaluate the impact of profit incentives on technical efficiency in the production of hospital care.

1.2 STATEMENT OF THE PROBLEM

Given the lack of resources in Zimbabwe there is a great need to figure out how to allocate the available resources effectively and efficiently in order to maximize the returns from investment in this sector. Seemingly the major challenge related to hospitals has been inefficiency in the utilization of the existing resources of the hospitals rather than availability of these resources. As a result, this calls for a thorough investigation in the operations of health institutions that takes a significant proportion of the health budget.

Hospitals in the country claim a significant proportion from the resources allocated for health expenditure. Out of the total health expenditure, hospitals consume about 59 % (National Budget Statement, 2007), which is a considerable proportion by any standards. Mission facilities on the other hand, perform a function that should be undertaken by the government. For this reason, the government reciprocates with financial grants and other technical expertise to subsidize the activities of the mission facilities. The public sector provides annual grants to mission facilities for recurrent expenses. State-owned and mission hospitals are therefore resource intensive and as such require an efficient management system. Given the fact that resources for the health care, public and private, are scarce in Zimbabwe, policy makers have recognized the importance of utilizing these resources in a manner which maximizes their health benefits for the country's population. This means ensuring that resources are devoted to those activities which will bring about the greatest improvement in health care status especially when their costs are relatively modest. In addition, large amounts of resources should not be assigned into health care activities which confer only minor benefits at a high cost per individual treated.

Furthermore the overall socio-economic situation in Zimbabwe is characterized by remarkably high poverty rates, numerous and devastating epidemics like HIV/AIDS and insufficient health resources. Some of the challenges facing the hospital sector in Zimbabwe are shortages of drugs, longer waiting periods before treatment, high user charges, expensive equipment lying idle, poor utilization of resources and shortages of medical personnel. Against this background, it is therefore necessary to investigate the levels of technical efficiency at which health facilities in Zimbabwe operate and provide mechanisms for efficiency improvement.

1.3 OBJECTIVES OF THE STUDY

The primary motivation of this study is to find out whether the scarce resources in the health sector are being efficiently utilized. In other words, the general objective of this study is to measure the levels of technical efficiency at which health facilities operate and provide mechanisms for efficiency improvement. The specific objective is therefore to examine the technical efficiency of both for-profit and not-for-profit hospitals in Zimbabwe with a view to comparing and contrasting the relative levels of efficiency within these two categories of health providers and to determine possible causes of inefficiency and ways for improvement. The specific objectives are to;

- Examine the technical and scale efficiencies of public and private (mission & for profit) hospitals in Zimbabwe.
- To evaluate the relationship between technical efficiency and hospital ownership or to determine the relative efficiency of private and public hospitals.
- Find out the returns to scale each hospital is exhibiting.
- Identify some of the factors that are likely to influence the (in)efficiency of public and private hospitals.
- To make policy recommendations on how efficiency can be improved if it is a challenge.

1.4 HYPOTHESES

The study will test the following hypotheses:

- The for-profit hospital sector performs relatively efficient than the non-profit hospital sector as implied by the property rights hypothesis. This is based on the premise that the public or state-owned hospital sector is inherently inefficient and that the market dynamics can help improve the efficiency thereof.

- The non-profit private hospital sector (mission hospitals) performs relatively better than the public hospital sector.

- State or public run institutions are inherently inefficient. This has been accounted to the fact that they pursue, among other goals, populist goals that normally conflict with the goal of strict cost minimization. Further more it has also been argued that no individual's income is correlated with performance.

- The form of hospital ownership has a significant bearing on the performance of hospitals.

1.5 SIGNIFICANCE OF THE STUDY

The application of economics in health is a new discipline of academic study and as such not many studies have been undertaken to investigate the application of economics to health in Zimbabwe. The significance of this study hinges upon the fact that previous research in this area has been confined mainly in the developed countries. Few researches have been partially done in developing countries particularly in Africa. Therefore this study provides an invaluable contribution to the field of health economics in Zimbabwe. The hospital sector has been selected because of its influence on the productivity of workers and on total health expenditure. This follows from the recognition that improved health status has a significant contribution to economic development (Grossman, 1972).

The other distinguishing characteristic of this study are its objective of coming up with a broader understanding of the relationship between hospital ownership form and efficiency. Most of the previous literature in hospital efficiency has been concentrated on public or state run hospitals. However, the widespread dynamism of the health sector today necessitates an enquiry into the private hospital sector. The widespread failure of public hospitals has generated larger concern in the alternative option which is the private sector. Moreover there is an increasing recognition of the relevance of market dynamics in the health sector and on performance thereof. It is imperative to note that the major drive of this study hinges on investigating the impact of profit incentives on hospital efficiency.

1.6 OUTLINE OF THE CHAPTERS

The rest of the study is organized in six chapters as follows: The prime purpose of chapter two is to describe a review of the health sector in Zimbabwe since independence (1980). The third chapter provides a detailed theoretical and empirical literature on the conceptualization of efficiency in general and on hospital efficiency in particular. The forth chapter presents a description of the data and the methodology used to measure hospital efficiency and the justification of using that procedure. In chapter five, we estimate the DEA and regression models followed by the presentation and analysis of results obtained. The dissertation`s key findings, policy recommendations, suggestions for further research and conclusion are contained in chapter six.

CHAPTER TWO

THE HEALTH CARE SYSTEM IN ZIMBABWE

2.1 INTRODUCTION

Zimbabwe is currently facing a huge challenge of a virtually collapsing health care system. Against a backdrop of severe political instability and macroeconomic constraints such as rising poverty levels, unstable exchange rates, high inflation breaching the 11 million percentage mark (Reserve Bank Zimbabwe, 2008) and unemployment reaching an 80% mark (Central Statistics Office, 2008). On key health indicators the country is drifting away from the targets for achieving the Millennium Development Goals. Life Expectancy at birth has dropped to 37 and 34 years for males and females respectively, the lowest in the world according to World Health Organization (WHO, 2006). The brain drain phenomenon is currently the topical issue in Zimbabwe where deteriorating economic, social and political conditions are aggravating the emigration tide. The health delivery sector is possibly the worst affected by migration as health workers are emigrating in search of greener pastures in Southern Africa, Western Europe, North America and Australia (Chikanda, 2004). Poor working conditions and low remuneration are cited as the main push factors.

Soon after independence in 1980, the government of Zimbabwe was one of the major and leading providers of health services as the government sought to redress the inequities in health care that existed prior to independence. The government's main policies and strategies for the health sector during the post independence period were spelled out in its 1984 white paper entitled "Planning For Health Equity in Health" and in its 1987 " Health For All Action Plan." The key sect oral objectives were to improve access and equity in the provision of basic health services. Thus the policy emphasized accessibility and equity rather than efficiency. Efficiency combined with accessibility and equity eventually became the major subject of much focus for policy makers. The ability of the government budget to maintain the remarkable progress achieved in the health sector since independence depends on the efficiency with which services are delivered.

Despite the achievements of the post-independence era, there has been evidence of inefficiencies in the deployment and use of resources within the health care sector. These inefficiencies have generally been signaled with shortages. However while there are admittedly major shortfalls in material, human and financial resources there is however potential for improved efficiency in the use of resources such as staff and drugs. There is a need to cultivate a culture of service planning that is focused on improvement to ensure a more rational deployment of resources as well as the cost-efficiency of their use.

The objective of this chapter is to search for ways in which Zimbabwe can obtain more "value for money" in its expenditure for health care particularly in hospitals which account for the overwhelming majority of health expenditure. This entails improving both allocative and technical efficiency, especially in the public sector but also among private health providers. In the following section aspects of technical efficiencies in Zimbabwe's health sector are examined and options for enhancing such efficiency are explored. Focus shall be confined to technical efficiency only since this study is centered on technical efficiency. A detailed analysis of both concepts of efficiency shall be explored in the literature review.

2.2 AN OVERVIEW OF THE HEALTH STRUCTURE

The institutional make-up of Zimbabwe's health sector is highly heterogeneous comprising the central government hospitals, municipal hospitals and clinics, church missions and NGO`s as well as private hospitals and clinics owned by industrial mining and agricultural enterprises for their employees and families. A full-fledged traditional health sector is also in existence. Most traditional health practitioners in this sector are members of the Zimbabwe National Traditional Healers Association (ZINATHA). The ZINATHA was formed to regulate the activities of traditional health practitioners and to create a legal environment that would enable these practitioners to operate as part of a comprehensive healthcare system. With the worsening of the

HIV and AIDS epidemic, NGOs also play an increasingly more important role in the prevention and control of HIV as well as dealing with the social impact of the epidemic. As a result NGOs and AIDS Support Organizations are actively involved in the provision of HIV and AIDS, TB and Malaria prevention services as opposed to curative. The system of public and mission health services is funded through several sources. Direct grants or budgets from the Ministry of Health and Child Welfare (MOHCW) are the principal source of funds. The historical background of Zimbabwe's health care system was greatly influenced by the British. Larger hospitals were built in the main cities and these have become the main referral centers today.

In Zimbabwe the ministry that is responsible for the health sector is the Ministry of Health and Child Welfare (MOHCW). Its mission: "To promote good health and quality of life for all Zimbabweans by ensuring equal access to health facilities." In the past donor funding used to complement government support for the health sector. However, owing to the political instability currently prevailing in the country donor funding has since dwindled. User fees are a second source of funds, in virtually all cases. Hospitals are allocated budgets for inputs such as salaries, supplies and provisions based on historic levels of spending and purchasing is controlled. The purchasing function is also controlled whereby medical supplies and drugs are obtained by requisition through the National Pharmaceutical Company of Zimbabwe (NatPharm). Other purchases are made by issuing requisitions to government-approved vendors.

One of the obstacles to achieving universal health care in Zimbabwe is inadequate funding in the entire public health sector. Currently, public facilities are over-stretched by the demands of a growing population and the increase in the HIV and AIDS pandemic. There is little development of these facilities due to the acute shortage of resources. The need to mobilize extra resources for the health sector is however unavoidable. Over the years, the government has been looking to the private sector to participate in either the provision and financing of health care. Private for profit providers, although concentrated in urban areas, have been in existence for many years. Non-profit private providers in the form of mission hospitals are also an integral part of the health delivery system in Zimbabwe. The government now accepts that there is very little scope to further increase the role of these mission hospitals (NHS, 1998). An alternative solution is to promote wider involvement of the private for-profit sector in providing and financing health care in the country. In the 1980s, the government had little faith in the private for-profit sector; however, today the private for-sector is acknowledged as a complementary partner in the health delivery system. In order to realize maximum gains from the private-for-profit sector, there is a need to foster an enabling environment for their participation.

The following section examines the changes in the economic environment which have taken place in Zimbabwe since independence in April 1980, concentrating on those aspects which are relevant to health. It also describes the post-independence restructuring of the health sector itself.

2.3 A REVIEW OF THE HEALTH SECTOR SINCE INDEPENDENCE (1980)

The most marked characteristic of the health sector prior to independence were the geographical inequities in the distribution of services and racial discrimination in accessing health services. The health status of Zimbabwe's population mirrored a noticeable unequal socio-economic structure characterized by appalling racial inequalities between whites and blacks and significant inequalities between the rural and urban populations (National Health Strategy, 1998). Thus the health sector that existed at independence encompassed policies designed to provide health care for a minority. Hospitals were defined according to four categories based on location: central, general, district and rural. In urban areas, hospitals and polyclinics provided mainly curative service while rural populations relied on mission hospitals, rural hospitals and a few district hospitals. Rural facilities played a minor role in promotive and preventive health services and the quality of curative care was generally poor. Thus while the majority suffered excess mortality and morbidity; the well-off white minority enjoyed a health status similar to that of the populations of developed countries. The health services also showed the familiar pattern, with expenditure concentrated on sophisticated facilities in the towns, leaving the rural majority with practically no services at all.

At independence, the government adopted a primary care approach and organized public and mission facilities into a four-tier system, with primary care in the first level, district hospitals and health services as the second, provincial facilities in the third tier and central hospitals in the forth level where services of increasing complexity were offered requiring more specialized personnel and equipment . In principle, individuals seeking medical care are not supposed to go to higher level facilities without being referred from a lower level. Due to inefficiencies within the referral system bypassing the lower levels is very common. As a result central hospitals treat a mix of highly specialized and routine cases that could be treated at lower levels.

Shortly after independence in 1980, the government of Zimbabwe like many other developing countries was determined to achieve equity in health through extensive development of government-owned and government-financed health services. The government was the main provider of health resources through direct budget support to government institutions and through a system of grants to local authorities and mission institutions. The “Growth with Equity and Primary Health Care" policy which wee adopted, emphasized rural health care development and resulted in the development of a number of health care programs, which included mobile clinics, integration of the traditional midwives into the formal health system and training of village health workers. Primary Health Care (PHC) was adopted because of its principles of affordability, accessibility and acceptability. The components of PHC included family planning, treatment of minor diseases, health awareness and education, immunization, water supply and sanitation.

Another evident characteristic of the health care sector after independence was the skewed distribution of resources (Planning for Equity in Health, 1984). Rural populations with the greatest health needs received the least health care resources. On the other hand, urban areas received more resources because they had more health facilities. The distribution of health personnel for instance, was biased towards urban areas. While 66% of doctors worked in the urban central hospitals, only 12% worked in rural areas. The doctor: population ratio in Harare and Bulawayo was just under 1:4000 while in rural areas the ratio was approximately 1:62000. Thus the health profile of 1981 reflected inequities in health status, health care provision, social status and allocation of health care resources. These inequities were apparent in the distribution of resources as can be seen in Table 2.1 below.

Table 2.1 Basic data on Government hospitals, 1981

HOSPITAL CATEGORY	HOSPITAL	BEDS	DOCTORS	NURSES	HOSPITAL EXPENDITURE AS A % OF GOVERNMENT EXPENDITURE
CENTRAL	4	3000	223	1568	60
GENERAL	11	2038	39	789	21
DISTRICT	28	2400	16	422	10
RURAL	46	2029	0	235	3

Source: Planning for Equity in Health, 1986.

During the first ten years of independence there was a remarkable change in the health sector as a whole. The country witnessed a rapid growth in absolute numbers of health facilities and programs designed to promote public health. During this period, the government of Zimbabwe extended basic health care to poor deprived rural areas by constructing and staffing about 290 clinics, reconstructing about 160 war-damaged clinics and upgrading an additional 160 health facilities. According to research conducted in the 1980s (Planning for Equity in Health, 1986) showed that the infant mortality rate amongst children under one year in 1980 was 96 out of 1000 live births. By 1990, this figure had gone down to 55. Maternal mortality rate was 120 out of every 100,000. By 1990, this figure had gone down to 86. Life expectancy in 1980 was 50 years and by 1990 it was at least 60 years. This achievement was attributed to Public Health Care (Planning for Equity in Health, 1986).

Unfortunately in the 1990s due to the impact of the HIV and AIDS pandemic, episodes of serious drought, the Economic Structural Adjustment Program (ESAP), public industrial action by health personnel and the huge brain drain from public to the private sector and to outside the country. The government increasingly found itself unable to adequately address the health requirements of its population because health sector resources failed to match demand. As a result, in 1999, infant mortality rose to 120 out of 1,000 live births, and maternal mortality rates rose to over 100 women out of 100,000. Life expectancy rates have decreased significantly to 37 years for males and 34 for females (WHO, 2007). It was during this period that the government began to search for alternative ways of financing, providing, and managing health care services in the country. At this point, the government warmed up to policies that considered the growth of the private health care provider. In the late 1990s, the government acknowledged that despite its heavy investment to ensure public provision of health services, the private sector was playing a significant role in the provision of health care services. Supporting the role of the private sector in providing and financing care and services was acknowledged as a possible means to controlling government health costs. Private medical care has since proved itself an important aspect of the health care system not only in Zimbabwe but also in many other developing countries.

2.3.1 Health Expenditure

In the years immediately following independence, public spending for health increased rapidly by 38 % in 1981 and 13% in 1982 (in real terms) as the government sought to expand basic health services to communities that had been previously neglected. A series of preventive health initiatives were also launched in such areas as rural water and sanitation, child immunization, pre-natal care, malaria and other communicable disease control. Health share of GDP also rose from 3.1% in 1981 to 4.2% in 1986. The increases in public expenditures for health also outpaced overall GDP growth, averaging 8.7% a year in real terms from 1981 and 1985 (Zimbabwe For All Action Plan, 1986).

While the private sector, including both modern and traditional practitioners, for profit and non-profit institutions have had an important part in providing and paying for these growing health services, the public sector-especially the central government has played the leading role in this process. For instance in 1987-1989, the public sector was responsible for 63% of all health-related expenditure, including MOH (40%) and other central ministries (13%) and municipal and local government (10%). The private sector, which provided the balance of health services expenditure of (37%), was composed of private doctors (25%), industries, mines and commercial farms (8%) and NGO`s (4%) (Zimbabwe National Health Profile, 2002).

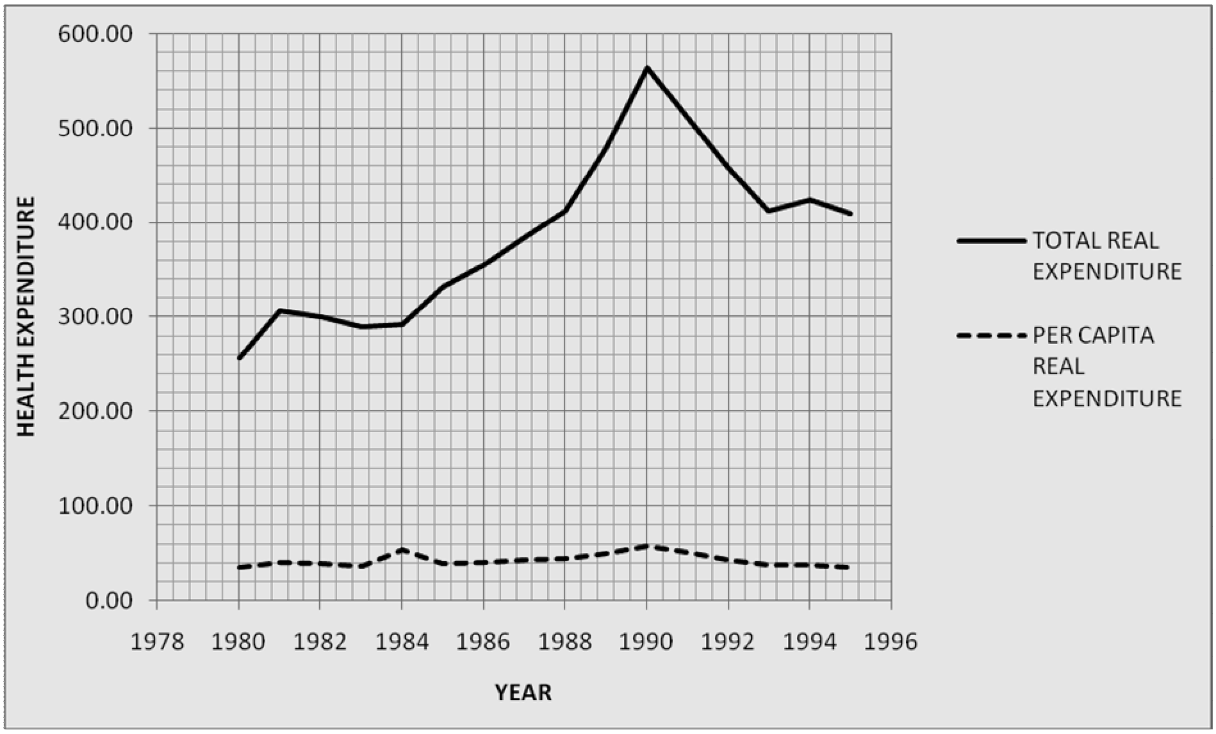
In recent years, however the pace of government spending for health has been slowing down markedly and the per-capita level of expenditure has diminished because of the severely constrained macroeconomic situation facing the government. This has been compounded by a sustained high rate of population growth of over 3% a year (Central Statistics Office, 2004). The visual depiction of the trend of health expenditure and its related components are tabulated in Table 2.2 and Figure 2.1 below. Today, government spending for health has become constrained by the worsening economic situation and as a result levels of real per capita public spending have been falling over the years.

Table 2.2: Real Recurrent Health Expenditure (at 1990 prices), Ministry of Health & Child Welfare.

YEAR	TOTAL REAL EXPENDITURE (Z$ MILLIONS)	PER CAPITA REAL EXPENDITURE	HEALTH EXPENDITURE AS % OF GDP	HEALTH EXPENDITURE AS % OF GOVT EXPENDITURE
1980/81	256.24	35.62	2.00	5.30
1981/82	306.60	41.44	2.20	5.60
1982/83	300.10	39.44	2.10	4.80
1983/84	289.61	36.74	2.20	4.80
1984/85	291.97	53.92	2.20	4.90
1985/86	331.00	39.48	2.30	5.30
1986/87	355.64	41.14	2.50	5.10
1987/88	384.10	43.08	2.80	5.50
1988/89	412.18	44.82	2.80	5.40
1989/90	478.88	50.50	2.60	5.90
1990/91	564.49	57.72	2.80	6.20
1991/92	511.97	50.76	3.00	5.10
1992/93	458.18	44.00	2.50	5.30
1993/94	412.18	38.45	2.40	5.10
1994/95	424.33	38.31	2.20	4.50
1995/96	409.71	35.86	2.20	4.20

Source: Chandiwana S et al, 1997

Figure 2.1 Total & Per-capita Real Expenditure.



2.3.2 Human Resources in the Health Sector

The World Health Report 2000 (WHO 2000) argued that health human resources are key determinants of the success or failure of health systems. The performance of health care systems is a function of the availability, know-how, skills mix and motivation of personnel delivering the services.

Zimbabwe, like many other countries in the region, is badly affected by a serious shortage of health workers. Many of the health indicator improvements achieved during the first ten years of independence are on the decline and one of the contributing factors to this is the shortage of skilled and experienced health workers at a time when demand for such services is increasing due to a growing population and the challenges posed by HIV and AIDS. The public sector provides as much as 65% of health care services in the country (MOHCW 2004), and so any shortage of public sector health workers affects a great majority of the population. Health professional groups with the highest loss rate in Zimbabwe are doctors, nurses and pharmacists. Although other non-clinical health workers make a significant contribution, it is the shortage of these key clinical professionals that limit accessibility to health care for the majority of patients, especially those that are economically disadvantaged and located in geographically-deprived areas. Since 1999, increased numbers of skilled health workers have migrated to regional and international destinations. In 2004, the MOHCW published figures that showed that 2,825 work permits were processed for Zimbabwean health professionals to enter the United Kingdom. The figure represents about 25% of the professional health workforce in the Zimbabwean public sector (National Health Strategy for Zimbabwe 1997-2007, 1998) Attrition of health workers in the public sector is due to factors that mainly relate to lack of incentives, poor salaries and conditions of services (Paulinus, et al 2000).

The shortage of medical personnel in the country has been attributed to inadequate training facilities for the various health personnel and to significant differentials in salary and conditions of service between the public and private sectors. As part of the post-independence expansion of health services, the government initiated a number of strategies to increase the health sector workforce. Since independence, there has been a high level of investment in health manpower and a steady growth in the number of health workers employed in the public sector. Despite this increase in numbers, professional manpower is being lost from the public to the private sector as well as to neighboring countries. For example during 1992, out of 1534 registered medical practitioners only about 46% worked in the public sector (National Health Strategy, 1995). Health workers complain of low wages and of being overworked and unable to fully manage their work environment. The inability of the public service to produce workable strategies for the retention of skilled health workers is a major challenge for the future health service. To address these problems, the government will need to establish economic stability and to develop a comprehensive health staffing strategy.

Many of the problems identified by the 1987 “Health For All Action Plan” review commission, still persist in the health sector and are a hindrance to the full utilization of available manpower.

There are continuing problems of low staff morale, low productivity, poor distribution of staff and poor salaries of health workers. The exodus of specialists' health manpower from the public to the private sector is occurring at a time when this sector is offering limited service to the rural population representing the majority of Zimbabweans. There are currently no incentives to entice health personnel to move from urban to rural areas. About two thirds of Zimbabwe's physicians and state registered nurses (SRN) and a third of the country's state certified nurses (SCN) work in the private sector where pay and conditions of service are generally favorable compared to the public sector. As a result the public institutions continue to experience shortages of staff especially doctors, nurses and especially at provincial, district and health center levels. Thus one of the challenges facing the government is to devise a system of incentives and regulations that will ensure adequate staffing of public health institutions given that government training facilities are graduating many doctors and nurses annually.

2.3.3 Drugs

The shortage of drugs characterizes health systems of many African countries as their procurement, storage and distribution systems are poorly organized. According to Sahn and Bernier (1993) only 12 % of the drugs purchased by African countries reach consumers in good quality. Like many African countries, the Zimbabwean public health sector faces chronic scarcity of drugs, lack of foreign currency and high dependence on aid. The storage facilities especially in rural areas are poor and most of the drugs reach the patients in bad quality which further compounds the drug shortage. One specific area in which health planners have recognized for a number of years where technical efficiency could be improved has been the utilization of pharmaceuticals, on both an in- and out-patient basis. At the time of independence it was felt that expensive brand-name drugs catering to an urban elite, were absorbing most of funds and foreign exchange available for pharmaceuticals. The drug cost per patient treated was therefore unnecessarily high and other persons who needed drugs were not receiving them. This was because of national shortages, weaknesses in the distribution system and inability to pay. In order to correct this situation the government proposed in its policy statement "Planning for

Equity in Health" the adoption of an "essential drugs list" for use by government facilities, offering potential for greater internal efficiency and equity. The policy has been implemented through the Zimbabwe Essential Drugs Action Program (ZEDAP), which since its establishment in 1986 has developed and published a national essential drugs list (EDLIZ) for use by government health workers. ZEDAP trains nurses and doctors in cost-effective prescribing practices. ZEDAP also makes possible important cost savings through bulk procurement and the use of international competitive bidding.

While ZEDAP has brought many positive changes in the efficiency and equity of drug use in Zimbabwe, several problems still remain. The current system for allocating the country's increasingly scarce foreign exchange does not satisfy the drug needs of government health facilities. The foreign exchange content of Zimbabwe's pharmaceuticals is high whether the drugs are imported in finished form or are manufactured locally using imported materials and the competition of limited foreign exchange is intense. The fact that government hospitals and clinics experience periods of shortages of drugs is neither equitable nor efficient: Equity goals are inhibited because peripheral facilities serving lower-income families suffer the most severe shortages, while efficiency is impaired because drug shortages at primary care facilities tend to encourage referrals to higher-cost secondary and tertiary hospitals.

2.3.4 Health infrastructure

Zimbabwe made progress especially in redressing the marked inequities in health care that existed prior to independence. By 1990, the government had extended basic health care to underserved rural areas, by constructing and staffing over 290 clinics, reconstructing about 160 war-damaged clinics and upgrading existing rural clinics to ensure access to basic services by the entire population (NHS, 1998).

About two decades since independence the government increased access to health services and facilities. To date, according to the National Health Strategy for Zimbabwe 1997-2007 ( ), 85 % of Zimbabweans live within 8 km of a health facility, while half of the population in rural areas lives within 4 km of a health facility. Since 1980, 250 additional Rural Health Centers have been built throughout the country. During the period 1986-1997 a total of 25 district hospitals with a bed complement ranging from 52 to 140 were built or upgraded.

The basic premise for infrastructure development is the need for each province to have a minimum package of infrastructure comprising of at least one health centre within 8 –10 km of each household, one hospital for each district and one hospital with specialists' service for each province. At least most of the districts in the country have a government hospital. However there is need to address the issue of maintenance. A number of facilities have deteriorated because of lack of maintenance and staff. The maintenance and refurbishment of existing facilities has continued to lag behind with some facilities in an advanced stage of dilapidation.

2.4 CONCLUSION

This chapter discussed the structure, conduct and performance as well as the challenges facing the Zimbabwean health sector in terms of its composition and problems. The sector was found to be characterized by poor economic performance and chronic shortages of basic resources such as personnel, equipment and a decaying infrastructure base. These challenges are aggravated by the lack of financial resources and an inefficient use of existing resources as well as the decline in foreign aid. This is consistent with the statement of the Minister of Finance DR H. Murerwa in his presentation of the 2007 Budget, "The present high inflationary environment coupled with shortages of foreign currency and manpower is compromising provision of quality health service to the people. Improvement in levels of funding for medical drugs and supplies as well as hospitals and clinical equipment remains critical" (National Budget Statement, 2007).

The next chapter will discuss the literature behind efficiency studies at the level of health care facilities.

CHAPTER THREE

LITERATURE REVIEW

3.1. INTRODUCTION

This chapter outlines the theoretical and empirical literature behind efficiency studies at the level of health care facilities. The theoretical section conducts a review of the literature on efficiency and the two main tools for empirical measurement; the Data Envelopment Analysis (DEA) and the Stochastic Frontier Approach (SFA). The empirical section explores studies, applications and observations that have been conducted by different authors in different countries regarding hospital efficiency.

3.2. THEORETICAL LITERATURE

3.2.1. Conceptualizing Efficiency

This section of the chapter discusses the various concepts of efficiency and their theoretical base. The section also considers the two frontier approaches to efficiency measurement: the Cost and Production approaches. Efficiency measurement is derived from the cost or production boundary. The theory of Duality ascertains this relationship between production and costs. The concept of technical efficiency can be defined directly in terms of the production or cost frontier. The word "frontier" is applied in either case because the function sets a bound on the range of possible observations. For instance, production can take place only below or on the frontier. Similarly, costs can be observed above the cost frontier but not below the frontier because it is impossible to achieve costs lower than the minimum input requirements implied by the production frontier. The amounts by which an organization lies below its production frontier or the amount by which it lies above its cost frontier can be regarded as a measure of relative efficiency.

3.2.1.1. The Production function Approach

The first empirical treatment of the production function as a frontier is found in the work of Farrell (1957) and Fieldhouse (1962). A production function can be defined as a process of physical transformation in which inputs are combined to generate output. The production function is then interpreted as a purely technical relationship which defines efficient transformation possibilities, given the set of feasible techniques (technology). In the case of inefficiency, the production function may be written as an inequality:

$y_i \leq f(X_i; \beta)$ ....(1)

Where y_i is observed output at establishment *i*, and X_i is a vector of inputs and β a vector of parameters which describe the transformation process. $f(.)$ is the production function and has the interpretation of a frontier or $y_{\max}$. At inefficient operations, potential output ($y_{\max}$) will exceed observed performance (y_i). Hence technical inefficiency implies ($y_i - y_{\max}$) is negative. The difference between observed and potential performance can be treated as a residual in the production function, which is equivalent to the technical efficiency ratio. If these residuals are denoted ε_i then in terms of the production function in [1] above, the technical efficiency ratio can be written:

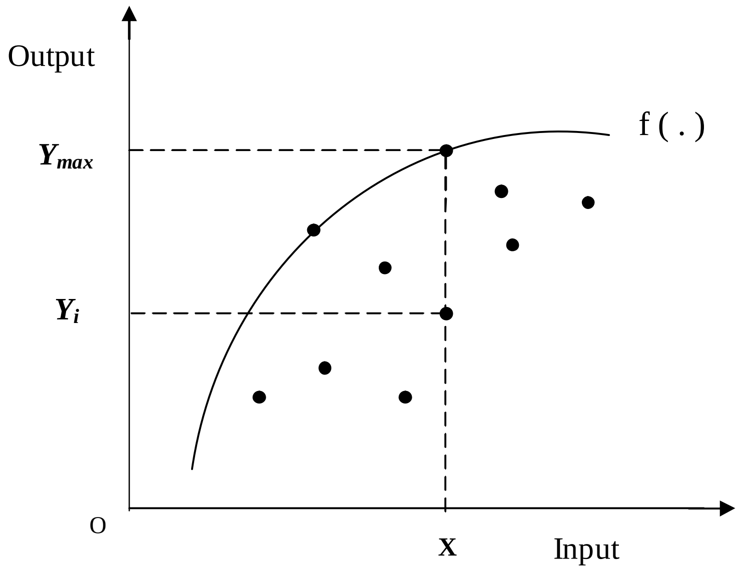
$$\varepsilon_i = \frac{y_i}{f(X_i; \beta)} \quad \text{...(2)}$$

The ε_i is always non-positive to ensure that observed output cannot exceed potential that is, $y_i > y_{\max}$ is not possible.

In Figure 3.1 below, Decision Making Unit (DMU[1]) *i* is producing output y_i which for input allocation *OX* is far less than frontier output y_{max} as can been seen clearly on the diagram. The difference between actual and potential output, ε_i is negative and hence production at unit *i* is relatively inefficient. The implication is that when the efficiency residual is equal 0 the production unit is efficient since actual and potential outputs are equal

[1] Charnes, Cooper and Rhodes introduced the term "decision making units" (or DMU) which is now widely used in literature. A DMU is to be regarded as an entity for converting inputs into outputs.

Figure 3.1: Efficiency and the Production Frontier



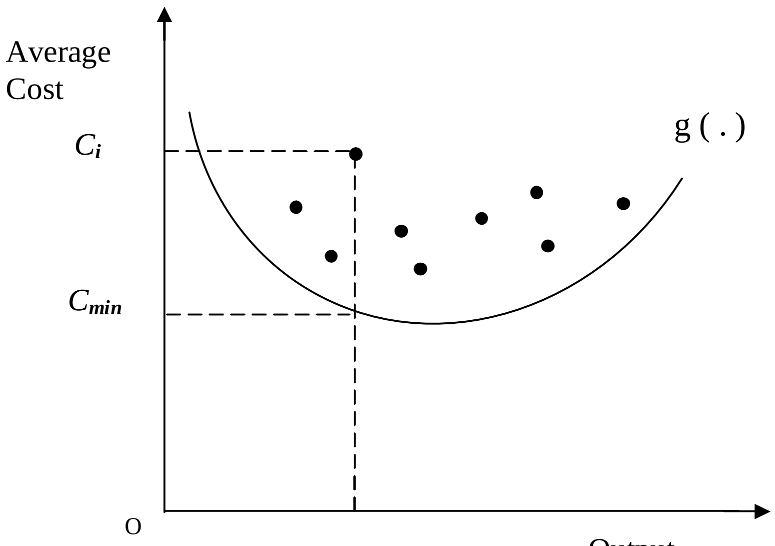
Source: Ganley, 1992

3.2.1.2. The Cost function Approach

The theory of duality between cost and production implies that there exists a dual cost function to the product transformation function in [1] above. A cost function relates the minimized total cost in a firm to output and factor prices. If excess costs are possible then the cost function may be written as an inequality: $c_i \geq g(z_i;\alpha)$

Where c_i represents average cost at establishment i, z_i are determinants of costs and α is a vector of parameters and $g(.)$ has a frontier interpretation denoting minimum costs ($c_{\min}$). The efficiency ratio is defined by the residuals (θ_i) in the cost function. That is: $\theta_i = \frac{g(z_i;\alpha)}{c_i}$. This formulation is equivalent to the ratio of potential to observed costs. In the presence of inefficiency, observed costs are greater than potential and the efficiency ratio is less than unity. This means that the efficiency residuals are positive.

Figure 3.2: Efficiency and the Cost Frontier



Source: Ganley, 1992

Figure 3.2 depicts an analysis of efficiency using the cost frontier approach, where observed costs, c_i at unit *i* are greater than the minimum costs on the appropriate part of the frontier. Since frontier or boundary costs are the minimum feasible, observed costs cannot fall below minimum costs, i.e. $c_i \geq c_{\min}$. This relationship is important in order to maintain the frontier interpretation of the cost function and implies that the residuals in the cost function are non-negative:

$$c_i = g(z_i;\alpha) + \theta_i \text{,} \quad (3)$$

$$\theta_i \geq 0 \text{ for all } i \quad (4)$$

The θ_i is a non-negative quantity as portrayed by the relation in 4 above, and it captures the inefficiency. The implication is that when this efficiency residual, θ_i is equal 0 the Cost unit is efficient since actual costs are equal to the possible minimum.

3.2.2. Concepts and Definitions of Efficiency

3.2.2.1. Technical and Allocative Efficiency

According to Farrell (1957) who pioneered most of the work on efficiency measurement, the efficiency of a firm consists of technical efficiency and allocative efficiency. In Farrell's framework, a firm's efficiency is measured relative to the efficiency of all other firms in the industry, subject to the restriction that all firms are on or below the frontier. In the context of health care, WHO (1999) defines allocative efficiency as when resources are devoted to right activities while technical efficiency is when a given health intervention or health outcome is obtained through few resources.

A production plan is said to be technically efficient if the inputs which are employed produce maximum output or maximum output is produced using the least amount of factor inputs. **Technical inefficiency** (TE) is due to excessive input usage. Within the context of healthcare services, technical efficiency may then refer to the physical relationship between the resources used (say capital, labor and equipment) and some health outcome. **Allocative efficiency** (AE) on the other hand, reflects the ability of an organization to use these inputs in optimal proportions given their respective prices and the production technology. In other words, allocative efficiency is concerned with choosing between the different technically efficient combinations of inputs used to produce the maximum possible outputs. Taken together, allocative efficiency and technical efficiency determine the degree of productive efficiency also known as total economic efficiency. Thus, if an organization uses its resources both allocatively and technically efficient, then it can be said to have achieved total economic efficiency (EE). The basic ideas underlying

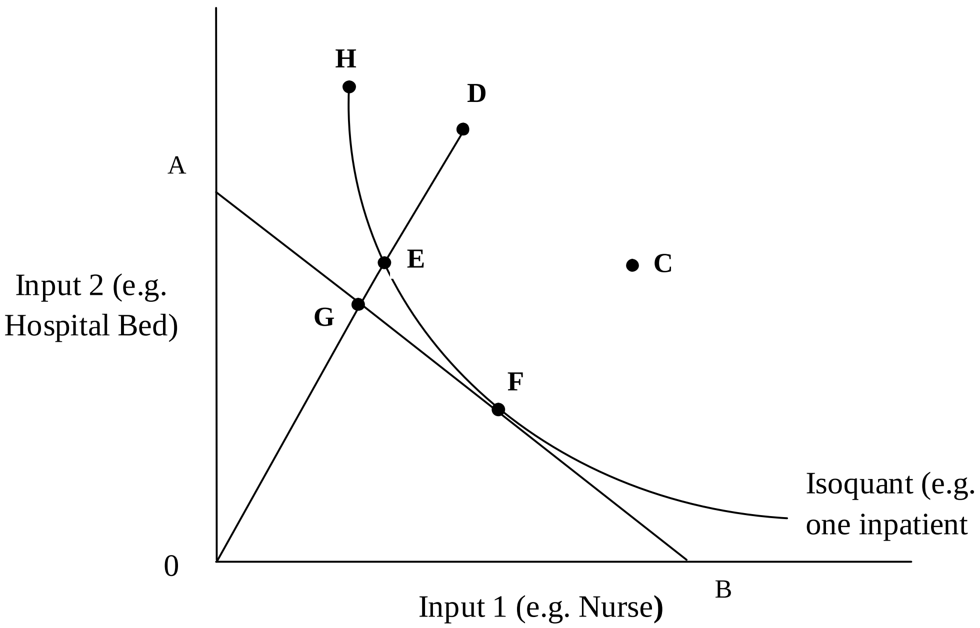
Farrell's concept of technical and allocative efficiencies under the assumption of constant returns to scale are illustrated in Figure 3.3.

3.2.2.2. X-Efficiency and Dynamic Efficiency

X-efficiency occurs when technical efficiency is not being achieved due to a lack of competitiveness and hence a lack of incentives to reduce cost. The concept of x-efficiency is accredited to Harvey Leibenstein (1978) as developed in his book *General X-efficiency Theory and Development*. A good example is a monopoly business structure which makes supernormal profits and therefore has little incentive to get rid of excess labour. As a result the business`s average cost will be higher than necessary. Monopolies are protected from competitive forces by entry barriers and thereby allowing for x-efficiency to occur. However, in a competitive market scenario, firms are continually under pressure from their rivals to produce at the lowest cost possible, thus x-inefficiency does not occur. The introduction of market or competitive dynamics in the health care service provision may help achieve efficiency gains.

Dynamic efficiency considers whether firms are likely to develop more efficient techniques over time. It is therefore necessary for firms to constantly introduce new technology and reduce costs over time. It is a concept that advocates for investment innovation, research and development as vital for efficiency attainment. It can also be defined as the ability to adapt quickly and at low costs to changed economic conditions and thereby maintain output and productivity performance despite 'economic shocks'. It is motivated by increased competition which then acts as an incentive for businesses to innovate and adapt.

Figure 3.3: Farrell's efficiency measurement



Source: Coelli, 1996

In Figure 3.3 above, a hospital produces its output (inpatient day) using a combination of two inputs (Nurse and hospital bed). A technically efficient hospital is one that is located on the Isoquant, that is, on the frontier such as **E, F** and **H**. Hospitals operating at points **C** and **D** are technically inefficient. For the hospital operating at point **D**, the measure of technical efficiency (TE) is given as;

$$TE_D = \frac{OE}{OD}$$

This denotes the ratio of minimal input required to the actual input use, given the input mix used by **D**. The ratio **ED/OD** represents the percentage by which all inputs could be reduced without a reduction in output. If the hospital at point **D** is to be efficient it has to relocate itself to point **E**. Technical efficiency takes values between zero and one, i.e. $0 \leq TE \leq 1$.

Given input prices, the isocost line **AB** represents the minimum cost of producing one unit of output. Allocative efficiency demands that production takes place at the point where the isoquant line is tangential to the isocost line. Given this definition, hospitals operating at points **H** and **E** are technically efficient but allocatively inefficient. Only the hospital operating at point **F** is both technically and allocatively efficient. The Allocative efficiency of hospital operating at point **D** is given as:

$$AE_D = \frac{OG}{OE}$$

The ratio GE/OE represents the percentage reduction in production costs that would occur if production were to occur in the allocatively efficient point **F**. Farrell (Coelli, 1996) proposed that economic efficiency (EE) is measured as:

$$EE_D = \frac{OG}{OD}$$

The overall (economic) efficiency (*EE)* has the advantage that it easily decomposes into technical and allocative efficiencies.

$$\frac{OG}{OD} = \frac{OE}{OD} \times \frac{OG}{OE} \quad \text{That is, } EE = TE \times AE$$

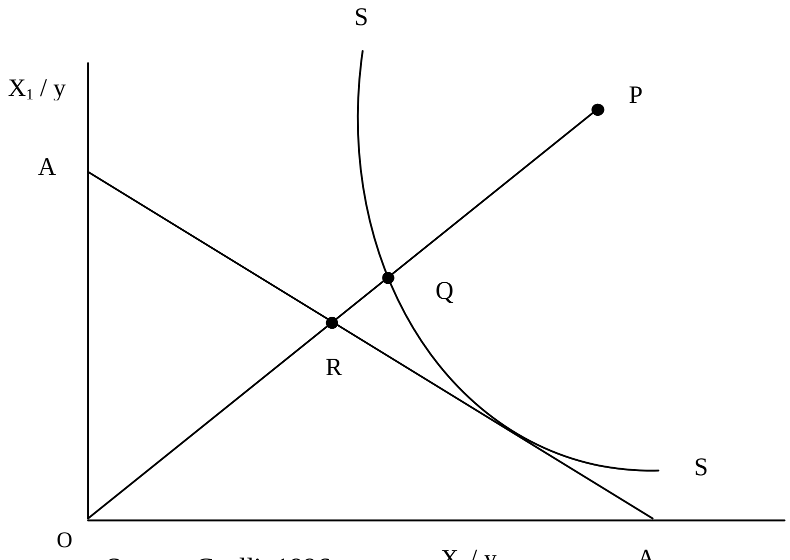
The measures obtained from Figure 3.3 represent input-oriented measures of efficiency. They are input-oriented as their focus is on the measurement of variations in input use between different hospitals for a standardized output.

3.2.3. Input-output efficiency measurement

3.2.3.1. Input-oriented measure

An input orientation approach measures input reductions that are necessary for a production unit to become efficient without a reduction in output. Input inefficiencies show the degree to which inputs must be reduced for the inefficient hospital to lie on the best practice frontier. Suppose a hospital uses quantities, defined by point P to produce a unit of output, the technical inefficiency of the hospital can be expressed in percentage terms by the ratio $\frac{QP}{OP}$ which represents the percentage by which all inputs could be reduced. Ideally the technical efficiency of a DMU is measured by the ratio: $TE = \frac{OQ}{OP}$ which is equal to $\left[1 - \frac{QP}{OP}\right]$. It takes a value between 0 and 1 and therefore provides a measure of the degree of TE of the hospital. A value of 1 indicates that the hospital is fully technically efficient since it would be lying on the efficient isoquant SS.

Figure 3.4.: Farrell's efficiency measurement

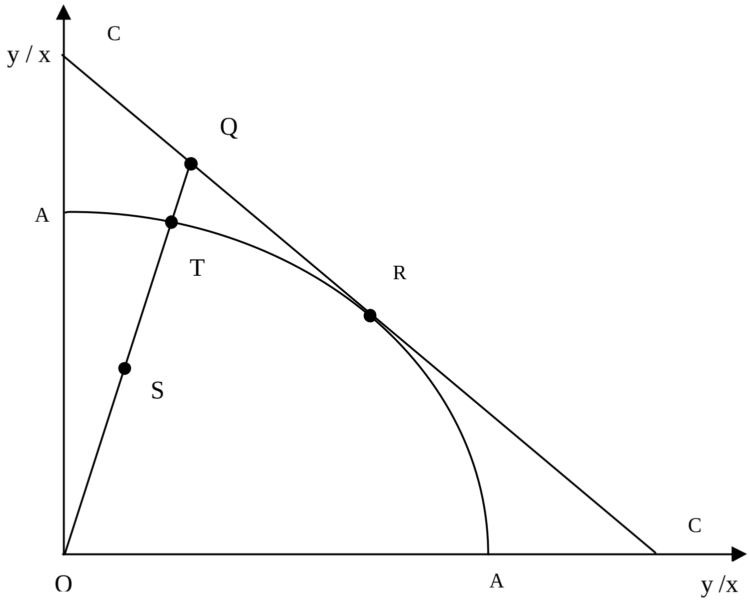


Source: Coelli, 1996

3.2.3.2. Output-oriented measure

An output orientation measures the expansion of output that is necessary for efficiency improvement holding inputs constant. Output inefficiencies represent the needed increase in output for the inefficient hospital to become efficient.

Figure 3.5: Farrell's efficiency measurement



Source: Coelli, 1996

In Figure 3.5 above, the curve AA represent the maximum possible output attainable given the resources available. All points located inside the curve are technically inefficient relative to points on the frontier. For example, the distance ST represents technical inefficiency since it shows how far the point S is from achieving the maximum possible output with the given resources. At point S resources are being underutilized. Thus technical efficiency at point S is:

$$TE_S = \frac{OS}{OT}$$

Figure 3.6 Classic Framework of Efficiency by Farrell

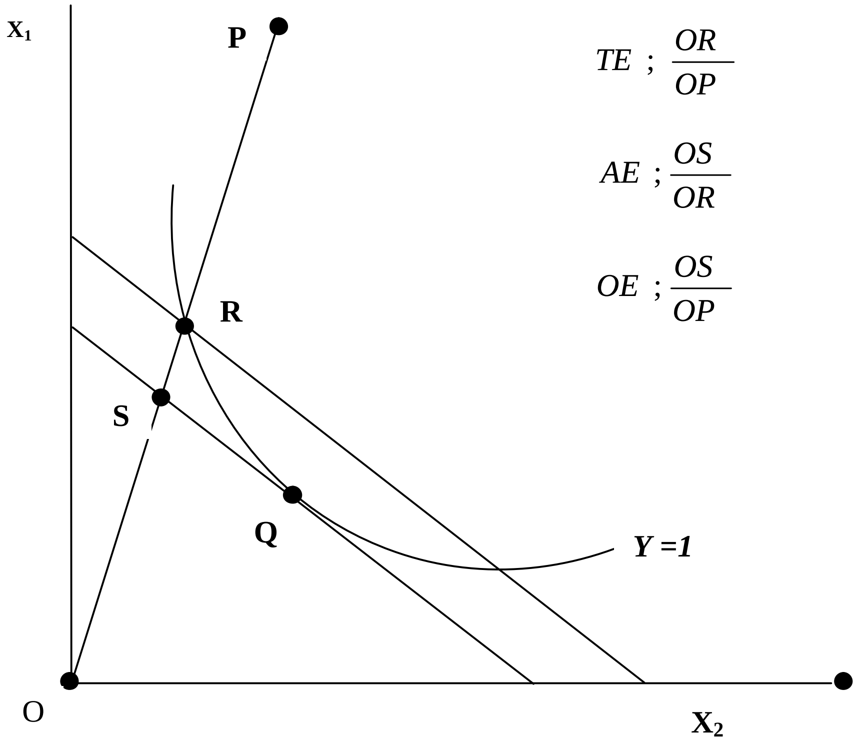


Figure 3.6 shows the classic framework by Farrell which makes it possible to decompose overall efficiency into technical and allocative (price) efficiency. Consider the case of a simple output Y that is produced by using two inputs (X_1, X_2). Under the assumption that the production function $Y = f(X_1, X_2)$ is linearly homogeneous, the efficient unit isoquant Y=1 shows all technically efficient combinations. Point P represents a firm that also produces at Y=1 but uses higher levels of input and is therefore less efficient in a technical sense. The magnitude of technical efficiency can be expressed as the ratio between optimal and actual resource use (OR/OP). By taking into account the isocost line (representing relative factor prices) we can identify allocative efficiency. Any point on the line Y=1 has technical efficiency but only Q receives technical efficiency *(TE)* at minimum cost. Allocative (price) efficiency can be expressed as the ratio between minimum and actual cost (OS/OR) and overall efficiency *(OE)* is the product of technical and allocative efficiency.

3.2.4. Public Sector Efficiency

The magnitude of the significance of public sector efficiency cannot be overemphasized. Public sector efficiency affects performance of the economy through two different channels: by means of positive output changes in the public sector itself, and through effects on the private sector. Public services are responsible for a large proportion of the overall output of an economy. The functioning of the public service sector affects productivity in the private sector through changes in taxes. For example, taxes which are required to fund public services distort relative prices in the economy influencing economic incentives, namely the willingness to work and invest (Coelho, 2007). Moreover, goods and services such as education and health are well known to have a positive influence on the development of economic activity. The extent to which public expenditure has positive growth effects through this second channel depends not only on the size of public expenditure but also on the efficiency of public expenditure. Public service efficiency is therefore an essential driver of the average productivity of an economy.

There has been widespread concern regarding to the performance of public administration in recent years especially on the utilization of resources. Public sector expenditure for instance needs to be efficiently administered if it is to have a positive effect on economic growth. In most cases public sector reforms tend to have multiple ends thereby making it difficult to investigate their efficiency. Public sector reform entails the deliberate changes of the structure and processes of public sector organizations with the objective of getting them to run better (Pollitt, 2002). Public sector efficiency measurement emphasizes the need for a definition of public sector performance which can be quantified. The statistical techniques to investigate public sector efficiency which are discussed in this dissertation depend to a greater extent on a clear definition of inputs and outputs. There is need therefore by the public sector production unit to identify the input factors in quantifiable units it uses in the production of its output. Once identified, public sector efficiency is defined as attainment achieved compared to the maximum that could have been achieved for the observed level of resource use (Tandon et al. 2003).

3.2.4.1. Measurement of Efficiency in the Public Sector.

The main thrust behind public reforms is to promote organizational efficiency and effectiveness towards attaining national development goals. Measuring an organization's efficiency is about the relationship between the outputs it produces and the inputs it uses. An efficient organization would be one that produces the maximum possible outputs given its inputs, or one that produces a certain level of output with the minimum amount of inputs. The process of measuring an organization's efficiency involves three stages. Firstly, its inputs and outputs need to be defined and measured. Secondly, there is need to define the set of feasible input-output combinations that is, the production efficiency frontier. In other words, the researcher must answer the question; what outputs could be achieved for any given set of inputs? Finally, the organization's actual inputs-outputs combinations are compared with the set of feasible inputs-outputs combinations.

These stages are relatively not so complicated in the case of private organisations operating in competitive markets. Even when organisations involve multiple inputs and outputs, prices are usually available for aggregating these operations and therefore efficiency can be easily estimated. The provision of public services is characterised with anomalies that hinder the process of measuring efficiency. Public services are normally provided free of charge or at subsidised prices at the point of delivery, which brings significant obstacles to the determination of their societal value and consequently to the aggregation of their output (Coelho. 2007). Public service inputs normally pose fewer problems as long as they are purchased in competitive markets where prices are available for determining their relative value. However, there are situations where government may have considerable monopsony power; it may bargain with powerful trade unions; or prices may be determined by complex regulatory mechanisms. This in itself distorts the true valuation. Apart from difficulties in measuring inputs and outputs, there are a number of methodological issues relating to the estimation of the production frontier that are likely to have a significant impact on public service efficiency measurements. Some of the most important choices concern the determination of output weights; modelling the production process; controlling for environmental constraints; and allowing for dynamic effects (Smith and Street, 2005).

The next sections will discuss in detail these empirical measurement challenges surrounding efficiency evaluation in the public sector.

3.2.4.1.1. Inputs

The measurement of the inputs used by the public sector presents fewer challenges as most inputs are sold on markets and prices are readily available. In competitive markets where buyers and sellers are price takers and suppliers of inputs bear the full costs of their decisions, input prices reflect marginal social costs and thus can be used directly to value inputs. However, as pointed by Dawson *et al* (2004), there are circumstances where government uses its buying power and others where prices are determined by a complex mechanism (Dawson *et al.*, 2004). Prices determined under these scenarios will not reflect the true valuation.

The valuation of inputs used in the evaluation of efficiency in the public sector is normally not a problem. This is because government departments procure these inputs in competitive markets at the prevailing prices, the input prices will therefore reflect the true valuation. However, in the case where the government is the sole buyer or where a regulatory board determines the input prices, the value thereof will not be reflective of the true scenario. The efficiency measures derived there-from will be misleading.

3.2.4.1.2. Outputs

Production in the public sector is difficult to evaluate in terms of its level and its efficiency. Most public sector departments have a distinctive characteristic of producing multiple outputs and in most cases these outputs are qualitative and lack the physical characteristic of "countability" (Ganley and Cubbin, 1992). The quality of these outputs has not been easy to estimate with reasonable precision. The quality of the output can also be seen as some function of the vector of outcomes it produces (Lancaster, 1971). Constancy in quality of these outputs is important inorder to ascribe any deviation from the frontier to inefficiency only and not to any quality differentials.

Furthermore, there has been a concern on the lack of differentiation as far as the concept of output and outcome is concerned. It is noted that most researchers pay more attention on outputs rather than outcomes. For instance, the outputs of university education may be school enrolments, or number of learners completing a particular level or grade. In the health sector output may be the number of discharged patients or inpatient days. However it must be emphasized that the outcomes should be based on how much students have learned in the case of education and how many patients recuperated enough to attain their former health. The difficulty involved in measuring outcomes objectively severely constraints most researchers from adopting this ideal concept.

Furthermore, due to the nature of the "outputs" generated by the public sector, the societal value of these services is also difficult to determine since there is no market transaction (Atkinson, 2005). In competitive markets, prices measure the consumers' marginal valuations or social values associated with the consumption of outputs. In the public sector, there are no final markets or prices to reveal consumers' marginal valuations of outputs, and so their value needs to be estimated. There are two ways of doing that: either measure the outputs and attempt to estimate the marginal valuations attached to them or measure the outcomes produced by each unit of output and attempt to estimate the marginal valuations of the outcomes (Dawson et al., 2004). Changes in the value of outcomes affect the allocation of resources within the public sector and the relative size of the public sector, and therefore should be taken into consideration when efficiency is being measured.

3.2.4.1.3 Weighting

Furthermore, while different outputs can be aggregated using observed prices as weights for private organisations that operate in competitive markets, however, for most public sector services, attaching weights to output has proved to be problematic. Firstly, this is because there are no prices for valuing outputs or outcomes in the public sector. Secondly, public sector organisations usually face multiple objectives and lack a consensus on the prioritisation of those objectives. Ultimately, the selection of objectives and the determination of their weights should be the responsibility of politicians who are charged with reconciling conflicting claims on public resources (Smith and Street, 2005).

The problem of market-price weights in the public sector can be overcome through a number of ways: expert opinion, client opinion, econometric frontier analysis and linear programming or data envelopment analysis. The judgemental weights of “expert” opinion (such as policy makers and practitioners) or of clients themselves can be used with caution. Adopting these subjective weights is not without its cost. In most cases it has been found that policy makers are either unwilling or unable to reveal policy-output priorities. On the other hand, regular investigations of client opinion on their valuation of public outputs are a rigorous exercise for most developing countries to undertake with reasonable precision. O`Mahony and Stevens, (2004) conclude that the use of “judgemental” weights is likely to be controversial at best and open to abuse at worst. The main purpose of assigning weights is essentially to attach societal values. These societal values then act as an indication of the relative worthiness of particular public outputs.

The parametric and non-parametric methodologies to be discussed in the later sections of this dissertation can be used to generate these weights. Under these two methodologies weights are generated as a by product of the estimation procedure as opposed to incorporating a pre-defined vector of weights. Some economists have regarded this feature an attractiveness of these methodologies (Cooper et al, 2000). The Data Envelopment Analysis (DEA[2]) treats the observed inputs and outputs as constants and chooses optimal values of the variable weights to maximise the efficiency of the production unit to the performance of the others. The optimal weights chosen for each production unit therefore represent a value-system which provides the most optimistic possible rating of that production unit relative to peer organisations (Ganley and Cubbin, 1992). This flexibility does not come without a cost as it weakens the conclusions that will be drawn about the relative efficiency derived. In effect, in small samples this may lead to some units being deemed efficient simply because they are different (in their input or output mixes) from other units. A possible way of overcoming this problem would be to impose

[2] DEA is a non-stochastic and non-parametric mathematical method that incorporates many inputs and outputs and enables an overall evaluation of technical efficiency. The methodology shall be discussed in the later sections in detail.

weights, but would bring us back to the problem of determining such weights removing the flexibility advantage of DEA.

3.2.4.1.4. Environmental Effects

The measurement of public service efficiency is further complicated by the influence of variables that lie outside managerial control (so-called environmental variables) - e.g. characteristics of individuals being served; external environment (geography, climate, and culture); and activities of other related organizations. Differences in these variables between organizations lead to differences in their production possibility frontiers (for each level of expenditure). The way in which the effect of these so-called environmental variables should be allowed for in measuring efficiency is often controversial (Coelli *et al.*, 1998). The Production (Cost) frontiers of organizations operating in less-favourable "environments" will lie inside (outside) those of more favourably endowed organizations. The effect of these variables needs to be taken into account when modeling efficiency otherwise efficiency will be over/underestimated. The way in which these variables should be included in the efficiency models is not consensual (Ozcan et al, 1992, Buck, 2000, Fried et a*l*., 2002). Environmental differences are particularly difficult to control. For instance, schools will differ in the average ability of children in their catchment area and ambulance service response times will vary according to population density. It is necessary to hold these factors constant in order to derive meaningful inferences about the determinants of efficiency.

3.2.4.2. Defining the "Best Practice" or the Efficient Frontier

With an appropriate set of inputs and outputs at hand, the next task is to define the efficient set of inputs and outputs against which an organization can be compared. This is, however, unknown. The procedure therefore is to compare an organization with an ideal comparator constructed from information on other organizations operating in the same field (and with similar size and environmental factors). In practice, this is often difficult, as similar organizations may be few and far between, especially in the public sector.

Measures of public sector efficiency can be derived by conceptualizing the administrative system as a “production unit”. A police department for example can be regarded as the production unit for security and public schools as production units for education. The overall relationship between the inputs and outputs of any production process can be summarized in a “production function”. The frontier production function represents the maximum level of output that can be obtained from a given level of inputs. The distance between a country’s actual level of goal attainment and the production frontier is called its “efficiency”. Efficiency of the public sector is simply:

$$Efficiency = \frac{Attainment - Minimum}{Maximum - Minimum}$$

The DMU that utilizes the fewest amounts of inputs to obtain a given level of output or the one that achieves maximum output with given level of inputs becomes the best practice. Every other production unit would have to be related to this best practice in order to establish by how many units it is deviating from the best practice. The more it deviates from the best-practice the less efficient it is. However it must be noted that this is a measure of relative efficiency as opposed to absolute efficiency. The production unit exhibiting “best practice” can be less efficient if it is placed in a different peer group with higher performers.

3.2.5. Efficiency in the Health Sector

Governments engage in the health sector in various ways, the most significant of which is in guaranteeing that the entire population is protected against the financial risks of sickness and medical treatment. The second most significant role is in the provision of medical services, by owning and operating medical care providers. In addition to intervening in the funding and provision of health services, governments can tax goods with adverse effects on health, and regulate the health sector – defining the operational framework of insurance companies; issuing licenses for medical care providers; and (dis)approving the commercialisation of new drugs and devices (Coelho, 2007).

There is an increasing recognition that improved health status contribute significantly to economic development. At the Millennium Summit in 2000, Member States of the United Nations reaffirmed their commitment to eradicate world poverty and improve the health and welfare of the world's poorest by 2015 (WHO, 2005). Health is at the centre of the Millennium Development Goals (MDGs). The achievement of the health-related MDGs among other things requires the availability of adequate resources for the health sector to improve access and quality of care. Therefore issues of efficiency need to be looked at before strategies to mobilize additional resources are considered.

Good health is important because it is an intrinsic element of human well-being. As a component of human capital, health is a key factor in the creation of wealth. In a study of the connection between health and wealth, Pritchett and Summers (1996) conclude that wealthier nations are healthier nations. A stressed healthy system and an unhealthy labor force do not augur well for any meaningful economic activity. As health is a form of human capital, its disruptions will inevitably dislocate all the fundamental links this sector has with the larger economy. This follows from the recognition that improved health status has a significant contribution to economic development (Grossman, 1972).

Health expenditure is also typically seen as an area where public intervention leads to better economic performance. Government intervention is justified both by equity and efficiency considerations. Health insurance markets are prone to failures due to problems created by information asymmetries such as moral hazard and adverse selection. These failures have a negative impact on the efficiency of the sector. Perverse incentives may be created, encouraging over/under consumption/provision of health services. Public sector provision and regulation of health services can mitigate some of these failures and help improve general health condition of the population thus creating a direct positive impact on human capital and hence on growth.

The most important role for government in the health sector, however, is in the market for health care and its subsidiary health insurance. Medical care and insurance markets are plagued with informational problems: moral hazard and adverse selection in health insurance; incomplete information on the part of patients in health care; asymmetric information between consumers and producers about patients needs; and inability to determine the quality of services, even after they have been provided (Cutler, 2002). These informational problems lead free competitive markets for medical care and insurance to failure. Health insurance markets are prone to moral hazard problems – people are likely to consume more health services when insured than they would do if they had to pay the full price of these services. In addition, health insurance markets are also known to suffer from adverse selection problems.

In medical care, informational problems also hinder the coordinative efficiency of free competitive markets. The presence of information asymmetries between patients and physicians as to the complexities of medical care diagnosis and treatment give market power to the physician thereby perpetuating what is called supplier induced demand (SID). A combination of profit-maximization and information imperfections contributes to disseminating perverse incentives that lead competitive markets to failure. Private for-profit suppliers will have an incentive to compromise on quality if consumers and/or government lack the regulatory and monitoring ability to detect such behaviour. Most empirical research in the health sector have demonstrated that the introduction of market-type mechanisms to public integrated health systems has positive effects on efficiency whereas further movement towards a market model of health care insurance and provision is shown to depress efficiency (Coelho, 2007).

The following section discusses SID and its impact on efficiency in the provision of health care.

3.2.5.1. Supplier-Induced Demand (SID) and Efficiency

According to economic theory, *ceteris paribus* an increase in suppliers of a service is expected to be followed by a decrease in the price of that service and therefore in the supplier`s earnings. While this holds for most markets it cannot explain the market for health care. Supplier Induced Demand (SID) describes the concern of policy makers and researchers about the adverse effects of financial incentives on provider behaviours. Such adverse incentives could be either overproduction or underproduction of desired services. In the case of overproduction, providers (physicians) may perform more procedures or activities than is needed to treat a condition. On the other hand physicians may under-provide a service depending on the payment method between the physician and the patient. SID is defined as the provision of services that consumers would not demand if they were fully informed. SID`s predominant concern is overprovision. Although overprovision may occur if more services means more gain to the provider (as in the case of fee-for-service[3] payment), under-provision may occur if delivering less service is in the providers interest (as in capitation[4] payment). The physician can induce the patient to use more or fewer services depending on the direction of the incentive.

Most of the services provided by the physicians are “induced” in the sense that physicians are acting as agents to consumers who lack the proper knowledge of the product, which in this case is healthcare. Only the demand that exists beyond what the well-informed patient would have chosen is defined as supplier induced demand (SID). The precise definition of SID is and has been open to debate over its existence and there isn’t a consensus over its precise definition. Below are two common definitions: Donaldson and Gerard (1993) defined supplier induced demand is the amount of demand created by doctors, which exists beyond what would have occurred in a market in which consumers are fully informed. Alternatively McGuire (2000) conceptualized supplier induced demand as existing when the physician influences a patient’s demand for care against the physician’s interpretation of the best interest of the patient. Therefore under this latter definition, the theory of SID says that health providers use their

[3] Payment based on itemized services or a charge per service provided.
[4] Payment based on the number of individuals registered with a service provider

superior knowledge to influence demand for self-interests. Physicians therefore have the ability to generate demand in response to fee changes, declining market shares, or simply changes in the labour-leisure choices. The most well known observation related to the SID hypothesis is the positive correlation between per capita health care utilization and physician density, that is, the greater the number of suppliers of physicians into the market the greater the health utilization per capita among the patients whether the new equilibrium price increase or decrease.

SID imposes a cost on the organization and the society at large and as a result it has serious repercussions on organizational and social efficiency. In the case of over-production, induced demand implies that more resources are wasted when the physician recommends more consumption of medical services. For instance a patient may not necessarily have need for X-ray examination or a certain complex operation. However, the doctor will recommend it if it means more revenue on his part. Thus SID erodes the consumer surplus and distorts the computation of efficiency scores. In empirical research of efficiency measurement researchers have not been able to successfully measure SID objectively despite their endeavor to incorporate it into the model of efficiency.

3.2.5.2. Health Care and Profit Incentives

One of the influential theories has been proposed by Paul et al (1973) who suggest a traditional neo-classical model where hospital clinicians[5] attempt to maximize residual profits essentially their own income thereby performing the dual role of manager and entrepreneur. The financial goal of the for-profit hospital is generally understood as to earn a profit while the financial goal of the nonprofit is to breakeven. Earning a profit is stated as one of the major goals for the for-profit hospitals but not for the nonprofit hospital. Modern health economists have argued that nonprofit hospitals do have an incentive to earn a profit but the more important difference concerns the different ways the two entities treat profits. The for-profit hospitals turn over any profit to its owners as private income while the non-profit hospital keeps any profit or surplus to

[5] Clinician is a term used to describe medical professionals such as surgeons, physicians, nurses, dieticians, or X-ray technicians including veterinary medicine who are engaged in actual patient care as opposed to researchers and academicians

expand services or its equity endowment (Preker, 2007). It is this distinction in the treatment of profit between for-profit and non-profit hospitals which forms the predominant basis of this study. It has been argued by economist that the fact that individual income under non-profit hospitals is not related to performance is a disincentive to minimize cost and attain efficiency. The hospital industry world over has been undergoing significant structural changes in the way in which health care is provided. One important structural change has been the growth of for-profit hospitals. Health economists have suggested that a continued movement in the direction of profit incentives in the production of hospital care may improve the industry`s performance. These individuals arguing from the theory of property rights believe that non-profit is inherently inefficient because no individual`s income is tied to economic performance. Furthermore, it is believed that the non-profit hospital manager may cause the hospital to pursue goals other than strict cost minimization.

3.2.5.3. Market Approaches and Health Care

Market approaches are taken here to be reforms that aim to introduce competitive pressures into public health services. The starting point for the discussion is some of the European health reform that aims to introduce competition. Those who advocate for provider markets argue that they generate both substantial increases in efficiency and increased consumer choice and influence over health services. The means by which these beneficial outcomes are considered to occur are twofold (Broomberg 1994). Firstly, it is argued that provider markets will give rise to competition amongst providers for contracts, and that competition will enhance efficiency on the supply side. Secondly, the replacement of direct management by contractual relationships between purchasers and providers is argued to promote increased transparency of prices, quantities and quality in trading, as well as managerial decentralisation, both of which will also enhance efficiency. To achieve the ideal outcome of a quasi-market, namely efficiency, provision of choice and responsiveness to consumers, it is argued that there must be competition on both the purchaser and provider side (Bartlett and Le Grand, 1993), though there is some debate as to whether competition on the purchaser side is as important as competition on the provider side.

Encouraging providers to compete for consumers requires a system for rewarding providers who attract more customers, and hence for paying them for providing care to identifiable consumers. It generally requires public providers to charge cost-covering fees. It has been argued by pro-competition economists that there is unfair competition in developing countries, which hampers the development of the private medical sector, because public services are provided free or highly subsidized in most cases. The fact that most governments compensate public hospitals for delivering services does not put these non-profit hospitals on the same level playing field with their for profit counterparts. Hence charging fees in the public sector is argued to be necessary to promote competition. On the purchaser side the idea of creating competition on the purchaser side is dependent on the people's ability to behave as informed consumers. In developing countries large proportions of the population are uneducated or poorly educated, have limited access to sources of information such as the mass media and newspapers, and are distant from providers in terms of socio-economic status and educational level.

Health care economists have suggested the following wide variety of market approaches: Inviting bidders to build and provide primary and secondary level services for the general public; Engaging the private hospital sector in contracts for a variety of different types of care undertaken by the public health providers such as primary care, whole hospitals, particular types of patient care, particular diagnostic procedures; Offering public facilities to private sector entities to run on a long term lease, involving temporary transfer of ownership; offering management contracts to the private sector to run public facilities; encouraging contracting out of non-clinical services in public hospitals and contracting for delivery of disease control activities such as immunization.

This section of the study on health service delivery has demonstrated the importance and benefits of efficient utilization of health public sector resources on economic growth. The discussion revealed that market failures at the service provision level generally stem from: information asymmetries that cloud the ability of consumers and government to effectively assess the quantity and quality of service provided; Lack of market approaches or dominance of market power following monopolistic and oligopolistic market structures. Particularly a lack of both provider and purchaser competition was identified as a necessary stimulant for efficiency. These failures requires the public sector to play an active role in the provision of the service, either by replacing , supplementing , or regulating private provision as well as enacting policies to inject market dynamics into public institutions. While profit incentives were seen theoretically as having a positive bearing on efficiency empirically the debate is inconclusive. It is this question of the impact of profit incentives on efficiency which is the primary objective of this dissertation.

3.2.6. Public Sector Efficiency Measurement

Having presented the theoretical basis to measuring efficiency for a single "production unit", this section reviews the statistical techniques to evaluate efficiency among various "production units" or DMUs. Efficiency in econometric terms can be estimated either as a "deterministic" frontier, or as a "stochastic" frontier. The deterministic method presumes that all deviations from the frontier are attributed to inefficiency. In contrast, the stochastic approach assumes that some deviation from the frontier is attributed to random factors (e.g. natural disasters or exogenous factors) and others to inefficiency.

Building upon the work of Farrell (1957), two main methods have dominated the empirical measurement of efficiency. These are; the non-parametric method known as Data Envelopment Analysis (DEA) and the parametric method known as the Stochastic Frontier Approach (SFA) (T. Coelli, 1996). Non-statistical methods such as DEA tend to be deterministic, whereas statistical methods, such as SFA tend to be stochastic, allowing for statistical "noise". Several studies have sought to compare DEA and SFA. There is no consensus on whether DEA or SFA is the best tool for efficiency measurement (Folland, 2001). Reliable efficiency measures are considerably important to policy makers and public service managers. First, they draw attention to the fact that it may be possible to achieve a higher level of goal attainment without increasing input resources. Second, with the measurement of efficiency, it is possible to investigate exogenous determinants of inefficiency. And third, the regular measurement of efficiency over time is essential for an effective monitoring of public management reforms which aim to increase efficiency (Tandon, p 683).

3.2.6.1. Stochastic Frontier Approach (SFA)

The SFA is a parametric methodology which adopts an econometric approach. The econometric approach specifies a production function and normally recognizes that deviation away from this given technology (as measured by the error term) is composed of two parts, one representing randomness (or statistical noise) and the other inefficiency. The main problem in measuring inefficiency is to separate genuinely inefficient behavior from other random factors affecting costs or profits. The SFA proposes that the observed costs of a hospital may deviate from the costs frontier either because of random fluctuations' or because of inefficiency.

The stochastic frontier production function which was proposed by Aigner et al (1977) can be expressed as:

$$y_i = x_i \beta + (v_i - \mu_i)$$

y_i : average maximum output.

x_i : a vector of input quantities of unit *i*.

β : a vector of unknown parameters.

V_i : are random variables which are assumed to be *i.i.d*[6] N $(0, \delta^2)$ and independent of μ_i.

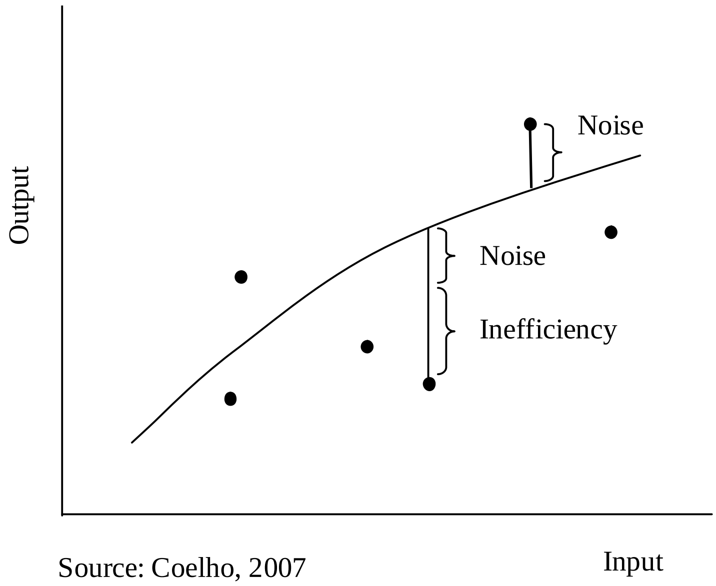
μ_i : are non-negative random variables and assumed to account for inefficiency.

The random error term is generally thought to encompass all events outside the control of the organization, including both uncontrollable factors directly concerned with the "actual" production function (such as differences in operating environments which are exogenous and therefore outside the control of the production system) and econometric errors (such as misspecification of the production function and measurement error). The SFA requires that the sample size be sufficiently large to avoid problems of degrees of freedom. The efficiency

[6] Independent and identically distributed.

measures are computed in terms of the distance that lies between the observation and the estimated function. Its biggest advantage lies in the fact that it introduces a disturbance term representing noise, measurement error, and exogenous shocks beyond the control of the production unit.

Figure 3.7: Stochastic Frontier Approach (SFA)



Source: Coelho, 2007

3.2.6.2. Data Envelopment Analysis (DEA)

DEA is a linear programming technique initially developed by Farrell (1957) and later by Charnes, Cooper and Rhodes (1978) to evaluate the efficiency of public sector non-profit organizations. It was originally intended for use as a performance measurement tool for organizations that lacked a profit motivation, for example, governmental organizations such as public schools and hospitals. However, since its introduction it has been developed and expanded for a variety of uses including application in for-profit institutions.

DEA is a "non-stochastic non-parametric mathematical method that incorporates many inputs and outputs and enables an overall evaluation of technical efficiency" (Dalmau-Matadora *et al*, 1990). The DEA method was first developed by Charnes, Cooper and Rhodes in 1978 (Coelli, 1996). The model had an input orientation and assumed constant returns to scale (CRS). Subsequent development of the model by Banker, Charnes and Cooper (1984) gave origin to the variable returns to scale (VRS) model. The DEA is *non-parametric* in the sense that it does not assume that the underlying technology "belongs to a certain class of specific functional form which depends on a finite number of parameters such as the well-known Cobb-Douglas functional form" (Diewert *et al*, 1983). This is important because as Bowlin (1986) has argued: "the functional relationships underlying public production may be unusually complex and difficult to specify". It is also *non-statistical* in the sense that it makes no explicit assumption on the probability distribution of errors (Sengupta, 1987a).

In principle the DEA provides an overall performance index which eschews several common pubic sector measurement difficulties. DEA has a number of advantages over the Stochastic Frontier Approach:

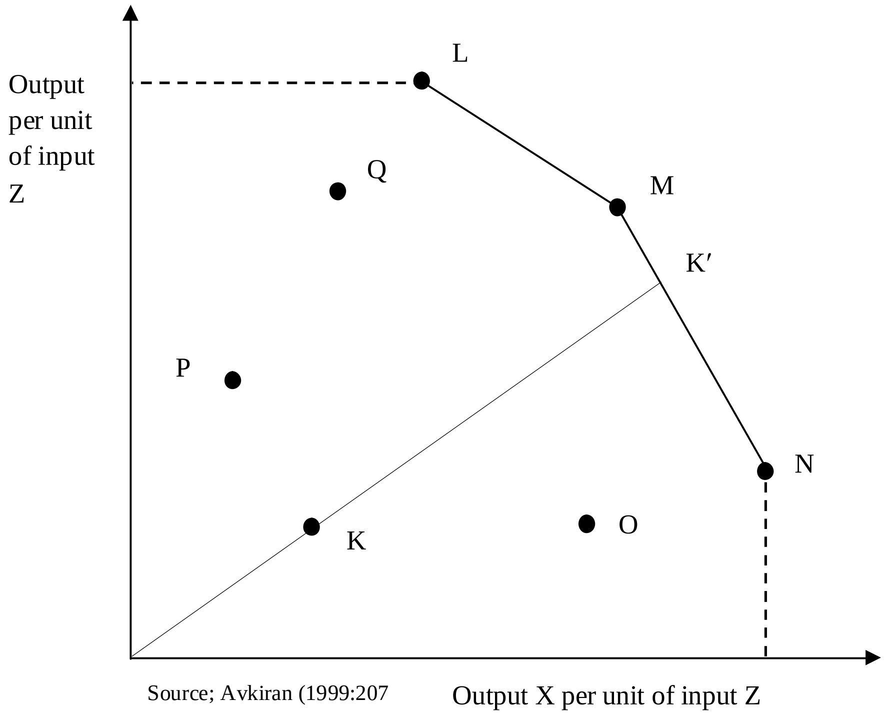
- It does not require explicit specification of functional relations between inputs and outputs (as in regression approaches) and does not also require the decision maker to express his own weighting scheme for inputs and outputs (as in index number approaches).
- The DEA approach is unit invariant. In normal circumstances the DEA relative efficiency coefficient is derived unaffected by units of measurements in the underlying data. According to T.Coelli (1996, 23) "changing the unit of measurement e.g. measuring quantity of labor in person hours instead of person years will not change the value of the efficiency measure."
- The DEA can accommodate the presence of multiple outputs much more readily than can parametric models of production, a useful property for analyzing health care institutions. Hospital "production" is characterized by multiple inputs and multiple outputs.
- The DEA does not require price data which are difficult to obtain for hospital inputs and outputs; rather DEA requires only data on inputs and outputs.

In general, DEA compares a set of organization's actual input used to produce their output levels during a common time period. The technique locates those units that are relatively more or less efficient compared with the most efficient ones in the set (Sherman, 1984). D.E.A determines the following:

- The best-practice and the most productive group of service units
- The less-productive service units compared to the best-practice units.
- The amount of excess resources used by each of the less-productive units
- The amount of excess capacity or ability to increase service outputs in less-productive units without utilizing added resources.
- The set of best practice service units most similar to the less productive units.

DEA can be illustrated by the illustration in Figure 3.8 below. Figure 3.8 shows a DEA model with a solid line connecting the efficient decision making units (DMUs) L, M and N that represent achieved efficiency. For instance DMU K is classified as inefficient and needs to move to K′ on the frontier to be classified as efficient.

Figure 3.8: A DEA Model showing an efficiency frontier



Source; Avkiran (1999:207

Table 3.1 below contains a summary comparative analysis of the main tools for efficiency measurement; the Data Envelopment and the Stochastic Frontier Approach.

Table 3.1: A Comparison of Data Envelopment Analysis and Stochastic Frontier Approach.

Category	**Data Envelopment Analysis (DEA)**	**Stochastic Frontier Approach (SFA)**
Description	A linear programming method that constructs a nonparametric production frontier by fitting a piece-wise linear surface over the data points.	An econometric method that estimates a production frontier of the form: $y = f(x) + v - u$, where y is the output f(x) are all the inputs, v is an error term capturing unpredictable perturbations and u captures technical inefficiency. A cost frontier (short run or long run) or a distance function can be used.
Data needs	Quantity data on inputs and outputs for a sample of firms. If price data are available, one can use it to calculate allocative efficiency.	For a production function or distance function: quantity data on inputs and outputs for a sample of firms, ideally over a number of years. For a long-run cost frontier: total costs, input prices, and output quantities. For a short-run cost frontier: variable costs, variable input prices, and fixed input quantities and output quantities.
Advantage	Identifies a set of peer firms (efficient with similar input and output mixes) for	Attempts to account for noise. Environmental variables are easier to deal

	each inefficient firm. Can easily handle multiple output Does not assume a functional form for the frontier or a distributional form for the inefficiency error term.	with. Allows for the conduct of traditional statistical tests of hypotheses. Easier to identify outliers. Cost frontier and distance function can deal with outputs.
Drawbacks	May be influenced by noise Traditional hypothesis tests are not possible. Requires large sample size for robust estimates, which may not be available early on in the life of a regulator.	The decomposition of the error term into noise and efficiency components may be affected by the particular distributional forms specified and by the related assumption that error skewness is an indication of inefficiency.

3.3. EMPIRICAL LITERATURE

3.3.1. Empirical Studies

In this section of the study a brief empirical literature review relating to hospital characteristics and the economic performance of hospitals is presented. This review is intended to show the types of hypothesis previously considered and the inconclusive nature of previous work in the area of hospital efficiency. Efficiency studies in developing countries and especially in Africa are very few. The available empirical literature has been enriched with studies mainly from the developed world. Furthermore, a significant number of researchers have confined their studies on the public health facilities rather than the private health sector. This is mainly explained by the reluctance of the private sector in exposing information governing their activities. The empirical evidence relating to the performance of public and private hospitals falls into three categories: studies finding that private hospitals apparently perform better; that public hospitals apparently perform better; and that there is no significant difference in their performance.

Comparisons between frontier efficiency measurement techniques have been made in studies. For example, Gonzalez Lopez-Valcarcel and Barber Perez (1996) compared the DEA-based technical efficiency measures with stochastic frontier cost efficiency indexes in a sample of Spanish general hospitals. Linna et al (1998) also examined the DEA measures and stochastic frontier estimates of cost efficiency in Finnish care hospitals. Both studies concluded that the choice of approach did not significantly influence the results. Ganley and Cubbin (1992) also investigated these two alternative research methods and concluded that the DEA and SFA “are in substantial agreement on several important issues.” Bowlin et al (1985) compared theDEA with the SFA for 15 hypothetical hospitals. The comparison revealed that both methods performed well in discriminating between efficient and inefficient hospitals but that the DEA offered the additional advantage of being able to identify the sources of inefficiencies by highlighting which resources were being used in excess. Sherman (1985) extended the analysis of the 15 hypothetical hospitals and showed that the DEA offered more accurate estimates of relative efficiency and target levels, but that regression analysis estimates were more stable.

Thanassoulis (1993) investigated the comparison of Regression Analysis (RA) and Data Envelopment Analysis (DEA) as alternative methods for performance assessment. The same data set which relate to 15 hypothetical hospitals used by Bowline et al (1985) were adopted in the study. Output variables consisted of the number of teaching units, regular patients and severe patients. The inputs of the hospitals were reflected by the total cost of the resources it used in pursuit of its output. The comparison revealed that the DEA offers on the whole, more accurate estimates of efficiencies, marginal values and targets. One further advantage of DEA that was identified in the study was that its estimates of marginal values and target levels are not affected by correlations and multi-collinearity between inputs or outputs.

3.3.1.1 Non-Profit Hospital Superiority

It is a common belief that private health institutions are more efficient than public health institutions. However, this belief is not necessarily correct according to some empirical evidence. A similar study by Charles et al (1987) was conducted and the DEA methodology was applied to determine the relative efficiency of for-profit and non-profit hospitals. The primary objective of their study was to investigate whether profit incentives had any significant impact on efficiency as postulated by theory. The sample of 457 United States hospitals classifies into 300 non-profit, 121 for-profit and 36 government-controlled hospitals. They estimated whether a relationship existed between technical efficiency and ownership form (profit incentive). Their analysis controlled for hospital size, case-mix, and geographical region. The study emphasizes the need to control for case mix because hospitals treating a relatively more complex mix of patient cases may require more inputs per unit output. The mean technical efficiency scores for non-profit, government and for-profit hospitals were 74.52, 73.67 and 66.07 respectively, indicating that non-profit hospitals possessed the greatest range of efficiency scores and accounted for the greatest number of the top 34 most efficient hospitals. For-profit hospitals accounted for 55 % of the least efficient hospitals. Findings from regression analysis revealed that no significant relationship existed between ownership form and technical efficiency. Competition as measured by the Herfindahl index (Concentration ratio) did not significantly influence efficiency. The only variable that was significantly related to efficiency was the hospital size variable as proxied by the number of hospital beds. The authors made a conclusion that for-profits are not more efficient than the not-for-profit hospitals as suggested by economic theory.

Valdmaris as cited by Folland (2001) applied DEA to two groups of hospitals in Michigan. One group represented 33 private (for-profit) health providers and the other group included 25 public hospitals. The inputs used were as included: Inpatient days, number of physicians, number of registered nurses and number of hospital beds. Outpatient visits and Total discharges. The regression model was estimated to control for factors such as quality variations, competition and hospital size. The average technical efficiency scores were 86.6 % and 98.5 % for private and public hospital respectively (Folland, 2001). Public hospitals performed better than private hospitals. Public hospitals consumed 11.9 % fewer resources than private hospitals. Regression results indicated that with the exception of competition the size of a hospital and quality of care were significant factors of efficiency.

Zere (2000) examined hospital efficiency in three provinces of South Africa. His study evaluates the technical efficiency and productivity of a sample of public sector hospitals using the non-parametric techniques of DEA and DEA-based *Malmquist Productivity Index* (MPI). A tobit regression was also estimated to control for potential “contaminants” and to identify those factors that may be associated with (in)efficiency. The sample consisted of 86 hospitals classified into three levels: Community hospitals with emergency services only (Level I), community hospitals with outpatient services (Level II) and non-academic secondary and tertiary hospitals (Level III). Recurrent expenditure and bed size were used as inputs. Outputs include inpatient days and outpatient visits. The results indicated that there was a marked deviation of performance among hospitals within each level. An average overall technical efficiency of 0.74, 0.68 and 0.70 was computed for Level I, II and III hospitals respectively. This implies that on average inefficient hospitals consume 26 % – 32 % more resources. Most hospitals operated at a non-optimal scale with decreasing returns to scale dominating in Level II and III hospitals. The study demonstrated that occupancy rate is an efficiency determinant factor that is positively related to efficiency whereas length of stay is adversely related to efficiency.

Mbalame (2002) employed the DEA methodology to compare and estimate the efficiency levels at which Mozambican health units operated. The sample size consisted of 30 general hospitals and 104 health centers. Hospitals were assumed to produce 5 outputs; outpatient visits, inpatient days for medicine, inpatient days for paediatrics, inpatient days for surgery, and inpatient days for other specialties using five inputs; number of beds and health personnel classified as superior, medium, basic and elementary. The VRS DEA results indicated a mean efficiency score of 93.37 % and 80.82 % for general hospitals and health centers respectively. The results showed that health centers were relatively more inefficient than general hospitals.

Zere *et al* (2006) conducted a study to evaluate a sample of 30 public hospitals in Namibia using the Data Envelopment Analysis technique. The DEA model used three inputs: total recurrent expenditure, beds and nursing staff and two outputs: total outpatient visits and inpatient days. To test for the robustness of the DEA technical efficiency scores the Jackknife[7] analysis was used. Results for the constant returns to scale (CRS) DEA models estimated for the period 1997/98 to 2000/2001 indicate average technical efficiency scores ranging from 62.7% to 74.3%. The jackknife analysis indicates that the stability of the estimates and that the efficiency frontier has not been affected by extreme outliers (Spearman rank correlation coefficient = 0.99). The CRS technical efficiency scores reveal combined inefficiency that is due to both pure technical inefficiency and inefficiency that is due to inappropriate hospital size. Increasing returns to scale was the predominant form of scale inefficiency observed. The study further revealed that the prevalent scale inefficiency is increasing returns to scale. In the presence of increasing returns to scale, expansion of outputs reduces unit costs and thereby increases efficiency. However, increasing the level of outputs requires an increase in the demand for health care, which is

[7] To test for the robustness of the DEA technical efficiency scores, the Jackknife analysis was used. In the jackknife analysis, a limited number of samples are obtained by omitting one observation at a time. In this case the efficient hospitals are dropped one at a time from the analysis and the efficiency scores re-estimated. The similarity of the efficiency rankings between the model with all the hospitals included and those based on dropping each of the efficient hospitals is then tested by using Spearman rank correlation coefficient. A correlation coefficient of 1 implies that the rankings are exactly the same. A value of zero indicates the absence of correlation between the rankings and reverse ranking is implied by a value of -1.

beyond the control of the hospital management. Merger of hospitals whose location is close to one another to one another may be an option worth of consideration.

3.3.1.2 For-Profit Hospital Superiority

Wilson and Jadlow (1982) examined the relative efficiency of for-profit and non-profit hospitals in the provision of nuclear medicine services. The mode of evaluation employed by Wilson and Jadlow was the non-parametric DEA methodology. This analysis revealed how close each type of hospital came to producing the maximum possible amount of output given its choice of input mix. In addition to comparing non-profit and for-profit private entities, they also evaluated public facilities. They discovered that in terms of technical efficiency (maximum output for given inputs), government providers performed worse than private non-profits, which performed worse than for-profits. The average efficiency scores were 83.3 %, 87.2 % and 92.7 % for public, private non-profit and for-profit hospitals respectively. They found that for-profit nuclear medicine services were significantly closer to maximum production than non-profits were. The authors conclude that institutional changes in the direction of profit incentives are likely to improve the performance of hospital care provision.

A study along similar lines by Nyman and Bricker (1989) employed the DEA to investigate the impact of profit incentives on the technical efficiency in the production of nursing home care. The evidence in their study suggested that for-profit nursing homes had significantly higher scores than non-profit homes. According to their study, for-profit nursing homes used about 4.5 % fewer labor resources per patient day than non-profit homes. Both studies support the property rights hypothesis that for-profit homes are inherently more efficient than non-profit ones. The authors identified the following crucial factors as key drivers of efficiency:

i) The *status* of firm. For-profit nursing homes have explicit reasons for minimizing costs, hence they have reasons for wanting to produce efficiently.

ii) The *reimbursement policy* associated with the clients. If nurses are reimbursed simply for costs incurred and given a certain return on capital independent of present period behaviour, there is no reason to minimize costs even for profit maximizing firms. This is because in such a case profit cannot be raised by minimizing costs.

iii) *Occupancy rate*. The occupancy rate is the number in the home on a certain day divided by the actual number of beds. It is assumed that nursing homes tend to staff 100% occupancy rates, then the degree to which the firm's actual occupancy rate is less than this target occupancy rate will have an effect on the firms staffing hours per patient day.

iv) *Patient case mix* has an effect on resource use. It is important to hold constant this factor if meaningful inferences are to be used.

v) *Quality* of care is another dimension of output that needs to be held constant. Greater quality care requires more inputs per unit of output, to the extent that higher quality service providers may have lower efficiency scores not because they are less efficient but because they provide better services.

vi) Finally the authors identified the need to introduce *competitive dynamics* as a remedy to the market failure in health sector, a problem normally created by the presence of oligopolistic and monopolistic structures

Vincenzo and Dino (2006) investigated the levels of efficiency by adopting the Data Envelopment Analysis non-parametric method. A large sample of 85 hospitals in Italy divided into 61 public and 24 private (7 non-profit and 17 for-profit) was considered. The author's model consisted of three outputs namely:

- Total care discharges.
- Number of days of treatment in hospital known as inpatient days.
- Number of treatments provided by emergency services.

The following five inputs were considered in the model.

- Number of physicians
- Number of Nurses
- Number of other employers.
- Number of hospital beds as a proxy for capital
- Total admissions as a proxy for hospital demand.

They distinguished between three components of technical inefficiency of hospitals: *internal* inefficiency attributable to the responsibility of hospital management, *external* inefficiency that could be due to past health care policy decisions and to exogenous demand and *scale* inefficiency which is due to under or over sizing of hospitals with respect to their actual activity levels. Their study revealed that non-profit private hospitals exhibited a level of total inefficiency higher than public hospitals. Private for-profit hospitals produced the least total inefficiency scores. However, both non-profit and for-profit hospitals were characterized by higher scale inefficiency than public hospitals. Under the VRS returns to scale model, the results indicated that 23 of all public hospitals exhibited decreasing returns to scale (DRS), 42 hospitals of which 80% were private hospitals exhibited increasing returns to scale (IRS) and finally 20 hospitals exhibited constant returns to scale (CRS). It was concluded that the problem of scale inefficiency mainly characterized the private hospital sector with many hospitals small in relation to their output levels. Private for-profit hospitals had the highest score of total efficiency than the public and the non-profit hospitals.

Fizel and Nunnikhoven (1992) examined nursing home efficiency in the state of Michigan. Their study attempted to determine the technical efficiency of for-profit and non-profit nursing homes to shed some light on the debate about the validity of the theory of property rights. The study examined a sample of 163 nursing homes classified into 104 for-profit homes and 59 non-profit homes using the non-parametric techniques of Data Envelopment Analysis. A regression analysis was estimated to identify some factors that may be associated with (in)-efficiency such as quality variations, ownership status and competition. Inpatient days and intermediate care patients were used as outputs whereas number of hours of registered nurses, aides and orderlies were captured as inputs. The number of beds was not included in the model. The DEA results computed an overall efficiency score of 0.68 and 0.48 for for-profit and non-profit homes respectively. The for-profit nursing homes used 20 % fewer resources than the non-profit homes. Using a property rights framework, the authors theorized that since for-profit homes have exclusive rights to income generated, with the resulting incentive to meter input productivity and rewards. Given the threat of take-overs, an incentive existed to produce efficiently. On the other hand, in a non-profit home the owner’s rights to income are attenuated (and ultimately non-transferable). The study highlighted the need to exercise caution when interpreting efficiency scores derived where the quality dimension of output has been assumed to be constant. This is because higher quality can translate in a lower efficiency score due to the additional resources required to improve quality.

Mills and Liu (1998) applied the DEA methodology to 62 private and public general hospitals in Korea. The hospitals were assumed to use the following inputs; beds, doctors, registered nurses, nursing aides, pharmacists, technicians and administrative staff to produce 16 outputs. The outputs were basically the various health specialties provided by the hospitals. Among other health specialties the outputs included internal medicine, pediatrics, general surgery and gynaecology. The DEA results showed that public hospitals were more scale inefficient than private hospitals. The average inefficiency scores were 23 % and 32.6 % for public and private hospitals respectively. The comparison between hospitals size and the overall technical efficiency scores showed a positive relationship between hospital size and efficiency scores. The optimal hospital size was estimated using the number of beds as a proxy for hospital size.

3.3.1.3 Non-Differential in Efficiency.

A study to test the hypothesis that for-profit hospitals were efficient than non-profit hospitals was undertaken by Register et al (1985). They considered the economic behavior of a sample of for-profit and non-profit hospitals in Oklahoma. According to their findings, there appeared to be no significant differences in the performance of the two hospital types. Thus the authors conclude that the two hospitals are equally efficient.

A similar study published in 1998 by Zelder (1998) applied DEA to a sample of 62 private and public hospitals in Canada. They tested whether these non-profit and for-profit hospitals operated according to the same production process, reasoning that if there was no significant difference in the manner in which inputs were combined in production, and then there could be no difference in efficiency. The numbers of beds were used as capital among other inputs such as the number of nurses, doctors, and other staff while total inpatient and out-patient visits and total discharges (excluding deaths) represented output. Their analysis indicated that no significant difference existed in the production technologies of non-profit and for-profit hospitals.

Another similar study was by Vitalino and Toren (1993), where the SFA was applied to a sample of for-profit and not-for-profit hospitals in the United States. The study was based on panel data collected on a two-year period. The study revealed a 29.6 % average level of cost inefficiency. They also found out that there was no change in efficiency between 1987 and 1990. The striking feature of their findings was that efficiency did not vary between for-profit and not-for-profit hospitals.

3.4. CONCLUSION

This chapter reviewed various conceptualized theoretical, and empirical evidence on the crucial subject of efficiency in the public sector. Critical areas in the health sector were explored. The health sector was selected because of its critical impact on the society as whole and on economic growth in particular. The most critical decisions in measuring efficiency involve the choice of the set of outputs to be measured and the choice of the set of weights reflecting the societal values attached to each output. In the public services, both of these are essentially political choices (Coelho, 2007). The quality of the results will also depend on the extent to which all the variables that significantly influence the productive process are taken into account and whether their influence is correctly modeled. Two common research methods of performance assessment the DEA and the SFA were discussed and their comparison was made. The comparison revealed that on the whole DEA shows more accurate estimates of efficiencies, marginal values and targets and that DEA offered the additional advantage of being able to identify the sources of inefficiencies by highlighting which resources were being used in excess. It is important however, to note that one should not lose sight of their agreements. There are in agreement on several substantial areas. Where possible both methods should be used and any disagreements on estimates should be analyzed in a way which throws light on the performance of DMUs (Thanassoulis, 1993). Finally the empirical literature has demonstrated the inconclusiveness of the subject of profit incentives on efficiency in health service delivery.

CHAPTER FOUR

METHODOLOGY

4.1 INTRODUCTION

This chapter discusses the methodology of the study. This encompasses a brief synopsis of the DEA methodology as well as the mathematical formulation of the DEA, different models that can be adopted, definitions of the variables that are applicable to the Zimbabwean situation and the model specification.

4.1.2 Mathematical Formulation of the DEA

In organizational management, comparative performance evaluation is crucial. Organizational performance is normally evaluated by making comparisons with other organizations in the same industry. Traditionally the efficiency of hospitals has been measured using ratio analysis such as doctor-patient ratio, cost per patient, and cost per day. However, ratio analysis is limited to two factors only, that is, to one input and one output cases. This inadequacy of ratio analysis gave way to the modern DEA which is a powerful aggregate comparative method for assessing the efficiency of organizations with multiple incomparable inputs and outputs. The general objective of DEA is to maximize the ratio of weighted outputs to weighted inputs for a particular DMU subject to the constraint that this ratio is less than or equal to one. The constraint is imposed because it is not possible that any DMU[8] can be more than 100% efficient.

[8] "Decision making units". It is a production unit or an entity for converting inputs into outputs such as banks, hospitals or schools especially when studying the performance of its branches.

According to Emrouznejad (2007), Efficiency is generally measured as:

$$Efficiency = \left(\frac{Output}{Input} \right) \qquad \text{.......................................4.1}$$

The above equation was modified because this research is using data which is made up of multiple inputs and outputs. The efficiency of a many-input, many-output decision making unit (DMU) is defined as the weighted sum of its outputs divided by a weighted sum of its inputs. Hence, equation 4.1 can be rewritten as:

$$Efficiency = \left(\frac{weighted\ sum\ of\ outputs}{weighted\ sum\ of\ inputs} \right) \qquad \text{...........4.2}$$

4.1.2.1 The Input-Oriented Constant Returns to Scale Formulation.

The input-oriented CRS model which was initially proposed by Charnes, Cooper and Rhodes (1978) has an input orientation and assumes constant returns to scale. The mathematical formulation of DEA assumes data on K inputs and M outputs on a sample of N firms or DMUs. The inputs and outputs for the *i*-th DMU can therefore be represented by the vectors x_i and y_i, respectively or more technically by an $K \times N$ input matrix and $M \times N$ output matrix. The purpose of DEA is to construct a non-parametric envelopment frontier over the data points such that all observed points lie on or below the production frontier (Coelli, 1996). For each *i*-th DMU there is need to obtain a measure of the ratio of all outputs over all inputs, such as $u'y_i / v'x_i$, where u is a $M \times 1$ vector of output weights and v is a $K \times 1$vector of input weights. According to Coelli (1996), the mathematical formulation for the Input-Oriented CRS DEA Model which will generate optimal weights is as follows:

$$\max_{u,v} (u'y_i / v'x_i),$$
$$st \quad u'y_j / v'x_j \leq 1, \ j = 1, 2, ..., N,$$
$$u, v \geq 0 \qquad \text{4.3}$$

Equation 4.3 was used to select the optimal weights which are the values of u and v such that the efficiency measure of the *i*-th DMU is maximized subject to the constraint that all efficiency measures must be less than or equal to one. The problem with this equation is that it has an infinite number of solutions, hence, to avoid this, a constraint $v'x_i = 1$ can be imposed thereby providing the following formulation:

$$\max_{\mu,\upsilon} (\mu' y_i),$$
$$st \quad \upsilon' x_i = 1,$$
$$\mu' y_j - \upsilon' x_j \leq 0, \ j = 1,2,...N,$$
$$\mu, \upsilon \geq 0, \qquad \text{4.4}$$

The notation change in u and v to μ and υ represent the transformation to a form which is known as the multiplier of the linear programming problem. An equivalent envelopment formulation can be obtained using duality linear programming as formulated in (4.5) below. The duality of (4.4) constructs a piecewise linear approximation to the true frontier by minimizing the quantities of *x* inputs required to meet stated levels of the *y* outputs (Ganley and Cubbin, 1992). The dual equivalent envelopment form will be:

$$\min_{\theta,\lambda} \theta,$$
$$st \quad -y_i + Y\lambda \geq 0,$$
$$\theta x_i - X\lambda \geq 0,$$
$$\lambda \geq 0 \qquad \text{4.5}$$

where θ is a scalar and λ is an N x 1 vector of constants. This envelopment form is more preferable than the equation 4.4 because it has got fewer constraints. It is equation 4.5 that is used to calculate the efficiency scores. The value of θ will be the efficiency score for the *i*-th DMU and should satisfy $\theta \leq 1$. A value of 1 indicates a point on the efficiency frontier and hence a technically efficient DMU. The linear programming problem is solved N times once for each DMU in the sample. To check whether a DMU with an efficiency score 1 is indeed fully efficient, the DEA model includes the calculation of input and output slacks. Therefore equation 4.5 is defined as follows:

$$\begin{aligned} & \min_{\lambda, OS, IS} -(M1'OS + K1'IS), \\ st \quad & -y_i + Y\lambda - OS = 0, \\ & \theta x_i - X\lambda - IS = 0, \\ & \lambda \geq 0, OS \geq 0, IS \geq 0 \end{aligned} \qquad \text{.............. 4.6}$$

OS is an M x 1 vector of output slacks, IS is a K x 1 vector of input slacks and M1 and K1 are M x 1 and K x 1 vectors of ones, respectively. This second stage linear program is also run for all the N DMUs under study. The multi-stage DEA which is in preference to the two-stage was used in the study. The multi-stage DEA method is more computationally demanding but its merit is that it identifies efficient projected points which have input and output mixes which are as similar as possible to those of the inefficient points (Coelli, 1996).

4.1.2.2 The Input-Oriented Variable Returns to Scale (VRS) Formulation.

The CRS assumption is only appropriate when all DMU`s are operating at an optimal scale, that is, at the lowest point of the long run average cost curve (LRAC). The prevalence of imperfect competition and financial constraints, for instance, may cause a DMU not to operate at an optimal scale. Banker, Charnes and Cooper (1984) suggested an extension of the CRS DEA model to account for variable returns to scale (VRS) situations. Adopting the CRS specification when all the DMU under study are not operating at an optimal scale will result in measures of technical efficiency which are “mixed up” with scale efficiencies (SE). Thus the VRS specification will allow the generation of technical efficiency scores that are independent of these scale efficiencies. The CRS linear programming problem, Equation 4.5, was modified to account for VRS by adding the convexity constraint: N1'λ = 1 to provide:

$$\begin{aligned} &\min_{\theta,\lambda} \theta, \\ &st \quad -y_i + Y\lambda \geq 0, \\ &\qquad \theta x_i - X\lambda \geq 0, \\ &\qquad N1'\lambda = 1 \\ &\qquad \lambda \geq 0 \end{aligned} \qquad \text{............} 4.7$$

Where N1 is an N x 1 vector of ones. This forms a convex hull which envelopes the data more tightly than the CRS hull, thus the technical efficiency scores of VRS will be greater or equal to the technical efficiency scores from the CRS model. Technical efficiency scores from a CRS DEA can be due to scale inefficiency or “pure” technical inefficiency. Scale inefficiency is the ratio of the CRSTE to the VRSTE. Therefore, if there is a difference between the CRS and VRS efficiency scores it indicates that the DMU is scale inefficiency (Coelli, 1996).

4.1.2.3 Economies of Scale in DEA

This DEA model also calculates the economies of scale at which hospitals under study were operating at. The scale inefficiencies computed in the aforementioned case which is obtained as a ratio of CRSTE to VRSTE does not enlighten one as to whether the scale inefficiency is due to increasing or decreasing returns to scale. This is determined by running the DEA problem with non-increasing returns to scale imposed. This is done by substituting N1'λ = 1 in equation 4.7 by N1'$\lambda \leq$ 1 to come up with the following formulation:

$$
\begin{aligned}
&\min_{\theta,\lambda}\ \theta, \\
&st \quad -y_i + Y\lambda \geq 0, \\
&\qquad \theta x_i - X\lambda \geq 0, \\
&\qquad N1'\lambda \leq 1 \\
&\qquad \lambda \geq 0
\end{aligned}
\qquad \text{..........} 4.8
$$

The nature of the scale inefficiencies due to increasing or decreasing returns to scale for a particular DMU can be determined by considering whether the NIRS TE score is equal to the VRS TE score. If they are unequal then increasing returns to scale exist for that DMU. On the other hand if they are equal then decreasing returns to scale will be present.

4.2 DEA MODELS

There are various types of DEA models which may be used depending upon conditions at hand. For example, if it is assumed that economies to scale do not change as size of the service facility increases then a constant returns to scale type DEA model is an appropriate choice (versus the variable returns to scale DEA model). Furthermore, if managers` priorities are to adjust their inputs (before outputs) then an input-oriented DEA model rather than an output-oriented model is appropriate. Thus there are two basic empirical models of DEA; The constant returns to scale (CRS) model also known as the Charnes Cooper and Rhodes (CCR[9]) model and the variable returns to scale (VRS) also known as the Banker, Charnes and Cooper (BCC[10]) model.

4.2.1 The Charnes, Cooper and Rhodes Model (CCR, 1978)

Building on the work of Farrell (1957) Charnes, Cooper and Rhodes (1978) constructed a constant returns to scale frontier by identifying a DMU which maximizes the ratio of output to

[9] After its pioneers Charnes , Cooper and Rhodes
[10] After its pioneers Banker , Charnes and Cooper

input. This ratio can be interpreted as the maximum average productivity and denotes the scale efficient branch since it is consistent with a position of constant returns to scale. The constant returns to scale assumption are only appropriate when all DMUs are operating at an optimal scale.

Productive efficiency is achieved when hospitals operate on the frontier of the production technology at a scale of operation characterized by constant returns to scale. Operating under IRS means that an inefficient small output vector is being produced, operating under DRS means that an inefficient large output vector is being produced. Thus the inefficiency associated with VRS (non-constant returns to scale) is captured by a measure of scale efficiency.

In the event that the DMU`s to be considered are not operating at an optimal scale, TE scores obtained from a CRS DEA can be decomposed into two components, one due to scale inefficiency and one due to "pure technical inefficiency." This is done by conducting both a CRS and VRS DEA upon the same data. If there is a difference in the two TE scores for a particular DMU, then this indicates that the DMU has scale inefficiency. Thus the scale inefficiency can be calculated from the difference between the VRS TE and the CRS TE score. Under the assumption of constant returns to scale, the efficiency of hospital j can be obtained by solving the following CCR model:

$$Max h_o = \sum_{r=1}^{s} U_r Y_{rj}$$

Subject to $$\sum_{r=1}^{s} U_r Y_{rj} - \sum_{i=1}^{m} V_i X_{ij} \leq 0 , \qquad \text{(4.9)}$$

$$\sum_{i=1}^{m} V_i X_{io} = 1$$

$$U_r, V_i \geq 0$$

where:

Y_{rj} (r = 1... s) - actual amount of output r produced by hospital j.

X_{ij} (I = 1…m) - actual amount of input i used by hospital j.

U_r - weight given to output r.

V_i - weight given to input i.

The formulation in (4.9) implies that the objective is to maximize output $Y = \sum U_r Y_{ro}$ (summed over r = 1... s) subject to unit input $X_O = \sum U_i X_{io}$ (summed over i = 1…m) while maintaining the conclusion that virtual output cannot exceed virtual input for any DMU. This implies that the conditions for Pareto optimality are fulfilled since further increases in this maximal value can be attained only if some of the input values X_{ij} are increased or if some of the output values Y_{rj} are decreased.

The first constraint implies that all hospitals are on or below the frontier, that is, the efficiency of all hospitals has an upper bound of one. The second constraint indicates that the weighted sum of inputs for the particular hospital equals one. The weights U_r and V_i are treated as unknowns and their weights are obtained in the linear program solution. The weights are the variables of the problem. They determine which input a particular DMU is best in utilizing or which output it is best in generating. The DMU is assigned higher rates to those inputs and output variables which it is more adept or best in utilizing or in generating, and lower rates to others. If there are N DMU`s, this linear programming model has to be run N times, once for each DMU.

This model is designed to evaluate the relative performance of some decision making unit (DMU) designated as DMU_O based on observed performance of j = 1, 2… DMU`s. The Y_{rj}, $X_{ij} > 0$ in the model are constraints which represents observed amounts of the r^{th} output and the ith input of the j^{th} decision making unit (DMU) denoted as DMUj in a collection of j = 1,..,n entities which utilizes these i =1..., m inputs and produce these r =1.., s outputs. One of the j = i..., n DMU`s is singled out for evaluation denoted DMUo. The value h_0^* obtained from this ratio satisfies $0 \leq h_0^* \leq 1$ and can be interpreted as an efficiency rating in which $h_0^* = 1$ represents full efficiency and $h_0^* < 1$ means inefficiency is present. The asterisk indicates an optimal value obtained from solving the model. The value of h_0^* has operational significance in that 1-h_0^* provides an estimate of the inefficiency for each DMUo being evaluated. It should be emphasized that the orientation of DEA is generally toward relative efficiency as determined by the above optimization. Thus for any DMU being evaluated the optimization implies that the evaluation will be affected by reference to the subset of j = 1…n DMUs. Since the formulation in (4.9) is a primal linear programming problem, it has a dual formulation which can be constructed by minimizing the quantities of the m inputs required to meet stated levels of the s outputs. That is;

Minimize: $$\theta - \varepsilon\left[\sum_{i=1}^{m} S_i^{-} + \sum_{r=1}^{s} S_r^{+}\right]$$

Subject to: $$\theta X_{io} - S_i^{-} = \sum_{j=1}^{n} X_{ij}\lambda_j \qquad \text{.......}(4.10)$$

$$Y_{ro} + S_r^{+} = \sum_{j=1}^{n} Y_{rj}\lambda_j$$

$$\lambda_j \geq 0,\ j = 1 \ldots n \qquad \text{(Weights on branches)}$$

$$S_i^{-} \geq 0, i = 1 \ldots m \qquad \text{(Input slacks)}$$

$$S_r^{+} \geq 0, r = 1 \ldots s \qquad \text{(Output slacks)}$$

Branch O is relatively efficient if and only if the efficiency ratio θ^* equals unity and the slack variables are all zero. That is, if and only if θ^*=1 with S_i^-= S_r^+ = 0, for all *i* & *r*. where the asterisk denotes optimal values of the variables in the dual program. It is necessary to have both θ^*=1 and zero slacks in order for DMU_o to be characterized as fully efficient (100%). θ^* is the technical efficiency score to be estimated for each DMU. Thus the solution to problem (4.10) is θ^*, the TE score and λ, the weights which are the variables of the problem.

4.2.2 The Banker, Charnes and Cooper Model (BCC, 1984)

Another version of DEA is the Banker, Charnes and Cooper (BCC, 1984) Model. The primary difference between the BCC model and the CCR model is the treatment of returns to scale. The CCR version bases the evaluation on constant returns to scale. The BCC version is more flexible and allows variable returns to scale. The CCR model assumes that the DMU`s are operating in the optimal scale. However this is not always the case. In order to estimate the VRS DEA model we run the following (primal) model:

$$Maxho = \sum_{r=1}^{s} U_r Y_{rjo} + U_O$$

$$St \quad \sum_{r=1}^{s} U_r Y_{rj} - \sum_{i=1}^{m} V_i X_{ij} + U_o \leq 0, \; j = 1...n \qquad \text{.......(4.11)}$$

$$\sum_{i=1}^{m} V_i X_{ijo} = 1$$

$$U_r, V_i \geq 0$$

$$U_o \frac{\geq}{\leq} 0$$

The additional term in the primal formulation (4.11) is unconstrained in sign. The sign of U_0 determines the returns to scale, where $U_O < 0$ indicates increasing returns to scale, $U_o = 0$ constant returns to scale and $U_O > 0$ for decreasing returns to scale. The CCR model in its dual formulation is then:

Minimize: $\theta - \varepsilon\left[\sum_{i=1}^{m} S_i^{-} + \sum_{r=1}^{s} S_r^{+}\right]$

Subject to: $\theta X_{io} - S_i^{-} = \sum_{j=1}^{n} X_{ij}\lambda_j$ ....(4.12)

$Y_{ro} + S_r^{+} = \sum_{j=1}^{n} Y_{rj}\lambda_j$

$1 = \sum \lambda_j$

$\lambda_j \geq 0,\ j = 1 \ldots n$

$S_i^{-} \geq 0, i = 1 \ldots m$

$S_r^{+} \geq 0, r = 1 \ldots s$

The difference between the CCR model and the BCC model is that the λ_js are now restricted to summing to one. This has the effect of removing the constraint in the CCR model that DMUs must be scale efficient. The DEAP 2.1 Computer program estimates both the dual and the primal problems for each model.

Running the above model (4.12) for each DMU, the BCC efficiency scores are obtained. These scores are also called “pure technical efficiency scores” since they are obtained from the model that allows variable returns to scale and hence eliminate the “scale part” of the efficiency from the analysis. That is, for the DMU to be considered as CCR efficient it must be both scale and technical efficient but for it to be BCC efficient it only needs to be technically efficient. The VRS model provides technical efficiency scores which are greater than or equal to those obtained using the CRS model.

4.3 PRECONDITIONS FOR APPLICATION OF DATA ENVELOPMENT ANALYSIS

The application of the DEA approach requires a set of axioms or conditions for the effective, interpretation, use and acceptance of the results of the DEA. The following guidelines and conditions are useful when applying the DEA.

i.) Positivity Property

Generally, the DEA formulation requires that the input and output variables be positive (greater than zero). If the variable is not positive the following procedure is performed.

A positive amount is added to the negative value so that the particular input or output variable becomes positive. This same adjustment must be made to the same input or output value for all decision making units included in the data set in order not to alter the efficiency frontier.

ii.) Isotonicity Property

It is required that the functions relating inputs to outputs have the mathematical property called isotonicity. This means that an increase in any input should result in some output increase and not a decrease in any output. If it is apparent that the isotonicity property is violated, the isotonicity requirement may be accommodated by using reciprocals.

iii.) Homogeneity of DMUs

DEA requires a relatively homogeneous set of entities. That is, all entities included in the evaluation set should have the same inputs and outputs in positive amounts in order to control for case-mix.

iv.) Numbers of DMUs

A sufficiently large number of DMUs are needed in order to ensure sufficient degrees of freedom for a meaningful analysis. If a small sample is included in the dataset, there is a danger that an excessive number of the DMUs will be considered efficient (receive a rating of one) because of an inadequate number of degrees of freedom. The discriminating power of DEA is limited when a small number of DMUs are considered. The sample size used in this study exceeds the rule of thumb given by Ganley and Cubbin (1992) who states that it should be larger than the product of the number of inputs and outputs

4.4 MODEL SPECIFICATION

The empirical model to be adopted in this study is that of Banker Charnes and Cooper (BCC Model) also known as the VRS DEA Model. The definitions of variables are contained in Table 4.1. Each variable has been chosen to reflect important characteristics of hospital care provision as indicated in Zimbabwean hospitals and in empirical literature.

The concept of efficiency used in this study is that of technical efficiency and the model for empirical measurement is the multi-stage, input-oriented and variable returns to scale (VRS) DEA. The assumption of the CRS DEA model is suitable for situations where all hospitals are operating at an optimal scale. This assumption is not realistic for most public institutions such as hospitals and therefore could not be adopted in this study. The input-orientation approach was used because in most cases hospitals do not control their output level but merely respond to the demand for medical care in their catchment areas. Hospital managers have limited control over the volume of their outputs. In many circumstances of health care services outputs are exogenous. For instance, output in terms of outpatient visits is not chosen by hospital managers. Rather it is intuitively determined by the health seeking behaviour of the public. It may therefore be meaningless to suggest that output be raised to increase efficiency.

4.5 DATA AND JUSTIFICATION OF VARIABLES

4.5.1 Data Sources

This study uses data on hospitals from all the provinces in the country, covering the period 2006 - 2008, the years for which relatively reliable data are available. The data was obtained from the Central Statistics (CSO), Health Department and the Zimbabwe National Health Profile (2007) published by the Ministry of Health and Child Welfare (MOH&CW) as well as from Zimbabwe Association of Church-related Hospitals (ZACH). However, it is imperative to note that most of the data on for-profit hospitals is not published due to their reluctance to release information governing their activities which might place them at a disadvantaged position with their competitors. As a result this information was therefore collected from different individual hospitals. This is predominantly the reason why private for profit hospitals account for a small sample in the study, an estimation of 30% of the relevant for-profit hospitals. Nevertheless this constraint in data collection does not compromise the quality of this study since the sample size is above that of the rule of thumb. The availability and completeness of the data determined the selection of the time period covered. The sample consists of 36 mission hospitals and 44 public hospitals made up of 13 central hospitals and 31 district hospitals and 20 private for profit hospitals making a total sample of 100 hospitals. With the exception of for-profit hospitals, the sample size in each case represents at 40% of the applicable hospital population.

4.5.2 Input and Output Variables

Hospitals do not normally fit an economist's standard notion of a firm and this gives rise to a host of challenges in properly defining its unique variables. Ellis (1992) stated that measuring the improvement in health status is unrealistic as health status is a multifaceted concept that is not easy to define and measure in an operationally feasible manner. Hospitals are multi-product firms treating a variety of patients with a variety of inputs. There is no consensus as to how outputs of hospital production should be accurately measured since the conceptual output, relative change in health, is unobservable. Thus the definition of output has always been problematic and has been a matter of long standing debate among researchers in empirical studies of the health care industry.

The input and output variables used in this dissertation comprise of variables support the assumption of the DEA method and the analysis of efficiency described in the literature. This dissertation used inpatient days and total discharges to constitute the major output variables for the hospitals. Among the input variables, the number of existing hospital beds has been used as a proxy for capital while the number of doctors and number of nurses has been used to reflect manpower inputs (labour). These inputs and outputs are measured in their physical units. The application of the DEA methodology does not need a homogeneous unit of measurement, the reason being that DEA is unit invariant that is, "changing the unit of measurement will not change the value of the efficiency measure" (T. Coelli, 1996). The inputs and outputs variables used in this study and their description have been provided in Table 4.1**.**

However, the above definition of variables if not adjusted for price variations, quality and case-mix presents three important methodological problems.

- Meaningful comparisons of efficiency among providers must be made while controlling for quality of care. Otherwise lower costs attributed to inferior quality could be erroneously ascribed to higher efficiency and vice-versa.
- Variations in case-mix (i.e. the types and complexity of the medical cases treated) must also be controlled for when comparing provider efficiency. Failure to do so may yield the

misleading result that those providers treating more complex cases appear less efficient than those treating simpler cases.

- Price or cost differences can make comparisons of efficiency among providers difficult particularly between private and public providers.

However the problem of price variations is not an issue in this study since data on input prices is not relevant when dealing with technical efficiency (as opposed to allocative efficiency). The problem of quality and case-mix in this dissertation has been handled by confining our study to only those health providers that are relatively homogeneous, which treat a relatively similar set of medical problems and uses, to some extent, similar medical technology.

Table 4.1: Definition and description of the variables

Variables	Description
Outputs	
Inpatient days [11]	The number of days of care charged to a beneficiary (patient) who is admitted to a hospital bed for inpatient care services or skilled nursing facility care services (always in units of full days).
Discharged patients	The number of patients who leaves the hospital and either returns home or are transferred to another facility.
Inputs	
Beds	The number of existing patient beds within the hospital ready for use.
Doctors	The number of medical doctors employed in the hospital (both specialists and general medical practitioners).
Nurses	The number of nurses employed in the hospital (plus nursing aids)

[11] It can also be defined as the number of days during which services are provided to an inpatient e.g. one patient occupying one room for two nights would be counted as two inpatient days.

4.6 ECONOMETRIC ANALYSIS OF THE DETERMINANTS OF INEFFICIENCY

Studies have shown that institutional factors at the discretion of the management as well as environmental factors beyond the control of the DMU affect a DMU`s efficiency (Ferrier and Valdmaris, 1996 and Rosko et al, 1995.) After computing efficiency scores for each hospital the question of the relationship between hospital ownership and technical efficiency may be addressed. An Ordinary Least Squares (OLS) regression analysis is conducted using efficiency scores for each hospital as the dependent variable. Therefore a DEA efficiency index is regressed on several explanatory variables.

Hospital Ownership

An OLS regression analysis is estimated using the inefficiency score for each hospital as the dependant variable. Several explanatory variables are chosen. First and of greatest importance to the present study is to investigate the significance of profit incentives on efficiency generation. The research question is "Does the profit motive influence the way hospital managers utilize their resources and therefore influence efficiency behaviour?" There is a view that economic inefficiency in the hospital sector is more prevalent among non-profit hospitals and that it may be improved by relying more heavily on profit incentives (Register et al, 1987). This argument, drawn from the theory of property rights claim that the non-profit hospital is inefficient by nature and that transferring the ownership (property rights) to the private for profit sector is the only way to improve efficiency. The expectation in this study is therefore that for-profit ownership will be associated with greater efficiency than not-for-profit ownership. A dummy variable designed to capture status of the hospital is therefore included. The inclusion of this variable forms the primary motivator of this study which is once again to evaluate the effect of profit incentives on efficiency.

Average Length of Stay

The average length of stay variable is also included as an explanatory variable. The assumption is that patients with longer lengths of stay require more resources because they represent chronic cases that do not improve. As a result hospitals with patients with longer length of stay may exhibit lower efficiency scores. Accordingly a variable representing the average length of stay in the hospitals is incorporated.

Bed Size

Next in order to control for the impact of hospital size on the degree of hospital efficiency, a hospital size variable is entered into the model. This variable seeks to capture the influence of hospital size on efficiency. The number of beds will be used as a proxy for hospital size. There is need to ascertain the variation in efficiency as the hospital size varies. Ideally, a hospital should operate at an optimal scale which is at the lowest of its long run average cost curve. However, in reality most hospitals exhibit scale inefficiencies in the form of increasing or decreasing returns to scale. The inclusion of this variable would inform the researcher whether the hospital in question is small or large relative to the optimal size. Moreover, if there is a significant relationship between hospital size and the mix of cases treated, it becomes important to restrict the sample to hospitals which are similar in size. This is because if small hospitals tend to treat a less severe mix of cases than larger hospitals and if for-profits tend to be smaller than nonprofits. Then behavioral differences attributed to differing ownership arrangements may be due to size and independent of ownership. In such a case it is important to control for the effect of size on efficiency.

Bed Occupancy Rate

Finally another determinant of efficiency is the occupancy rate. The occupancy rate will be measured by the number of patients in the hospital on a certain day by the actual number of beds.

Since this measure represents short-term fluctuations (that are largely beyond the control of the manager) rather than long-term trends, any relationship found between the occupancy rate and the efficiency score is probably not a major area of concern.

Thus, if one form of hospital ownership is superior in terms of technical efficiency, then the coefficients associated with the ownership dummy will reflect the significance of this difference. Due to data collection constraints some important variables relating to hospitals and their operating environment has been omitted. The empirical model (REGRESSION) therefore takes the following form;

$$EFF_i = \beta + \beta_1 BOR_i + \beta_2 ALOS_i + \beta_3 BED_i + \beta_4 BEDSQ_i + \beta_5 DUM_{PROF_i} + \beta_6 DUM_{GOVT\,i} + \beta_7 \rho_i + \beta_8 \eta_i + \varepsilon_i$$

where**:**

EFF_i	-	Efficiency Score
BOR_i	-	Bed Occupancy Rate (%)
$ALOS_i$	-	Average length of stay (days)
BED_i	-	Size variable reflecting number of beds
$BEDSQ_i$	-	Bed Squared
DUM_{PROFi}	-	Dummy for for-profit ownership
DUM_{GOVTi}	-	Dummy for government ownership
ρ_i	-	Interaction term associated with beds and for-profit dummy
η_i	-	Interaction term associated with beds and government dummy
ε_i	-	Error term to capture other possible factors not specified.

4.7 CONCLUSION

This chapter has looked at the methodology as well as the description of variables to be applied in this study. The study will apply the input oriented VRS DEA model to estimate the technical efficiency scores of all the hospitals in the sample. The efficiency scores derived are then regressed on a number of environmental and organizational factors to identify those factors influencing inefficiency. Thus the simple OLS regression method will be adopted in this study to investigate the significance of the impact of hospital size, profit incentives, patient average length of stay and bed occupancy rate on efficiency.

The succeeding chapter will run the DEA models discussed in this chapter using the DEAP version 2.1 computer software developed by T. Coelli (1996). The software to be used for regression analysis is the Intercooled Stata version 8.0.

CHAPTER FIVE

INTERPRETATION OF THE RESULTS

5.1 INTRODUCTION

The aim of this chapter is to present empirical results obtained from the DEA model over a sample of 100 hospitals classified as private for-profit, private not-for profit (mission) and public or state owned hospitals. Mission hospitals were considered because of their crucial role and given that they also receive some funding from the government. However, the main focus of this study centers specifically on for-profit and public hospitals. The presentation of DEA results is guided by our methodological framework that was developed in the previous chapter. The computation of efficiency scores was undertaken using *DEAP version 2.1* software package developed by T. Coelli (1996). The rest of the chapter is organized as follows; the second section presents the DEA results, the third section is a discussion of the results and the last section will then conclude this chapter.

5.2 RESULTS FROM THE DEA MODEL

A summary of the efficiency scores for each type of hospital is presented in Table 5.1. The DEA results revealed that there was a marked deviation of efficiency scores from the best practice frontier with for-profit hospitals having the highest mean pure technical efficiency (PTE) score of 71.1 %. The mean PTE scores for both mission and public hospitals were 64.8 % and 62.6 %, respectively. The mean PTE score of for–profit hospitals of 71.1 % implies that the average inputs for the private hospital could be potentially reduced by 28.9 % without affecting the level of the outputs while the potential saving for mission and public hospitals is 35.2 % and 37.4 %, respectively. Technical and scale efficiency scores for individual hospitals can be found in the Appendix 2. It is imperative to recall that efficiency scores ranges from 0 (totally inefficient) to

1 (100% total efficiency). The presence of inefficiencies indicates that a particular inefficient hospital has excess inputs or insufficient outputs compared to those hospitals on the efficient frontier.

Table 5.1: Distribution of Pure Technical Efficiency (PTE) scores by Hospital Type:

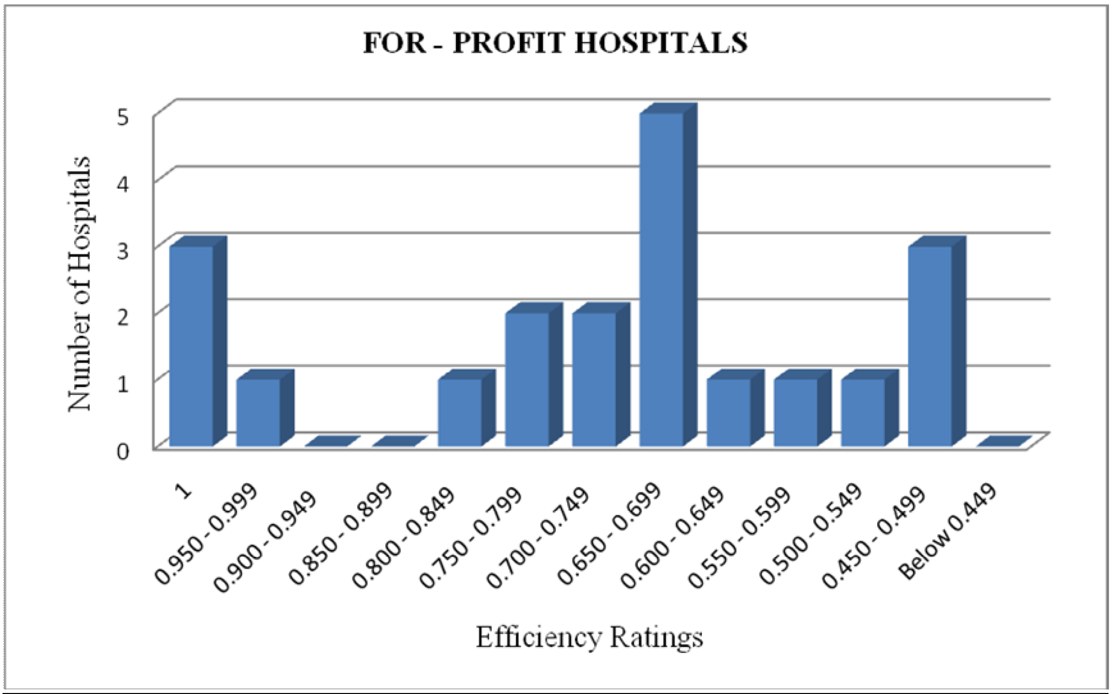
	NUMBER OF HOSPITALS			
EFFICIENCY RANGE	**FOR-PROFIT**	**MISSION**	**PUBLIC**	**ALL HOSPITALS**
1	3	6	4	13
0.950 - 0.999	1	1	3	5
0.900 - 0.949	0	0	0	0
0.850 - 0.899	0	2	0	2
0.800 - 0.849	1	1	3	5
0.750 - 0.799	2	2	2	6
0.700 - 0.749	2	2	4	8
0.650 - 0.699	5	1	3	9
0.600 - 0.649	1	3	6	10
0.550 - 0.599	1	4	3	8
0.500 - 0.549	1	5	1	7
0.450 - 0.499	3	2	3	8
Below 0.449	0	7	12	19
TOTAL	**20**	**36**	**44**	**100**
MEANS	**0.711**	**0.648**	**0.626**	**0.645**

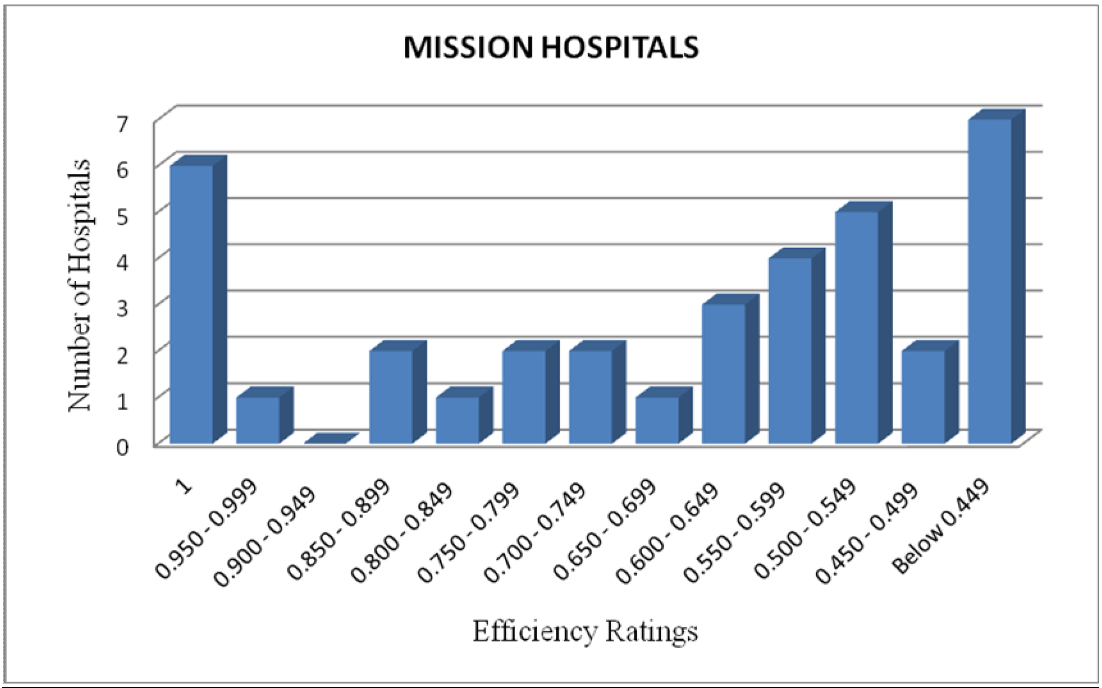
The proportion in each category of hospitals that were found to be efficient was 15%, 16.7 % and 9 % for For-profit, Mission and Public hospitals respectively. The DEA results revealed that Mission hospitals proved to exhibit the highest proportion of hospitals operating on the efficient frontier, followed by for-profit and lastly public hospitals. Overall, 13 % of all the hospitals included in the analysis were classified by the DEA program as being on the efficient frontier[12]. DEA results obtained for all the hospitals regardless of type, treated as one peer group showed a

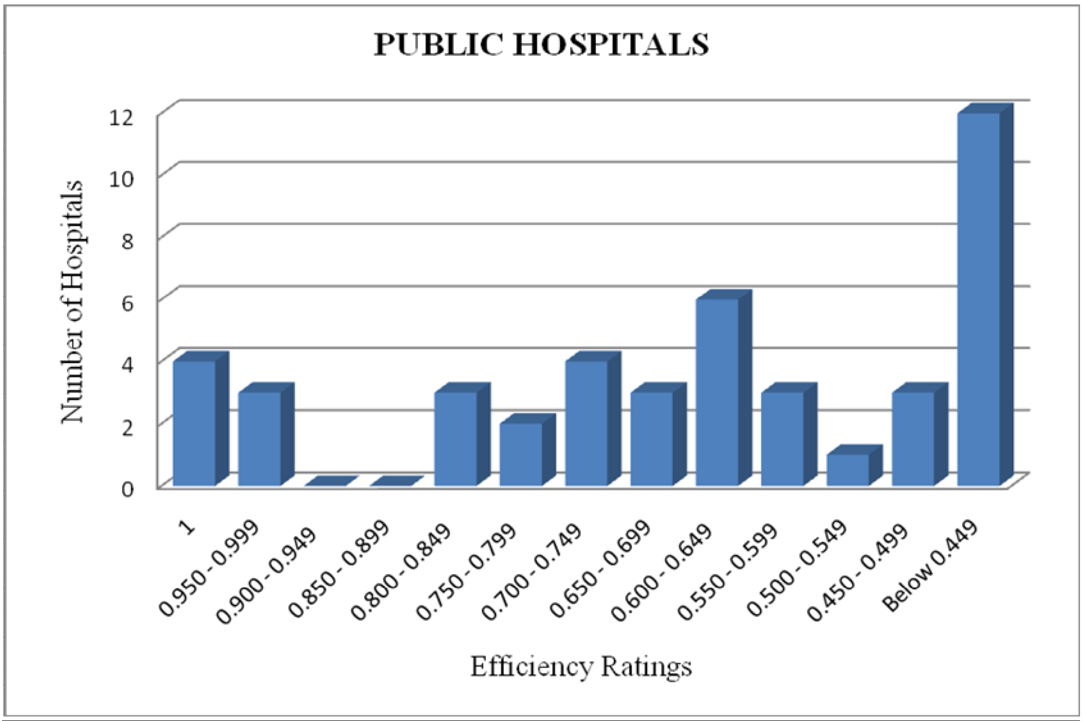
[12] The proportion of hospitals that received efficiency scores of 1 or 100%.

mean efficient score of 64.5 %. This efficiency score of 64.5 % implies that on average, these hospitals have the potential to reduce their utilization of all inputs by about 35, 5 % without reducing output. Figure 5.1 provides a visual depiction of the distribution of efficiency scores.

Figure 5.1: Distribution of efficiency scores by hospital type:




MISSION HOSPITALS
Number of Hospitals
0
1
2
3
4
5
6
7
1
0.950 - 0.999
0.900 - 0.949
0.850 - 0.899
0.800 - 0.849
0.750 - 0.799
0.700 - 0.749
0.650 - 0.699
0.600 - 0.649
0.550 - 0.599
0.500 - 0.549
0.450 - 0.499
Below 0.449
Efficiency Ratings


PUBLIC HOSPITALS
Number of Hospitals
0
2
4
6
8
10
12
1
0.950 - 0.999
0.900 - 0.949
0.850 - 0.899
0.800 - 0.849
0.750 - 0.799
0.700 - 0.749
0.650 - 0.699
0.600 - 0.649
0.550 - 0.599
0.500 - 0.549
0.450 - 0.499
Below 0.449
Efficiency Ratings

5.2.1 For-Profit Hospital Summary Results

The CRS technical efficiency scores reveal combined inefficiencies which are equivalent to overall technical efficiency (OTE) scores; inefficiency due to pure technical inefficiency and inefficiency that is due to inappropriate hospital size (scale inefficiency). Therefore further decomposition of overall technical efficiency (OTE) into pure technical efficiency (PTE) and scale efficiency (SE) enables an investigation of the sources of the inefficiencies of each hospital in the sample.

Table 5.2: Summary of DEA findings for For-Profit Hospitals:

FOR-PROFIT	**Overall technical Score(CRS)**	**Pure technical Efficiency (VRS)**	**Scale technical Score (CRS / VRS)**
Mean	0.614	0.711	0.856
SD	0.183	0.170	0.091
Max	1.000	1.000	1.000
Min	0.349	0.452	0.676
Hospital on frontier	**2**	**3**	**2**

An analysis of pure technical efficiency showed that out of the 20 for-profit hospitals included in the analysis whose results are tabulated in Table 5.4, 3 (15%) were technically efficient that is, there were on the frontier, whilst the remaining 17 (85%) were technically inefficient (see Table 5.4). Among the inefficient hospitals (see Table 5.1), 5 (25%) of the hospitals had a TE score below 60%, 10 (50%) between 60 and 79%, 1 (5%) between 80 and 89 %, and 1 (5%) between 90 and 99%. The inefficient hospitals had an average PTE score of 71.1% (see Table 5.1). This implies that on average, for-profit hospitals could reduce their utilization of all inputs by about 28.9 % without reducing output.

On the other hand, out of these 20 for-profit hospitals analyzed, none of the hospitals had a scale efficiency score below 50 %, 1 (5%) between 50 and 70, 13 (65%) between 70% and 90%, 4 (20%) between 90% and 100%, and 2 (10%) scored 100% (see Table 5.5). On the whole, 18 (90%) of the hospitals were scale inefficient; meaning their scale efficiency score was less than unity. The average scale efficiency score was 85.6% implying that under the assumption of variable returns to scale, for-profit hospitals could save on average 14.4 % of their inputs if they operate on the optimal scale. Table 5.5 shows the distribution of the scale efficiency scores.

DEA also demonstrates that 60 % of for-profit hospitals have efficient scores below the mean PTE score of 71, 1 % and are therefore run inefficiently. These hospitals need to either reduce their inputs or increase their outputs in order to become efficient. However since this study has adopted the input orientation owing to the fact that the hospital managers can only influence input utilization rather than output which is largely the choice of consumers, any inefficiencies realized would call for input reduction as opposed to output expansion.

Table 5.3: Distribution of overall technical efficiency scores for For-Profit hospitals

Overall technical efficiency scores (CRS)	**Percentage of hospitals**
0 – 50	30 %
50 – 70	50 %
70 – 90	10 %
90 – 100	0 %
100	10 %

Table 5.4: Distribution of Pure technical efficiency scores for For-Profit hospitals

Pure technical efficiency scores (VRS)	Percentage of hospitals
0 – 50	15 %
50 – 70	40 %
70 – 90	25 %
90 – 100	5 %
100	15 %

Table 5.5: Distribution of Scale efficiency scores for For-Profit hospitals

Scale efficiency scores (CRS/VRS)	Percentage of hospitals
0 – 50	0 %
50 – 70	5 %
70 – 90	65 %
90 – 100	20 %
100	10 %

5.2.2. Mission Hospitals Summary Results

A PTE score of 64.8% imply that on average mission hospitals are wasting 35.2% of inputs due to non-scale factors. Table 5.6 tabulates OTE, PTE and SE relating to mission hospitals. The OTE further incorporates (in)efficiencies emanating from size of operation in relation to its optimal size. For example an OTE of 35% as shown in Table 5.6 indicates that on the whole, consolidating both scale and non-scale factors, an input reduction of 65% are necessary for mission hospitals to operate on the efficiency frontier. Lastly, a SE score of 54.9% means that mission hospitals could save on average 45.1% if they operate on the optimal size. It is important to note that mission hospitals had the least SE and OTE. Tables 5.7 to 5.9 show scores of these three aforementioned efficiencies and their distribution.

Table 5.6: Summary of efficiency scores for mission hospitals

MISSION	**Overall technical Score(CRS)**	**Pure technical Efficiency (VRS)**	**Scale technical (CRS / VRS)**
Mean	0.350	0.648	0.549
SD	0.237	0.227	0.278
Max	1.000	1.000	1.000
Min	0.015	0.230	0.015
Hospital on frontier	1	6	1

Table 5.7: Distribution of overall technical efficiency scores for mission hospitals

Overall technical efficiency scores (CRS)	**Percentage of hospitals**
0 – 50	78 %
50 – 70	11 %
70 – 90	8 %
90 – 100	0 %
100	3 %

Table 5.8: Distribution of Pure technical efficiency scores for mission hospitals

Pure technical efficiency scores (VRS)	**Percentage of hospitals**
0 – 50	25 %
50 – 70	36 %
70 – 90	19 %
90 – 100	3 %
100	17 %

Table 5.9: Distribution of Scale technical efficiency scores for mission hospitals

Scale efficiency scores (CRS/VRS)	Percentage of hospitals
0 – 50	44 %
50 – 70	19 %
70 – 90	22 %
90 – 100	11 %
100	3 %

5.2.3. Public Hospitals Summary Results

The Tables below relates to results pertaining to public hospitals. Computed DEA results revealed that public hospitals exhibited the least PTE scores and hence accounted for the majority of inefficiency hospitals. For instance a PTE score of 62.6% imply that public hospitals could reduce their input utilization by 37.4% of inputs without affecting output level (see Table 5.10). Similarly an OTE of 50.3% as shown in Table 5.10 indicates that overally an input reduction of 49.7% is required for public hospitals to operate on the efficiency frontier. Moreover a SE of 82.9% means that public hospitals have the potential of reducing input usage by 17.1% if they operate on the optimal size. Tables 5.11 to 5.13 tabulate various efficiencies and their distribution for public hospitals.

Table 5.10: Summary of efficiency scores for Public hospitals

PUBLIC	**Overall technical Score(CRS)**	**Pure technical Efficiency (VRS)**	**Scale technical (CRS / VRS)**
Mean	0.503	0.626	0.829
SD	0.200	0.230	0.116
Max	1.000	1.000	1.000
Min	0.171	0.198	0.484
Hospital on frontier	2	4	2

Table 5.11: Distribution of overall technical efficiency scores for public hospitals

Overall technical efficiency scores (CRS)	**Percentage of hospitals**
0 – 50	52 %
50 – 70	36 %
70 – 90	5 %
90 – 100	2 %
100	5 %

Table 5.12: Distribution of Pure technical efficiency scores for public hospitals

Pure technical efficiency scores (VRS)	**Percentage of hospitals**
0 – 50	32 %
50 – 70	30 %
70 – 90	20 %
90 – 100	9 %
100	9 %

Table 5.13: Distribution of Scale technical efficiency scores for public hospitals

Scale efficiency scores (CRS/VRS)	**Percentage of hospitals**
0 – 50	2 %
50 – 70	18 %
70 – 90	50 %
90 – 100	25 %
100	5 %

The mean overall technical efficiency[13] scores for the for-profit, mission and public hospitals were 61.4 %, 35 % and 50.3 % respectively. This implies that if the inefficient hospitals were to operate as efficiently as their peers on the efficient frontier, inputs can be decreased by about 38.6%, 65% and 49.7% respectively without changing the quantity of outputs attained. In other words, the inefficiency rating $(1 - efficiency\ \ score)$ indicates the maximum reduction that can be affected uniformly across all inputs without causing a reduction in the volume of that production unit or branch. The score for mission hospital category proved to be the worst excessive input wastage of 65% which could be reduced and still attain the same output level.

The CRS technical efficiency scores reveal combined inefficiencies; inefficiency due to pure technical inefficiency and inefficiency that is due to inappropriate hospital size (scale inefficiency). Therefore further decomposition of overall technical efficiency (OTE) into pure technical efficiency (PTE) and scale efficiency (SE) enables an investigation of the sources of the inefficiencies of each hospital in the sample. About 90%, 97% and 95% of for-profit, mission and public hospitals respectively were found to be scale inefficient. In other words they suffered from inefficiencies emanating from inappropriate hospital size namely either for being too small or too large relative to the optimum size. The average scale efficient scores for the sample were 85.6%, 54.9% and 82.9% for for-profit, mission and public hospitals respectively (see Table 5.2, 5.6 and 5.10 respectively). This implies that if all hospitals had an optimal size, input usage would have decreased by about 14.4%, 45.1% and 17.1% respectively without decreasing the output attainment.

The figures in the Table 5.2, 5.6 and 5.10 for for-profit, mission and public hospitals respectively also suggest that the dispersion or variability of the overall technical efficiency (OTE) measures is wider for mission hospitals compared to those of other efficiency measures. Its standard deviation (SD) of 23.7 is the highest for OTE which is marginally different from that of public hospitals of 23.1 and the least is 18.3 for for-profit hospitals. A comparison of these values indicated that efficiency scores for for-profit hospitals had the least variation as opposed to both mission and public hospitals.

[13] Mean overall technical efficiency score is tantamount the CRS efficiency scores. It is further decomposable into pure and scale efficiency.

5.3 RETURNS TO SCALE

Disintegration of the OTE into PTE and SE necessitates an enquiry into the sources of inefficiencies for hospitals that are not operating on the efficient frontier. A CRS model assumes that hospitals are operating on the optimal size whereas VRS model decomposes efficiency scores into pure and scale efficiency score. In order to investigate whether the inefficiencies were due to increasing returns to scale (IRTS) or decreasing returns to scale (DRTS) the VRS DEA model was run. The findings indicate that only 2 (10%), 1 (3%) and 2 (5%) in the sample of 20 for-profit, 36 mission and 44 public hospitals operated at CRS (at optimal size) (see Table 5.2, 5.6 and 5.10). The findings further revealed that 17 (85%), (35) 97% and 36 (82%) of for-profit, mission and public hospitals operated at IRS and 1 (5%), none (0%) and 6 (14%) of the for-profit, mission and public hospitals exhibited DRS respectively (see Figure 5.1 and 5.3).

Increasing returns to scale is the predominant form of scale inefficiency observed. About 10%, 3% and 5% of for-profit, mission and public hospitals respectively are operating at CRS. The remaining proportion of hospitals operates either under IRTS or DRTS or both as mentioned afore. These figures indicate that a great proportion of hospitals are of inefficient size, that is, they are bigger or smaller than the optimal size. There is a wide prevalence of IRS among hospitals implying that a proportionate increase in all inputs is followed by more than a proportionate increase in outputs. This means that hospitals should be able to expand both their inputs and outputs. But it is necessary to note that hospitals outputs are unique, that is in order to increase the scale of operation demand has to increases also. However it is known that demand management is beyond the control of hospital management but is solely the decision of consumers of hospital services, the patients. In order to reap gains from IRS most studies have recommended mergers of small hospitals

Evidence of DRS was also found among for-profit and public hospitals with only one (5%) for-profit hospital and six public hospitals accounting for 14% operating at decreasing returns to scale. A hospital operating at decreasing returns to scale has an inefficiently large size. This means that a percentage increase in all inputs is followed by less than a percentage change in outputs. In order to improve the efficiency of the inefficiently large hospitals, there is a need to have more health units of a relatively smaller size. It is important to note that public hospitals were largely found to exhibit IRS and to a lesser extent DRS as opposed to results from other studies. The main reason for this could be attributed to the current shortage of skilled and experienced health workers(in relation to the number of beds) at a time when demand for health services is increasing due to a growing population and the challenges posed by HIV/AIDS. If the existing manpower resources could be augmented by additional medical personnel public hospital output such as discharged patients and inpatient days could increase remarkably. Health professional groups with the highest loss rate in Zimbabwe are doctors, nurses and pharmacists.

On the other hand, the DEA results obtained only noted one for-profit hospital exhibiting DRS. Normally a predominance of DRS in the private sector especially for-profit hospital can be explained by the exodus of specialists' health manpower from the public to the private for-profit sector resulting in more than proportionate medical personnel. This huge influx of manpower from the public hospitals to the private hospitals is normally attributed to significant differentials in salary and conditions of services between the public and private for profit sector. The current situation in Zimbabwe is such that there are continuing problems of low staff morale, low productivity, and the financial impoverishment of health workers in the public sector. About two thirds of Zimbabwe's physicians and state registered nurses (SRN) and a third of the country's state certified nurses (SCN) work in the private sector where pay and conditions of service generally compare favorably to the public sector. Attrition of health workers in the public sector is due to factors that mainly relate to lack of incentives, poor salaries and conditions of services (Paulinus et al 2000).

An analysis of returns to scale for each hospital category is shown in Figure 5.2 and 5.3 below. Both the percentages and the number of hospitals exhibiting a particular form of returns to scale are depicted in these same Figures 5.2 and 5.3 respectively.

Figure 5.2: Returns to scale for each hospital type (% of hospitals)

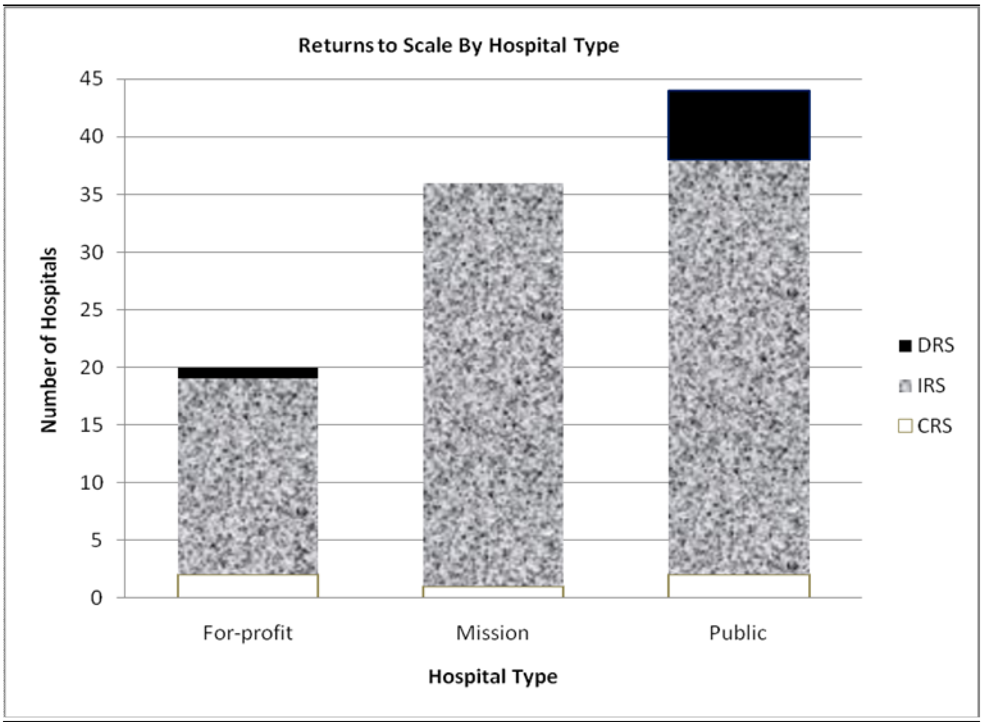
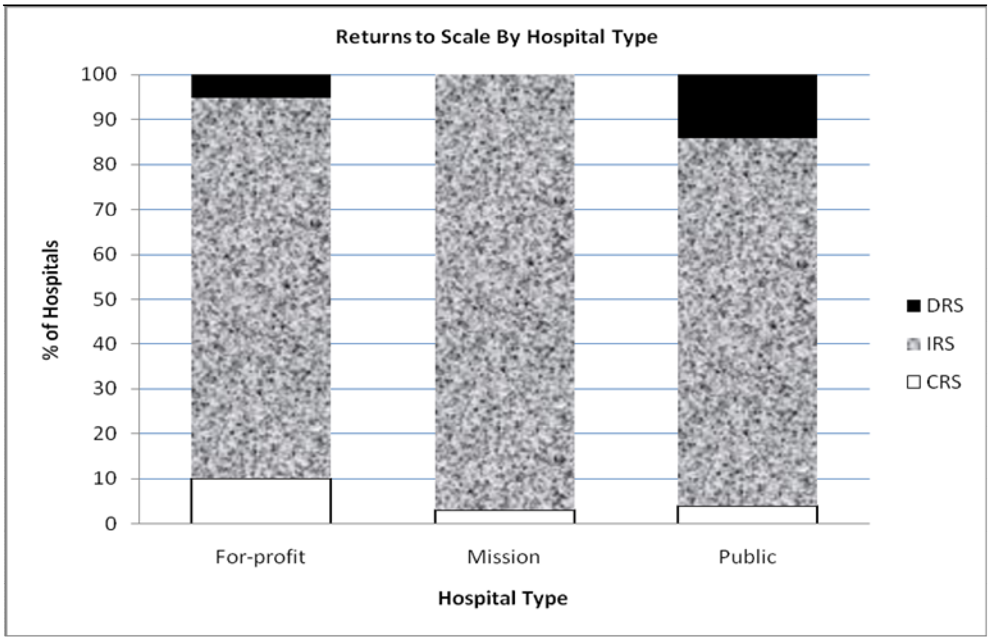


Figure 5.3: Returns to scale for each hospital type (Number of hospitals)



5.4 INPUT REDUCTION NECESSARY TO MAKE INEFFICIENT HOSPITALS EFFICIENT

In the current year, 2009, the government recognized the need to prioritize expenditures towards the health sector in order to revitalize the deteriorating crucial functions of the sector. It has therefore proposed to allocate an amount of US157.8 million to the Health Vote (National Budget Statement, 2009). Of this amount, US59.9 million has been earmark to target government central, provincial and district hospitals as well as rural centers. In this allocation, 60% catered for the procurement of drugs and medical supplies while 40% was for the general running expenses of these institutions. If the public hospital`s mean pure technical efficiency score of 62.6 % found in this study is to continue this year it would be reasonable to expect that about US$22.4026 million worth of financial resources (is heading for wastage) can been saved if the necessary input reductions “suggested” by the DEA program are considered. This substantial amount of input savings can go a long way in significantly contributing towards the urgent rehabilitation of central hospitals and the ongoing construction works of provincial and district hospitals as well as rural centers for which an amount of US$9 million was allocated in the 2009 National Budget Statement. Furthermore, these savings can also be channeled towards the procurement of sixty-one ambulances and eighty service vehicles for which an amount of US$4 million was allocated.

The findings of this study reveal that overall efficiency of all the hospitals in the sample ranges from 60 – 72 %. This implies that if the inefficient hospitals were to operate as efficient as their peers on the best-practice frontier, the health system could have reaped efficiency gains amounting to 28 – 40 % of the total resources used in running the hospitals. The possible input savings are depicted in Table 5.14 – 5.16. These input savings are aggregates for the each hospital type. The amounts of input savings for each hospital are given in the Appendix column. Table 5.14 – 5.16 shows the total input decreases needed to make inefficient hospitals efficient in order to become efficient. For example inefficient for-profit hospitals would need to decrease their current input levels by 31.53% less beds, 37.74% less doctors, and 37.44% less nurses.

These findings indicate that the amount of inputs could be decreased substantially without decreasing the quantity of outputs achieved. As can be seen from Table 5.14 – 5.16, each of the inputs exhibits a tremendous decrease – more than 30% in all cases. This includes both radial and slack movements. Radial movements indicate the proportional decrease in inputs without changing the mix of the inputs. On the other, hand slack movements arise because of the sections of the piece-wise linear frontier that run parallel to the axes. Both the radial and slack movements are reported in order to give an accurate indication of the technical efficiency of the hospitals. Slack scores are also calculated only for the inefficient hospitals. This is because efficient hospitals by definition do not have slack values.

Table 5.14: Input reductions needed to make inefficient hospitals efficient (FOR-PROFIT)

	Beds	Doctors	Nurses
Total Input Slack	35	3	27
Total Radial inefficiency	1105	17	122
Total Input	3616	53	398
TOTAL PERCENTAGE	**31.53%**	**37.74%**	**37.44%**

Table 5.15: Input reductions needed to make inefficient hospitals efficient (MISSION)

	Beds	Doctors	Nurses
Total Input Slack	15	11	181
Total Radial inefficiency	1984	44	433
Total Input	4683	109	1005
PERCENTAGE	**42.69%**	**50.46%**	**61.09%**

Table 5.16: Input reductions needed to make inefficient hospitals efficient (PUBLIC)

	Beds	Doctors	Nurses
Total Input Slack	312	26	1871
Total Radial inefficiency	5057	237	3694
Total Input	12134	711	9793
PERCENTAGE	**44.25%**	**36.99%**	**56.83%**

Input slack indicates additional amounts by which individual inputs could be reduced after eliminating radial inefficiency while holding output constant, whereas slack in the output constraints represents additional amounts by which particular outputs could be expanded after eliminating radial inefficiency while leaving inputs unchanged. The problem of input slack (slack movement) arises as a result of the sections of a piecewise linear frontier which run parallel to the axes. Coelli (1996) noted that if an infinite sample size were available and/or if an alternative frontier construction method was used which involved a smooth function surface, the slack issue would disappear. Slacks may be viewed as an artifact[14] of the frontier construction method chosen (DEA) and the use of finite sample sizes. It is against this background that many studies simply concentrate their technical efficiency measure on radial movements and ignore the slacks completely. Radial inefficiency (radial movement) refers to the proportionate reduction of the *i*-th hospital`s inputs holding output constant or the proportionate expansion of the *i*-th hospital`s output holding input constant. Among inputs each hospital type had a large amount of slack in staffing for nurses with the largest amount in public hospitals followed by missions with the least in for-profit hospitals. Burges et al (1996) noted that the variability in the uncertainty of demand for health care can show up as input slack since greater uncertainty would necessitate more standby capacity.

[14] Something observed in a scientific investigation that is not naturally present but occurs as a result of the investigative procedure.

5.5 DISCUSSION OF RESULTS

The findings of this study revealed that more than half of the hospitals are operating at less than optimal levels of pure technical and scale efficiency. The performance of some of the hospitals in the sample is actually observed to be very low, and raises much concern for planners and policymakers. With the existing levels of inefficiency, the achievement of the health policy objectives and health-related global and regional targets such as the Millennium Development Goals (MDGs) will be compromised. Hence, greater focus should be placed on efficient use of the existing resources.

Focusing on public hospitals, the Zimbabwean health sector is technically inefficient. However when compared with the results obtained from studies conducted in other countries, it is still relatively inefficient. On average the country has a pure technical efficiency score of 71.1%. This figure is lower than the score obtained by Mills and Liu (1998) which revealed a higher efficiency score of 77% and 67.4% for Korean public and private hospitals respectively. The efficiency score from our study was way less than for South Africa (Newbrander et al, 1993) and Mozambican hospitals (Mbalame, 2002) that had a score of 90.6% and 84.3% respectively. This comparison of Zimbabwean results with those from other countries is coupled with some inconsistencies since different inputs and outputs were used in these studies. DEA is sensitive to the variables included in the analysis. Mbalame (2002) used a 4 X 4[15] set of inputs and outputs, while a 3 X 2 set was used in our study. Furthermore the efficiency scores are time specific and they change over time making the results valid only at the point in time when the study was done.

The CRS TE scores as indicated in Tables 5.3, 5.7 and 5.11 indicate that throughout the period considered, less than half of the hospitals were located on the frontier (TE score = 100%). Furthermore, it is revealed that there are hospitals whose TE scores are extremely low. The results of this study indicate that many of the public hospitals operate at technical efficiency levels well below the efficient frontier. The findings of this study are in line with other studies in

[15] Four variables for inputs and four variables for outputs

sub-Saharan Africa, which indicate the wide prevalence of technical inefficiency among public hospitals. For example, Zere et al (2000) in their study of technical efficiency and productivity of public sector hospitals in South Africa found technical inefficiency levels ranging between 34% – 48%.

The study further reveals that the prevalent form of scale inefficiency is increasing returns to scale. In the presence of increasing returns to scale, expansion of outputs reduces unit costs. However, increasing the level of outputs requires an increase in the demand for health care, which is beyond the control of the hospital management. Merger of hospitals in close proximity to one another may be an option worth of consideration. However, this option may potentially pose some problems given the very low population density of the country. If larger hospitals are to be established in centrally located places, residents of some areas may incur additional costs in travel expenditure and in delayed treatment of emergency cases. These potential problems may to some extent be minimized by establishing primary care units linked to centrally located hospitals through an effective referral and patient transport system. In taking such decisions, however, the equity implications should always be viewed carefully. It is important that efficiency measures be integrated in policy formulation and pursued vigorously inorder to contribute towards improving the health status of the population.

5.6 REGRESSION ANALYSIS

The major force driving this study has been to investigate whether profit incentives in the provision of hospital care has a significant impact on efficiency. The findings of this study have so far revealed that private for-profit hospitals are more efficient than non-profit hospitals, both public and mission. A further comparison of overall technical efficiency scores between the public and the mission hospitals revealed that the mission hospitals are significantly more efficient than the public hospitals. These results are consistent with the property rights hypothesis. The property rights approach holds that the form of ownership is the predominant explanation for the varying performance of different organizations. The approach believes that the public sector is inefficient by nature and that transferring the ownership (property rights) to the private sector is the only way to improve efficiency. The aim of this section is to assess the significance of this claim as well as to extend this assertion by testing the hypothesis that profit incentives have a bearing on efficiency. The section that follows will probe into that enquiry using the Ordinary Least Squares (OLS) Regression.

The degree of correlation between independent variables is an important issue that has great impact on the robustness of the ordinary least regression (OLS) model. Thus a correlation analysis is imperative to establish appropriate regressors to be included in the model. It is critical that these regressors (independent variable) are not correlated among themselves inorder to dichotomize or separate their individual effects on the regressand (dependant variable). This is because the idea is to be able to establish with reasonable certainty the relationship between each of the hypothesized independent variables and the dependent variable. In this study, a correlation matrix was computed to investigate the interrelationships among the sets of variables to be included in the model and the table below presents the details (see Table 5.17).

Table 5.17: A matrix of correlation of independent variables

	BEDS	BOR	ALOS	DUM_{PROF}	DUM_{GOVT}	
BEDS	1.000					
BOR	**0.053**	1.000				
ALOS	**0.042**	**0.285**	1.000			
DUM_{PROF}	**-0.0765**	**0.261**	**0.380**	1.00		
DUM_{GOVT}	**0.4324**	**0.158**	**-0.516**	**-0.457**	1.00	

As can be seen from the Table 5.17 the variables exhibit very low correlations and therefore do not suffer from the problem of multi-collinearity which if not dealt with can cause spurious regression. The empirical model to be estimated therefore takes the following form; **Equation 1:**

$$EFF_i = \beta + \beta_1 BOR_i + \beta_2 ALOS_i + \beta_3 BED_i + \beta_4 BEDSQ_i + \beta_5 DUM_{PROF_i} + \beta_6 DUM_{GOVT\,i} + \beta_7 \rho_i + \beta_8 \eta_i + \varepsilon_i$$

where**:**

EFF_i	-	Efficiency Score
BOR_i	-	Bed Occupancy Rate (%)
$ALOS_i$	-	Average length of stay (days)
BED_i	-	Size variable reflecting number of beds
$BEDSQ_i$	-	Bed Squared
DUM_{PROFi}	-	Dummy for for-profit ownership
DUM_{GOVTi}	-	Dummy for government ownership
ρ_i	-	Interaction term associated with beds and for-profit dummy
η_i	-	Interaction term associated with beds and government dummy
ε_i	-	Variable to capture other possible factors not specified.

Table 5.18 shows descriptive statistics for the variables to be adopted into the model. Table 5.19 tabulates the estimated regression model that was obtained using the Intercooled Stata Version 8.0 Figure 5.4 provides a visual depiction of the distribution of efficiency scores for all hospitals in the sample. The figure shows that the distribution of efficiency scores among hospitals is

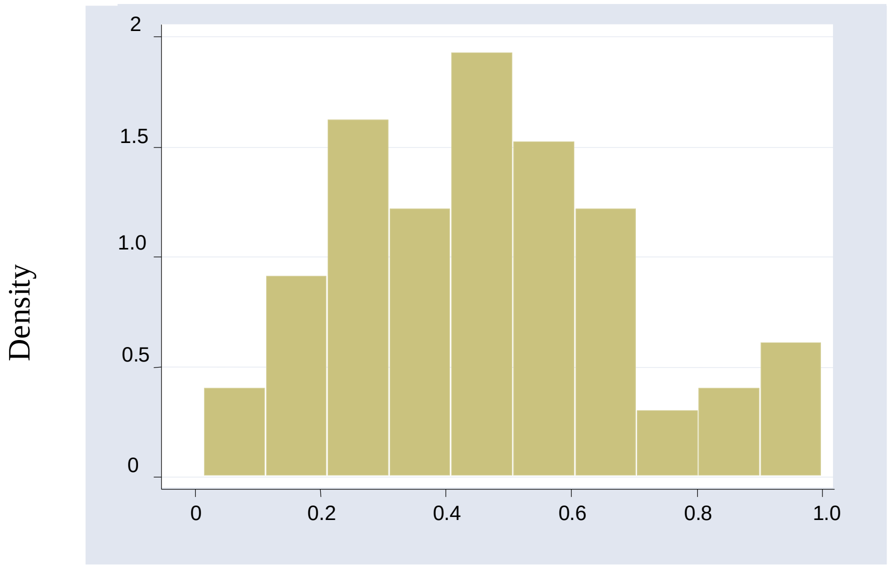
nearly normal. An analysis of each hospital type as shown by Figure 5.5, 5.6 and 5.7 indicate that each distribution approximates a normal distribution.

Table 5.18: Descriptive Statistics

Variable	Mean	Std. Dev	Min	Max
TE	0.47034	0.2313709	0.015	1
BED	204.33	186.196	50	1088
BEDSQ	76073.01	195281.8	2500	1183744
BOR	34.05277	17.3159	1.178	98.424
ALOS	5.05303	4.627642	0.126	34.638
DUM_{PROF}	0.21	0.4093602	0	1
DUM_{GOVT}	0.44	0.4988877	0	1
ρ	37.14	74.49575	0	238
η	121.39	218.6924	0	1088

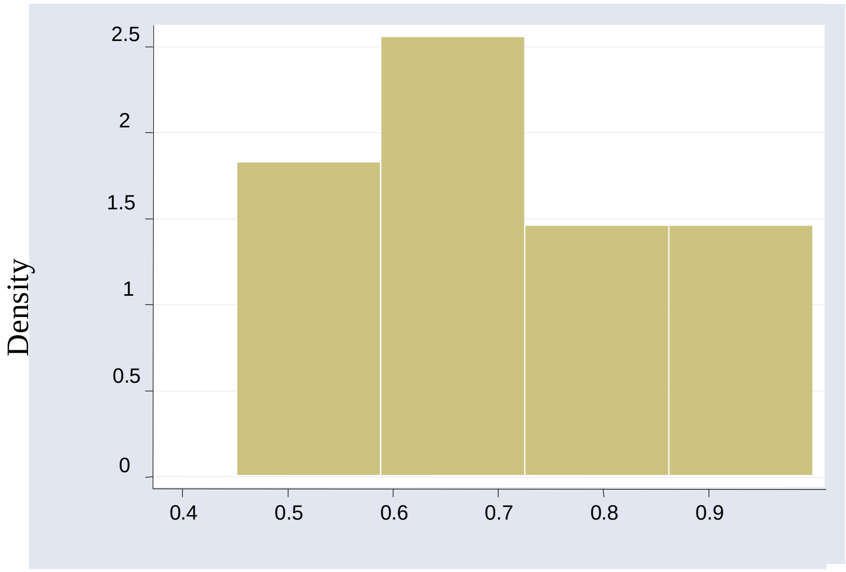
$n = 100$

Figure 5.4: ALL THE HOSPITALS



Efficiency Score

Figure 5.5: FOR-PROFIT HOSPITALS



Efficiency Score

Figure 5.6: MISSION HOSPITALS

Figure 5.7 PUBLIC HOSPITALS

After systematic trials of dropping insignificant variables and rerunning the model the following results were finally obtained as tabulated below in Table 5.19.

Table 5.19: Efficiency Regression Results, Dependent Variable: Efficiency Score

Independent Variable	**Coefficient**	**t-statistic**	**Significance**
Intercept/Constant	0.0641	3.00	S
Bed Occupancy rate (BOR)	0.0118	20.55	S
Average Length of Stay (ALOS)	-0.00266	-1.19	NS
$\mathbf{DUM_{PROF}}$	0.0861	3.42	S

S: statistically significant at 95% level of significance.

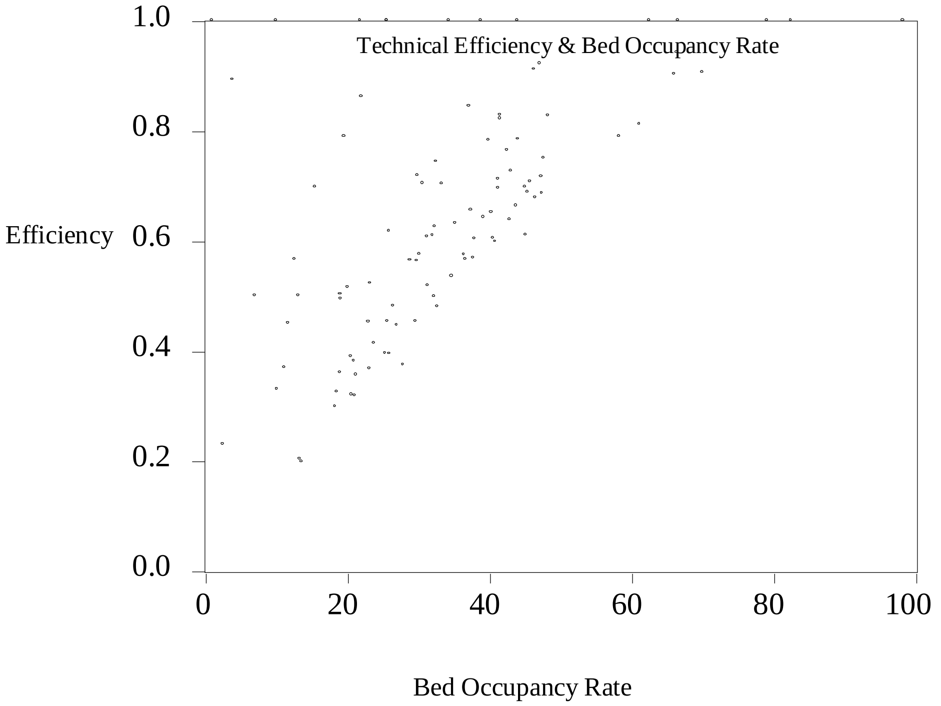
NS: not significantly at 95% level of significance.

R-squared: 0.8422

F (3, 96): 170.75

A final analysis of the results obtained revealed that DUM_{PROF} variable in Table 5.19 improved significantly following the dropping off of insignificant variables. Therefore at 5% level of significance the dummy variable is significantly different from zero. The regression results concluded that only DUM_{PROF} and BOR were found to be significant within the model. The sign of the significant variables were also consistent with the expected priori. Average length of stay is an important variable that determines hospital performance, while the sign of ALOS is consistent with theoretical and empirical literature; however the results indicate that its t-statistic value of -1.19 did not exceed the conventionally accepted value of 1.96 for 95% level of confidence and therefore was not significantly different from zero.

Figure 5.8: Scatter Plot Diagrams



5.7 CONCLUSION

In our study we have managed to examine efficiency figures of a sample of 100 hospitals divided into for-profit, mission and public. The main idea that motivated this study was to investigate the impact of profit incentives on technical efficiency in the provision of health care in the Zimbabwean health sector with the intention to shed some light on the debate about the validity of the theory of property rights. The evidence in this study suggests that for-profit hospitals have significantly higher efficiency scores than non-profit hospitals. A for-profit hospital uses on average about 8.5% fewer resources than public hospitals. The findings of this study also revealed that more than half of the hospitals are operating at less than optimal levels of pure technical and scale efficiency. The performance of certain individual hospitals in the sample is actually observed to be very low and raises much concern for planners and policy makers. The main focus of policy makers should be on public hospitals which had the lowest efficiency score lower than mission hospitals. With the existing levels of inefficiency, the achievement of the health policy objectives and health–related global targets such as the Millennium Development targets will be compromised. Hence, greater focus should be placed on efficient utilization of the existing resources before “clamoring” for additional resources.

CHAPTER SIX

CONCLUSION AND POLICY RECOMMENDATIONS

6.1 KEY FINDINGS.

The main motivation driving this study has been to address the subject of profit incentives and technical efficiency in the Zimbabwean health sector in order to cast some light on the debate about the validity of the theory of property rights. The importance of making such an enquiry arises from the fact that many studies in health care provision have suggested that economic efficiency in the hospital sector may be improved by relying more on profit incentives. Data Envelopment Analysis (DEA) was used to calculate efficiency scores for the various hospitals in the sample. The evidence in this study suggests that for-profit hospitals have significantly higher efficiency scores than non-profit hospitals. For-profit hospital with a PTE score of 71.1% uses on average about 8.5% fewer resources than public hospitals which had the least pure efficiency score of 62.6%.

The DEA calculated efficiency scores therefore shows that of all the hospitals included in the study, for-profit hospitals had the highest pure and overall efficiency score. The results suggest that there is a relationship between profit incentives and efficiency in the provision of hospital care. Therefore this study concludes that profit incentives contribute positively to efficiency. The results highlight the important role played by for-profit hospitals in the economy.

These efficiency measurements also help identify: I) efficient hospitals whose practice can be emulated by the inefficient hospitals; II) inefficient hospitals, whose performance need to be improved; III) the inputs that are being wasted and the magnitude of waste; and IV) the output increases needed to make inefficient hospitals efficient. This kind of evidence would then empower health policy makers and managers to develop concrete strategies for attaining maximum benefit with minimum input utilization thereby boosting the efficiency of hospitals.

6.2 POLICY IMPLICATIONS

The policy recommendations are made based on the empirical results from this study. For instance a result which indicates that for-profit hospitals are technically more efficient than non-profit hospitals could be used to justify policies designed to promote a shift towards greater reliance on profit incentives. On the contrary if for-profit hospitals are found to be no more efficient than non-profits, then economic support for the shift toward profit incentives would be discarded. It was determined in this study that significant differences in relative technical efficiency exist between for-profit and non-profit hospitals. From a policy standpoint, this conclusion suggests that initiatives designed to improve the hospital sector`s performances which are founded on the premise of profit incentives are likely to be of paramount value. In this respect the government should design policies to encourage the participation of non-governmental organizations (NGO`s), industries and individual private firms to establish hospitals or clinics to augment government efforts to provide health services of greater quality to the population.

Hospital based services can be made more efficient by eliminating duplication of services, reducing unnecessary surgery and hospitalization. Physicians and hospital administrators should be made more budget and cost conscious through changes in financial management practices. In regard to the excess resources, which are being wasted and not efficiently utilized in the production of hospital outputs, decision-makers have a number of policy options available to them:

In the light of scarce medical personnel in the country and given the need to strengthen health services at provincial, district and community levels, it would not be rational to offer excess medical staff the option of early retirement. Instead, excess medical officers and nurses should be transferred to health centers to provide primary health care or transfer excess labor force to understaffed hospitals. This would increase health coverage and quality of service provided at a time when demand for doctors and nurses is critical.

The study also identified evidence of scale inefficiencies arising from hospitals not operating on their most productive scale. In this respect hospital administrators or managers could consider scaling down their inputs and outputs if that particular hospital exhibits DRS. On the other hand if that particular hospital is exhibiting IRS then policy should focus on expanding both inputs and outputs. However the choice of which inputs or outputs to increase or decrease needs to be done carefully. It is imperative to appreciate the fact that hospital inputs are not within the capacity of hospital managers to manipulate rather it is only the input factors that are at their discretion to influence.

On the whole regarding the excess utilization of resources, reallocation of resources should be channeled towards health institutions where benefits are at the maximum. In other words policy focus should concentrate on reallocating the excess inputs to underserved or understaffed hospitals. However the policy is only feasible if the cost of maintaining the resources in its original location exceeds the cost of reallocation (Kirigia et al, 2000).

6.3 SUGGESTIONS FOR FURTHER RESEARCH

This study focused on only one dimension of efficiency, that is, technical efficiency which is not the ultimate goal of most organizations. The ultimate aspiration of any organization is to achieve economic efficiency which is the aggregate of technical efficiency and allocative efficiency. In order for studies centered on total efficiency to be conducted in Zimbabwe in the future, more emphasis should be laid on collecting information on the quantities of all the main outputs and inputs and the average or median prices per unit of each input from all public and private health facilities, health centers and hospitals. This would facilitate measurement of total economic efficiency.

Furthermore, in order to aid monitoring and evaluation of the effects of different health care reforms on the efficiency of individual health care facilities over time, a *Malmquist* Productivity Index analysis is necessary. The Malmquist Productivity Index helps to measure explicitly total factor productivity. It decomposes productivity growth into efficiency change and technical change. It would be necessary to collect data for a year (or more) before the introduction of specific reforms, and for subsequent years.

There is also need to determine the quality variations and their effect on the efficiency scores. This dimension of output must be held constant because provision of greater quality care requires additional inputs per unit of output .However the hospitals included in our study were fairly homogeneous in size and mix of services provided. This is important because hospitals offering higher quality of care may require more personnel time and other inputs than those offering low quality of care to the extent that higher quality hospitals may have lower efficiency scores not because they are less efficient but because they provide better services to their patients.

The study has demonstrated that DEA not only helps health policymakers and managers to answer the question "How well are the hospitals performing?" but also "By how much could their performance be improved?" We therefore recommend that a further analysis of the hospitals which are performing best and their operating practices be undertaken with a view to establishing a guide to "best practice" for others to emulate. Hence, it is recommended that the causes of the inefficiencies be unpacked and necessary efficiency measures be instituted to augment government's efforts to address problems of health care provision in the country.

In a nutshell, we found that on all the set of hospitals considered for-profit hospitals are the most performing in terms of efficiency. Perhaps fruitful research may then attempt to determine why this is the case: why are public and mission hospitals apparently not efficient than for-profit hospitals? It is important to note that hospitals are instrumental in efforts to address cost-effective interventions aimed at achieving the current United Nations Millennium Development Goals (MDGs) scheduled for 2015.

6.4 CONCLUSION

The major driving force in this study has been to investigate whether profit incentives in the provision of hospital care has a significant impact on efficiency. The evidence in this study revealed that private for-profit hospitals are more efficient than non-profit hospitals, both public and mission. A further comparison of performance within the non-profit hospital sector revealed that the mission hospitals are better performing than public hospitals. These findings are consistent with the main hypothesis outlined in this study namely that the for-profit hospital sector performs relatively efficient than the non-profit hospital sector as implied by the property rights theory. Therefore the form of hospital ownership has a significant bearing on the performance of hospitals. Future research is therefore necessary to determine why this is the case: why are public and mission hospitals apparently not efficient than for-profit hospitals? Perhaps factors such as pay structure differences, administration and many others would need to be considered.

REFERENCES

Aigner, D.J., and Chu, S. F., 1968. Estimating the Industry Production Function. *American Economic Review*, 58(4), pp.259-269.

Akin, J.N., Birdsall, N., and deFerranti., 1987. Financing Health Services in Developing Countries: An Agenda for Reform. Washington DC: The World Bank.

Banker, R.D., Charnes, A., and Cooper, W.W. 1984. Some Models for Estimating Technical and Scale inefficiencies in Data Envelopment Analysis. *Management Science*, 30, pp. 1078-1092.

Banker, R.D., Conrad R.F and Strauss, R.P., 1986. A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production. *Management Science*, 32(1), pp. 30 - 44.

Bartlett, W., and Le Grand, J., 1993. The Theory of Quasi-Markets. In Le Grand J and Bartlett W (eds) Quasi-Markets And Social Policy. Macmillan, Basingstoke.

Bowlin, W.F., Charnes A., Cooper W.W., and Sherman H.D., 1985. Data Envelopment Analysis and Regression Approaches to Efficiency Estimation and Evaluation. *Annals of Operations Research*, 2, pp. 113–138.

Bowlin,W.F., 1987. Measuring Performance; An introduction to Data Envelopment Analysis. Department of Accounting, University of Northern Iowa, Cedar Falls.

Broomberg, J., 1994. Managing the Health Care Market in Developing Countries. Departmental Publication No 11, Health Policy Unit, London School of Hygiene and Tropical Medicine.

Burgess, J., and Wilson P., 1996. Hospital Ownership and Technical Inefficiency, *Management Science*, 42(1) pp. 110–123.

Central Statistical Office. 1997. Quarterly Digest of Statistics, March 1997.

Central Statistical Office, 2008. Harare, Government of Zimbabwe.

Central Statistical Office, National Accounts, 2004 – 2006.

Charnes A., Cooper W., Lewin A.Y., and Seiford L.M., 1994. Data Envelopment Analysis: Theory, Methodology and Applications Kluwer Academic: Boston.

Charnes, A., Cooper.W., and Rhodes. E., 1978. Measuring the Efficiency of Decision- Making Units. *European Journal of Operational Research*, 2, pp.429-444.

Coehlo, C.C.S., 2007. *Modes of Governance and Public Service Efficiency*. Ph.D. Birmingham: University of Birmingham.

Coelli, T. J. 1996. A guide to DEAP version 2.1: A Data Envelopment Analysis (Computer) Program. [Online],
Available at : http://www.une.edu.au/econometrics/cepawp.html [20 February 2008].

Coelli, T. J. 1996. A Guide to DEAP version 2.1; A Data Envelopment Analysis (Computer) Program. [Online] CEPA Working Paper 1996 / 98, Department of Econometrics, University of New England. Available at: http://www.une.edu.au/econometrics/cepawp.html [20 February 2009].

Coyne, J., 1982. Hospital Performance in Multi-hospital systems; A Comparative Study of System and Independent Hospitals. *Health Service Research*, pp. 303 – 329.

Culter, D.M., 2002. Health Care and the Public Sector. In Auerbach, A.J., and Feldstein, M., (eds), *Handbook of Public Economics*, 4, Elsevier: pp. 2143-2243.

Cutler, D.M., and Horwitz, J., 1999. Converting hospitals from not-for-profit to for-profit form: why and what effects. In Cutler, D., ed., *The Changing Hospital Industry: Comparing Not-for-Profit and for-Profit Institutions*, Chicago: University of Chicago Press, pp. 45-79.

Dalmau-Mataradona, E., and Puig-Jonoy, J., 1990. Market Structure and Efficiency; Evaluating Potential Effects of Deregulation in a National Health Service. Department of Economics, University of Pompeu Fabra.

Dawson, D., Gravelle, H., Kind, P., O'Mahony, M., Street, A., and Weale, M., 2004. Developing New Approaches to Measuring NHS Outputs and Productivity – First Interim Report to Department of Health, CHE Technical Paper Series 31.

Diewert, E., Caves, D., and Christensen, L., 1982. The Economic Theory of Index Numbers and the Measurement of Input, Output And Productivity. *Econometrica*, 50(6), pp. 1393-1414.

Diewert, W.E., 1971. An Application of the Shepherd Duality Theorem: A Generalized Leontief Production Function. *Journal of Political Economy*, pp. 481-507.

Ellis R., 1992. Hospital Cost Function Estimation When Firms may not try to Minimize Total Costs. Manuscript. Boston, Massachusetts: Boston University.

Ellis, R.P., and Chawla, M., 1993. Public and Private interactions in the Health Sector in Developing Countries. Phase 1. Review of concepts and literature, and preliminary fieldwork design. Health Financing and Sustainability Project, Bethesda, Maryland.

Ellis, R. P., 1998. Creaming, Dumping, skimping: Provider competition on the intensive and extensive margins, *Journal of Health Economics*, vol 17: pp 537-555.

Emrouznejad, A., 2007. DEA Tutorial. [Online], Available: http://www.deazone.com/tutorial/graphical.html [27 September 2008].

Evans, D., Tandon, A., Murray, C., and Lauer, J., 2000. The Comparative Efficiency of National Health Systems in Producing Health: an Analysis of 191 Countries, GPE Discussion Paper Series 29, World Health Organization, Geneva.

Fand, S., Goodman A.C., and Stano M., eds., 2001. The Economics of Health and Health Care, Prentice Hall: New Jersey.

Farrell, M., and Fieldhouse, M., 1962. Estimating Efficient Production Functions under Increasing Returns to Scale. *Journal of Royal Statistical Association*, pp. 252-67.

Farrell, M., 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*, Series A, 120, Part 3, pp. 253-290.

Ferrier, G.D., and Valdmanis, V., 1996. Rural Hospital Performance and its Correlates. *The Journal of Productivity Analysis*,7, pp. 63-80.

Fried, H.O., Lovell, C.A.K., Schmidt, S.S., and Yaisawarng, S., 2002. Accounting For Environmental Effects and Statistical Noise in, Data Envelopment Analysis. *Journal of Productivity Analysis*, 17(1-2), pp. 157-174.

Ganley, J., and Cubbin, J., 1992. Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis, New York: Elsevier Science Publishers.

Gonzalez, L., and Perez, B., (1996), Evaluating Performance among Spanish Hospitals; A Comparison of DEA and SFA. *Journal of Health Economics*, 14, pp. 257-289.

Grosskopf, S and V. Valdmanis 1987. Measuring Hospital Performance: A Non-Parametric Approach. *Journal of Health Economics*, 6, pp. 89-107.
Available at: http://www.hsph.harvard.edu/takemi/pp132.pdf. [Accessed 9 March 2009].

Grossman M., 1972. The Demand for Health: A Theoretical and Empirical 1nvestigation. New York: National Bureau of Economic Research.

Harvey, L., 1978. General X-Efficiency Theory and Economic Development, New York: Harvard University Press.

Hofler, R. A., and Folland S.T., 1991. Technical and Allocative Efficiencies of United State Hospitals under a Stochastic Frontier Approach presented at the Midwest Economics Association, Fifty-fifth Annual Meeting (St Louis, MO).

John, L., Fizel. T., and Nunnukhoven S., 1992. Technical Efficiency of For-Profit and Non-Profit Nursing Homes. *Managerial and Decision Economics,* 13, pp 429 – 439.

Kirigia J.M., Emrouznejad A., Sambo L.G., Munguti N., and Liambila W., 2004. Using Data Envelopment Analysis to Measure the Technical Efficiency of Public Health Centers in Kenya. *Journal of Medical Systems*, 28, pp. 155-166.

Kooreman, P., 1994. Nursing home care in The Netherlands: A nonparametric efficiency analysis. *Journal of Health Economics*, 13, pp. 301-16.

Le Grand, J., and Bartlett, W. Eds., 1993. Quasi-Markets and Social Policy. Macmillan, Basingstoke.

Lewin, L.S., Derzon, R.A., and Margulies, R., 1981. Investor-Owned and Nonprofits Differ in Economic Performance. Hospitals, 55(13): pp. 52-58.

Linna, Miika, and Hakkinen, 1996. *Determinants of Cost Efficiency of Finnish Hospitals; A Comparison of DEA and SFA*. Natal Research and Development Centre for Welfare and Health, Helsinki: Finland.

Liu, X., and Mills, 1998. *Performance - related pay to hospital - based doctors: A case study in Korea*." Available at: http://www.hsph.harvard.edu/takemi/pp132.pdf. [Accessed 6 July 2008].

Mangwiro, C., 2001. *Public Sector Health Efficiency in Zimbabwe*. MSc. Department of Economics, Zimbabwe: University of Zimbabwe.

Mbalame, E.D., 2002. *Technical efficiency in the public health sector. The case of Mozambique*. MSc. Department of Economics, Zimbabwe: University of Zimbabwe.

Ministry of Health and Child Welfare, 1998. *National Health Strategy for Zimbabwe (1997 – 2007).*Working for Quality and Equity in Health. Government Printers.

National Budget Statement, 2008. Ministry of Finance. Zimbabwe.

National Health Accounts Draft (NHA), 1999. Ministry of Health and Child Welfare.

National Health Survey (1995), Ministry of Health and Child Welfare.

Newbrander, W., Barnum H., and Kutzin J., 1992. *Hospital Economics and Financing in Developing Countries*. World Health Organization, Geneva.

Nyman, J.A., and Bricker D., 1989. Profit Incentives and Technical Efficiency in the Production of Nursery Home Care. *The review of Economics and Statistics*, 71, pp. 586 – 594.

O'Mahony, M., and Stevens, P., 2004. International Comparisons of Performance in the Provision of Public Services: outcome based measures for education. *National Institute of Economic and Social Research*, London.

Osei D., d'Almeida S., George M.O., Kirigia J.M., Mensah A.O., Kainyu L.H., 2005. Technical Efficiency of Public District Hospitals and Health Centers in Ghana: A Pilot Study. Cost Efficiency Resource Allocation.

Owino, P. W., and Korrir J., 1997. *Public Health Sector Efficiency in Kenya: Estimation and Policy Implications*. IPAR Discussion Paper.

Ozcan, Y.A., 1992. Sensitivity analysis of hospital efficiency under alternative output/input and Peer groups: a review. *The International Journal of Knowledge Transfer and Utilisation*, 5(4): pp.1-29.

Ozcan, Y.A., 1995. Efficiency of Hospital-Service Production In Local Markets - The Balance-Sheet Of US Medical Armament. *Socio-Economic Planning Sciences*, 29(2), pp.139-50.

Ozcan, Y.A., Luke, R.D., and Haksever, C., 1992. Ownership and Organizational Performance: a comparison of technical efficiency across hospital types. *Medical Care Research and Review*, 30(9), pp. 781-94.

Ozcan, Y.A., McCue, M.J., and Okasha, A.A., 1996. Measuring the Technical Efficiency of Psychiatric Hospitals. *Journal of Medical Systems*, 20(3), pp.141-50.

Preker, A.S., 2007. Public Ends, Private Means: Strategic Purchasing of Health Services. Washington DC: The World Bank.

Pritchett, L., and Summers, L.H., 1996. Wealthier is Healthier. *Journal of Human Resources*, 31(4), pp.841-868.

Puig-Junoy J. (2000), ‘Partitioning Input Cost Efficiency into its Allocative and Technical Components: An Empirical DEA Application to Hospitals’, *Socio-Economic Planning Sciences*, 34, pp. 199-218.

Register, A., and Bruning E., 1987. Profit incentives and technical efficiency in the production of Hospital care. *Southern Economic Journal*, 53, pp.899 – 914.

Register, A., Sharp. A., and Bivin D., 1985. Profit Incentives and The Hospital Industry; Are We Expecting Too Much? *Health Services Research*, pp. 225 – 241.

Sahn, D., and Bernier R., 1993. Has Structural Adjustment Led to Health Sector Reform in Africa?

Sengupta, J.K., 1990. Transformations in Stochastic DEA models. *Journal of Economics*, 46, pp. 109-123.

Sherman, H.D., 1984. Hospital Efficiency Measurement and Evaluation. Empirical Test of a New Technique. *Medical Care Research and Review*, 22(10): pp. 922-38.

Smith, P**.,** and Street, A., 2005. Measuring the Efficiency of Public Services: The Limits of Analysis. *Journal of the Royal Statistical Society*, Series A, 168, pp. 401-417.

Thanassoulis, E., 1993. A Comparison of Regression Analysis and Data Envelopment Analysis as Alternative Methods for Performance Assessments. *Journal of the Operational Research Society*, 44, pp. 1129-1144.

Thanassoulis, E., 2001. Introduction to the Theory and Application of Data Envelopment Analysis. Dordrecht: Kluwer.

Valdmanis, V.G., 1990. Ownership and Technical Efficiency of Hospitals. *Medical Care Research and Review*, 28(6): pp. 552- 61.

Valdmanis, V.G, 1992. Sensitivity analysis for DEA models: An empirical example using public versus not-for-profit hospitals. *Journal of Public Economics*, 48, pp. 185–205.

Vitaliano, D.F., and Toren M., 1993. Frontier analysis: Measuring efficiency of for-profit and not-for-profit hospitals in USA. *Journal of Health Economics*, 13, pp. 341–343.

Vitaliano, D.F., and Toren, M., 1996. Hospital Cost and Efficiency in a Regime of Stringent Regulation. *Eastern Economic Journal*, 22(2): pp.161-75.

WHO., 2000. *The World Health Report – Health systems: improving performance*, World Health Organization, Geneva.

Wilson, G.W., and Jadlow J.M., 1982. Competition, Profit Incentives, and Technical Efficiency in the Provision of Nuclear Medicine Services. *Bell Journal of Economics*, 13(2), pp. 472-82.

Wilson, G.W., and. Jadlow, J.M., (1992), "Competition, profit incentives and technical efficiency in the provision of nuclear medicine services. *The Bell Journal of Economics*, 13, pp.472 – 482.

World Bank .1990. World Development Report. Oxford University Press, New York.

Zere, E., 2000. Hospital efficiency in Sub-Saharan: Evidence from South Africa.

Zere E., McIntyre D., and Addison T., 2001. Technical Efficiency and Productivity Of Public Sector Hospitals in Three South African Provinces. *South African Journal of Economics*, Vol. 69:2.

Zimbabwe Association of Church-related Hospitals (1998-2004).

Zimbabwe Health for All Action Plan, 1986. Ministry of Health.

Zimbabwe National Health Profile, 1998. Ministry of Health and Child Welfare.

Zimbabwe National Health Profile, 2002. Ministry of Health and Child Welfare.

Zimbabwe National Health Profile, 2007. Ministry of Health and Child Welfare.

APPENDICES

Appendix 1: Summary Data for All the Hospitals

Type	Total	Discharges	Inpatient days	Beds	Doctors	Nurses	BOR	ALOS
MISSION1	H1	4565	19499	80	2	16	66.78	4.27
2	H2	1379	6620	160	3	26	11.34	4.80
3	H3	1113	5446	58	1	6	25.73	4.89
4	H4	2258	5260	65	2	13	22.17	2.33
5	H5	863	5830	120	2	19	13.31	6.76
6	H6	778	4340	54	1	7	22.02	5.58
7	H7	6529	30728	144	4	52	58.46	4.71
8	H8	1461	4449	78	6	40	15.63	3.05
9	H9	861	29823	222	3	29	36.80	34.64
10	H10	2234	9954	106	1	1	25.73	4.46
11	H11	1603	8565	122	3	22	19.23	5.34
12	H12	1718	4715	101	2	12	12.79	2.74
13	H13	2423	8779	125	3	45	19.24	3.62
14	H14	2027	10681	125	4	20	23.41	5.27
15	H15	906	10901	80	2	11	37.33	12.03
16	H16	2633	17061	218	5	59	21.44	6.48
17	H17	4043	2343	240	5	51	2.67	0.58
18	H18	909	6575	175	4	31	10.29	7.23
19	H19	1679	8868	120	3	19	20.25	5.28
20	H20	4084	34241	218	3	24	43.03	8.38
21	H21	3247	13388	121	3	25	30.31	4.12
22	H22	788	3343	128	2	13	7.16	4.24
23	H23	2563	20151	190	2	54	29.06	7.86
24	H24	701	5041	70	3	21	19.73	7.19
25	H25	64	215	50	2	8	1.18	3.36
26	H26	4003	13521	115	3	30	32.21	3.38
27	H27	3479	17491	200	4	48	23.96	5.03
28	H28	3830	38421	150	4	22	70.18	10.03
29	H29	2302	17842	180	4	36	27.16	7.75
30	H30	713	2483	67	1	7	10.15	3.48
31	H31	3439	20550	241	5	61	23.36	5.98
32	H32	1416	9676	88	5	66	30.12	6.83

33	H33	726	5380	124	3	21	11.89	7.41
34	H34	169	842	57	2	15	4.05	4.98
35	H35	3326	31105	188	4	44	45.33	9.35
36	H36	2635	15567	103	3	31	41.41	5.91
PUBLIC 1	H37	32533	146810	988	96	754	40.71	4.51
2	H38	40738	154727	1088	91	852	38.96	3.80
3	H39	23067	86827	388	32	354	61.31	3.76
4	H40	15663	75442	210	46	834	98.42	4.82
5	H41	14548	23353	466	53	777	13.73	1.61
6	H42	17737	21823	443	48	495	13.50	1.23
7	H43	19097	142625	1038	81	586	37.64	7.47
8	H44	14868	199038	660	81	525	82.62	13.39
9	H45	15246	53953	790	21	289	18.71	3.54
10	H46	11300	35278	345	10	103	28.02	3.12
11	H47	18842	23168	213	6	137	29.80	1.23
12	H48	16739	23155	331	4	195	19.17	1.38
13	H49	10890	17963	267	7	124	18.43	1.65
14	H50	11080	14267	189	5	132	20.68	1.29
15	H51	12300	19748	212	6	136	25.52	1.61
16	H52	13420	24167	312	7	254	21.22	1.80
17	H53	11970	21567	156	5	137	37.88	1.80
18	H54	15300	25312	211	6	175	32.87	1.65
19	H55	13200	18679	198	4	173	25.85	1.42
20	H56	18900	25267	265	6	173	26.12	1.34
21	H57	15780	24126	313	4	151	21.12	1.53
22	H58	11800	31241	412	6	174	20.77	2.65
23	H59	12700	21495	134	4	142	43.95	1.69
24	H60	17900	26954	111	7	97	66.53	1.51
25	H61	10890	17664	123	4	113	39.35	1.62
26	H62	13708	18896	136	7	116	38.07	1.38
27	H63	10700	19985	157	4	101	34.87	1.87
28	H64	9066	13632	115	7	99	32.48	1.50
29	H65	10090	15276	98	4	75	42.71	1.51
30	H66	9800	13242	87	7	89	41.70	1.35
31	H67	10598	13400	79	4	79	46.47	1.26
32	H68	10299	15367	119	3	111	35.38	1.49
33	H69	10083	13460	78	4	65	47.28	1.33
34	H70	9883	12100	99	3	77	33.49	1.22

35	H71	11299	13426	141	2	121	26.09	1.19
36	H72	9011	11215	94	2	96	32.69	1.24
37	H73	9200	13249	87	2	78	41.72	1.44
38	H74	12189	18955	142	4	113	36.57	1.56
39	H75	13199	24100	146	3	123	45.22	1.83
40	H76	13129	27400	158	3	135	47.51	2.09
41	H77	14610	25746	172	4	146	41.01	1.76
42	H78	14790	25447	152	3	113	45.87	1.72
43	H79	11800	16546	113	2	98	40.12	1.40
44	H80	98277	12343	98	3	76	34.51	0.13
FOR-PROFIT 1	H81	4678	23518	215	2	23	29.97	5.03
2	H82	3312	45161	156	3	13	79.31	13.64
3	H83	2971	29716	123	4	16	66.19	10.00
4	H84	2986	47618	208	1	18	62.72	15.95
5	H85	4351	39274	225	2	21	47.82	9.03
6	H86	2129	19810	112	2	14	48.46	9.30
7	H87	2876	18973	195	3	26	26.66	6.60
8	H88	2100	21739	135	1	29	44.12	10.35
9	H89	2015	31237	198	2	16	43.22	15.50
10	H90	4120	19289	228	3	23	23.18	4.68
11	H91	4321	26752	238	2	13	30.80	6.19
12	H92	3893	27173	184	3	16	40.46	6.98
13	H93	2765	26781	177	2	25	41.45	9.69
14	H94	4942	25208	213	4	27	32.42	5.10
15	H95	3813	31278	188	3	17	45.58	8.20
16	H96	4310	23410	145	2	14	44.23	5.43
17	H97	2419	26718	157	5	18	46.62	11.05
18	H98	2519	18791	164	2	25	31.39	7.46
19	H99	2615	29871	172	3	23	47.58	11.42
20	H100	3721	21082	183	4	21	31.56	5.67
TOTAL	**100**	**862562**	**2601527**	**20433**	**873**	**11196**	**3405.28**	**505.30**

Appendix 2: Results from DEAP Version 2.1

TYPE	HOSPITAL	CRSTE	VRSTE	SCALE	RETURNS
Mission	H1	0.859	1.000	0.859	IRS
M	H2	0.148	0.369	0.401	IRS
M	H3	0.354	1.000	0.354	IRS
M	H4	0.295	0.861	0.342	IRS
M	H5	0.177	0.500	0.354	IRS
M	H6	0.287	1.000	0.287	IRS
M	H7	0.746	0.789	0.945	IRS
M	H8	0.200	0.697	0.287	IRS
M	H9	0.506	0.566	0.894	IRS
M	H10	1.000	1.000	1.000	CRS
M	H11	0.246	0.502	0.490	IRS
M	H12	0.191	0.566	0.337	IRS
M	H13	0.249	0.494	0.504	IRS
M	H14	0.300	0.522	0.574	IRS
M	H15	0.470	0.844	0.556	IRS
M	H16	0.273	0.356	0.767	IRS
M	H17	0.065	0.230	0.284	IRS
M	H18	0.131	0.330	0.395	IRS
M	H19	0.259	0.515	0.504	IRS
M	H20	0.596	0.638	0.934	IRS
M	H21	0.390	0.575	0.679	IRS
M	H22	0.106	0.500	0.212	IRS
M	H23	0.425	0.564	0.753	IRS
M	H24	0.248	0.789	0.314	IRS
M	H25	0.015	1.000	0.015	IRS
M	H26	0.418	0.609	0.686	IRS
M	H27	0.308	0.413	0.746	IRS
M	H28	0.885	0.905	0.978	IRS
M	H29	0.344	0.446	0.773	IRS
M	H30	0.158	1.000	0.158	IRS
M	H31	0.299	0.367	0.814	IRS
M	H32	0.368	0.718	0.512	IRS
M	H33	0.151	0.450	0.335	IRS
M	H34	0.052	0.892	0.058	IRS
M	H35	0.573	0.610	0.940	IRS
M	H36	0.524	0.711	0.736	IRS
Public	H37	0.503	0.604	0.833	DRS
P	H38	0.484	1.000	0.484	DRS
P	H39	0.756	0.811	0.932	DRS
P	H40	1.000	1.000	1.000	CRS
P	H41	0.171	0.198	0.864	IRS

P	H42	0.180	0.203	0.884	IRS
P	H43	0.464	0.655	0.709	DRS
P	H44	0.998	1.000	0.998	DRS
P	H45	0.242	0.325	0.744	DRS
P	H46	0.363	0.374	0.973	IRS
P	H47	0.418	0.453	0.921	IRS
P	H48	0.307	0.360	0.853	IRS
P	H49	0.251	0.298	0.843	IRS
P	H50	0.289	0.389	0.743	IRS
P	H51	0.347	0.395	0.880	IRS
P	H52	0.288	0.318	0.905	IRS
P	H53	0.508	0.568	0.894	IRS
P	H54	0.446	0.480	0.929	IRS
P	H55	0.365	0.453	0.806	IRS
P	H56	0.365	0.394	0.926	IRS
P	H57	0.330	0.381	0.866	IRS
P	H58	0.299	0.320	0.935	IRS
P	H59	0.594	0.663	0.895	IRS
P	H60	0.879	0.941	0.935	IRS
P	H61	0.532	0.642	0.828	IRS
P	H62	0.511	0.603	0.847	IRS
P	H63	0.470	0.535	0.879	IRS
P	H64	0.430	0.625	0.688	IRS
P	H65	0.575	0.764	0.753	IRS
P	H66	0.552	0.821	0.673	IRS
P	H67	0.630	0.911	0.692	IRS
P	H68	0.487	0.631	0.772	IRS
P	H69	0.637	0.921	0.691	IRS
P	H70	0.468	0.703	0.666	IRS
P	H71	0.413	0.617	0.670	IRS
P	H72	0.462	0.743	0.622	IRS
P	H73	0.578	0.828	0.699	IRS
P	H74	0.500	0.574	0.870	IRS
P	H75	0.616	0.697	0.883	IRS
P	H76	0.653	0.716	0.912	IRS
P	H77	0.557	0.598	0.932	IRS
P	H78	0.636	0.707	0.900	IRS
P	H79	0.588	0.782	0.751	IRS
P	H80	1.000	1.000	1.000	CRS
For-profit	H81	0.456	0.563	0.810	IRS
F	H82	1.000	1.000	1.000	CRS
F	H83	0.836	0.902	0.927	IRS
F	H84	1.000	1.000	1.000	CRS
F	H85	0.716	0.750	0.955	IRS

F	H86	0.628	0.827	0.759	IRS
F	H87	0.362	0.481	0.753	IRS
F	H88	0.676	1.000	0.676	IRS
F	H89	0.652	0.726	0.898	IRS
F	H90	0.349	0.452	0.772	IRS
F	H91	0.675	0.704	0.959	DRS
F	H92	0.577	0.651	0.886	IRS
F	H93	0.595	0.695	0.856	IRS
F	H94	0.422	0.498	0.848	IRS
F	H95	0.622	0.688	0.904	IRS
F	H96	0.653	0.784	0.833	IRS
F	H97	0.588	0.678	0.868	IRS
F	H98	0.447	0.607	0.736	IRS
F	H99	0.618	0.686	0.900	IRS
F	H100	0.404	0.518	0.780	IRS
F	**Mean**	**47.034**	**64.539**	**73.377**	

Appendix 3: Regression results from Stata

reg crste bed bed2 bor alos dprof dgovt beddprof beddgovt

Source	SS	df	MS		
				Number of obs	= 100
				F(8, 91)	= 62.72
Model	4.48609733	8	.560762167	Prob > F	= 0.0000
Residual	.813619156	91	.00894087	R-squared	= 0.8465
				Adj R-squared	= 0.8330
Total	5.29971649	99	.05353249	Root MSE	= .09456

crste	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
bed	4.47e-06	.0003006	0.01	0.988	-.0005927	.0006016
bed2	1.43e-07	2.08e-07	0.69	0.494	-2.70e-07	5.55e-07
bor	.011636	.0006832	17.03	0.000	.010279	.0129931
alos	-.0023198	.0028762	-0.81	0.422	-.008033	.0033933
dprof	.1978798	.1097032	1.80	0.075	-.0200321	.4157917
dgovt	.0353602	.0566499	0.62	0.534	-.0771679	.1478882
beddprof	-.0006541	.0006203	-1.05	0.294	-.0018862	.000578
beddgovt	-.0002074	.0003357	-0.62	0.538	-.0008742	.0004595
_cons	.0664079	.0423182	1.57	0.120	-.0176521	.1504678

Appendix 4: Regression results from Stata after dropping insignificant variables.

```
reg crste bor alos dprof

    Source |       SS       df       MS              Number of obs =     100
-------------+------------------------------           F(  3,    96)       =    170.75
     Model |  4.46327009     3   1.4877567           Prob > F          =    0.0000
  Residual |  .836446404    96  .008712983           R-squared         =    0.8422
-------------+------------------------------           Adj R-squared   =    0.8372
     Total |  5.29971649    99   .05353249           Root MSE          =    .09334
------------------------------------------------------------------------------
     crste |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       bor |   .0117925   .0005738    20.55   0.000     .0106535    .0129316
      alos |  -.0026636   .0022409    -1.19   0.238    -.0071117    .0017846
     dprof |   .0861293   .0251514     3.42   0.001     .0362041    .1360545
     _cons |   .0641439   .0213966     3.00   0.003      .021672    .1066159
```